Accepted Manuscript

This is the peer reviewed version of the following article:

Schneider, S. C., Rodrigues, A. , Moe, T. F., Ballot, A. and De Clerck, O. (2015),
DNA barcoding the genus Chara: molecular evidence recovers fewer taxa
than the classical morphological approach. J. Phycol., 51: 367-380,

which has been published in final form at https://doi.org/10.1111/jpy.12282.

This article may be used for non-commercial purposes in accordance with
Wiley Terms and Conditions for Use of Self-Archived Versions.

It is recommended to use the published version for citation.

1 **DNA barcoding the genus *Chara*: molecular evidence recovers fewer taxa than the**
2 **classical morphological approach**

3

4 Susanne C. Schneider[1*], Anuar Rodrigues[2], Therese Fosholt Moe[1], Andreas Ballot[1]

5

6

7 * corresponding author: susi.schneider@niva.no
8 1 Norwegian Institute for Water Research, Gaustadalleen 21, 0349 Oslo, Norway
9 2 Biodiversity Institute of Ontario, University of Guelph, 50 Stone Road East, Guelph, ON,
10 Canada

11

12

13 **Running title**
14 DNA barcoding the genus *Chara*

15

16

17

18 **Keywords**
19 algae, barcode, Charales, charophyte, ITS, matK, plant, rbcL, taxonomy

20

21

22

23 **Abbreviations**

24	AFLP	Amplified Fragment Length Polymorphism
25	CCDB	Canadian Centre for DNA Barcoding
26	ITS2	nuclear ribosomal internal transcribed spacer region 2
27	*matK*	plastid-encoded protein-coding gene
28	*rbc*L	plastid-encoded gene coding for the large subunit of RUBISCO

29

30 **Abstract**

31 Charophytes (Charales) are benthic algae with a complex morphology. They are vulnerable to
32 ecosystem changes, such as eutrophication, and are red-listed in many countries. Accurate
33 identification of *Chara* species is critical for understanding their diversity and for
34 documenting changes in species distribution. Species delineation is, however, complicated,
35 because of high phenotypic plasticity. We used barcodes of the ITS2, *matK* and *rbc*L regions
36 to test if the distribution of barcode haplotypes among individuals is consistent with species
37 boundaries as they are currently understood. The study included freshly collected and
38 herbarium material of 91 specimens from 10 European countries, Canada and Argentina.

39 Results show that herbarium specimens also for aquatic plants like *Chara* are useful as a
40 source of material for genetic analyses. *rbc*L and *matK* had highest sequence recoverability,
41 but *rbc*L had a somewhat lower discriminatory power than ITS2 and *matK*. The tree resulting
42 from the concatenated data matrix grouped the samples that according to a traditional
43 morphological approach consisted of 12 different taxa, into six main groups. A large
44 unresolved group consisted of *C. intermedia, C. hispida, C. horrida, C. baltica, C.*
45 *polyacantha, C. rudis, C. aculeolata,* and *C. corfuensis*. A second unresolved group consisted
46 of *C. virgata* and *C. strigosa*. The taxa within each of the unresolved groups shared identical
47 barcode sequences on the 977 positions of the concatenated data matrix. The morphological
48 differences of taxa within both unresolved groups include the number and length of spine
49 cells, stipulodes, and bract cells. We suggest that these morphological traits have less
50 taxonomic relevance than hitherto assumed.

51

~~1.~~ Introduction

Charophytes, extant and fossil members of the order Charales plus the members of the extinct orders Sycidiales and Moellerinales (Schneider et al.~~,~~ 2015a), are algae with a complex morphology, which are closely related to modern land plants (Timme et al.~~,~~ 2012). It was long believed that Charales were the closest living relatives to land plants, i.e. that the tremendous diversity of land plants all descended from a single charophyte alga (Karol et al.~~,~~ 2001). For this reason, charophytes received much attention in DNA studies. However, recent phylogenetic analyses instead support the Zygnematales as the direct ancestors of land plants (Wodniok et al.~~,~~ 2011; Timme et al.~~,~~ 2012).

Charophytes help maintain oligotrophic conditions in ecosystems by directing nutrients and carbon from the water to the sediment beneath charophyte meadows (Kufel et al.~~,~~ 2013). On the other hand, charophytes are also sensitive to environmental changes such as eutrophication (Blindow~~,~~ 1992). Consequently, many charophytes have become rare or even endangered in recent decades (Baastrup-Spohr et al.~~,~~ 2013) and further changes are predicted in a changing climate (Auderset and Rey-Boissezon~~,~~ 2015). Accurate identification of charophyte species is, however, critical for understanding their diversity and for documenting changes in species distribution. *Chara* species are also used as bioindicators of eutrophication and in the determination of ecological status according to the Water Framework Directive, both in streams (Schneider & Melzer~~,~~ 2003) and lakes (Stelzer et al.~~,~~ 2005), such that accurate species identification is of high practical relevance for ecosystem assessment and management.

Species delineation of charophytes is commonly based on morphological traits of the plant thallus. This is, however, more complicated than it might seem, because 1) there is considerable overlap in morphological characteristics used to discriminate species (Boegle et al.~~,~~ 2007), and 2) phenotypic plasticity in charophytes may be environmentally induced, e.g. by light, water temperature, nutrient concentrations and salinity (Wood and Imahori~~,~~ 1965; Schneider et al.~~,~~ 2015b). Such plasticity makes it difficult to know which morphotypes are environmentally induced and which ones are genetically controlled (Boegle et al.~~,~~ 2010a). For example, Boegle et al. (2010b) showed that *Chara baltica* Bruzelius and *C. horrida* Wahlstedt cannot be separated from each other with the genetic fingerprinting technique AFLP (Amplified Fragment Length Polymorphism), in spite of pronounced differences in the number and length of spine cells and stipulodes. The number and length of spine cells and

stipulodes are, however, two of the most important characters for traditional species delineation within the genus *Chara* (Wood and Imahori 1965; Krause 1997), such that results from morphological and genetic analyses conflict with each other.

DNA barcoding, i.e. the use of short regions of DNA to identify species by assigning individuals to known taxa through comparison of their barcodes with a reference library, has become a popular means to improve species identification (Saarela et al. 2013). The Consortium for the Barcode of Life (CBOL) Plant Working Group (2009) recommended *rbc*L + *matK* as the standard plant barcode, after a broad study of several candidate regions. The *rbc*L gene is plastid-encoded, and its function is to code for the large subunit of ribulose 1, 5 bisphosphate carboxylase/oxygenase (RUBISCO). *MatK* is a plastid-encoded protein-coding gene, that has been shown to reside within a group II intron of the plastid encoded tRNALys (lysine tRNA). The maturase *matK* presumably helps in splicing of multiple introns (Vogel et al. 1999). Although the nuclear encoded internal transcribed spacer regions (ITS) were dismissed as standard barcodes due to complicating problems, the importance of ITS as a supplemental marker to *rbc*L and *matK* was stressed by several working groups (China Plant BOL Group 2011; Kuzmina et al. 2012). We therefore additionally analysed ITS2 in an attempt to achieve higher discriminatory power. ITS is a region of non-coding DNA situated between structural ribosomal RNA genes.

Herbarium collections are potentially an excellent resource for providing material that can be used for DNA studies. It can, however, be difficult to obtain DNA of good enough quality from herbarium material that will result in the amplification of various genes/markers. For land plants it has been shown that the drying method strongly affects PCR success (Sarkinen et al. 2012). We included both fresh and herbarium material of different ages in this study because we expected the drying of samples may be even more critical for hydrophytes, such as *Chara*, but on the other hand the use of *Chara* herbarium collections would greatly facilitate the analysis of a large number of samples.

The objectives of our study were to 1) design ITS2 and *matK* primers for the genus *Chara*, 2) solve taxonomic problems, 3) test if herbarium samples are suitable as source of genetic material for *Chara*, and if so, to also test for relationships between the age of herbarium samples and their sequence recovery for ITS2, *matK* and *rbc*L, 4) test if results obtained from ITS2, *matK* and *rbc*L are consistent with each other, or if one marker discriminates better than

115 the others, and 5) test if the distribution of barcode haplotypes among individuals is consistent
116 with species boundaries as they are currently understood.

117

118 **~~2.~~ Material and Methods**

119 ***Taxon sampling***

120 The study included 91 samples from 10 European countries, Canada and Argentina (Table 1).
121 Eleven individuals were collected fresh, and dried in silica gel shortly after sampling; 29
122 individuals were collected from private herbaria that are stored at the Norwegian Institute for
123 Water Research (assembled by Michael Boegle and Susanne Schneider), and 51 individuals
124 were collected from the herbarium of the Natural History Museum, University of Oslo (Table
125 1). We only sampled herbarium material that was green, indicative of fast drying. Voucher
126 specimens exist for all samples. All data were managed in the Barcode of Life Systems
127 (BOLD) database in the project called "CHARA".

128

129 ***Taxonomy***

130 The samples were tentatively identified as representing 17 species of the genus *Chara* (Table
131 1). The number of individuals sampled per species ranged from 1 – 20 (Table 1). Wood and
132 Imahori (1965) subdivided the genus *Chara* into sections and subsections, and assumed a
133 close phylogenetic relationship among the members of each group. The material used in this
134 study contains specimen from the subsections Hartmania, Chara, and Grovesia. We focused
135 on taxa belonging to subsection Hartmania, because morphological traits used to delineate
136 these taxa have been reported to intergrade among individuals (i.e. intermediate forms are
137 observed; Boegle et al.~~,~~ 2007). Taxa within the subsection Hartmania are rather stout plants,
138 with a generally diplostichous cortex, elongated stipulodes, and geminate or fasciculate spine-
139 cells (see Fig. 1 for an explanation of typical morphological traits of a *Chara*); prominent taxa
140 of the subsection Hartmania include among others *C. hispida, C. intermedia,* and *C. baltica*
141 (Table 1). Other taxa from the subsection Chara (diplostichous cortex, solitary spines) and
142 Grovesia (triplostichous cortex) were included for comparison.

143 Many *Chara* taxa have been variously recognized as species, varieties, or forms, and there is
144 little consensus about appropriate rank among different flora treatments. The two most widely
145 applied taxonomic concepts are those of Wood and Imahori (1965) and Krause (1997). While
146 the former authors belong to the school of "lumpers" (lumping taxa into broad categories), the

latter is a so-called “splitter” (creating many narrowly defined categories). For example, Wood and Imahori (1965) discriminate 19 species world-wide within the genus *Chara*, whereas Krause (1997) recognizes 29 species in Europe alone.

In order to be consistent, and to provide barcode data that are taxonomically informative on an as detailed level as possible, our species delineation generally followed that of Krause (1997), with the following exceptions: i) *C. aculeolata* was differentiated by its longer spines and stouter appearance from *C. intermedia*, because it is an ongoing debate whether or not these two should be separated; Krause (1997) recognized this taxon as “form” within *C. intermedia*; ii) *C. corfuensis* was determined after Wood and Imahori (1965) because this taxon is not listed in Krause (1997); Wood and Imahori (1965) recognize this taxon as *C. hispida* var. *hispida* f. *corfuensis*; we gave this taxon species rank in order to be consistent with Krause’s (1997) taxonomic concept; iii) *C. vulgaris* var. *vulgaris* f. *calveraensis* was determined after Wood and Imahori (1965), because Krause (1997) does not list this taxon (the taxon is described from South America, and the treatment put forth by Krause (1997) only deals with European taxa).

Primer Design

Chara specific *matK* and ITS2 primers were designed based on sequences for *matK* and ITS2 that were obtained from GenBank (Table 2) and aligned in MEGA version 6 (Tamura et al.~~,~~ 2013). Resulting primers were named *Chara*_ITSF2, *Chara*_ITSR2, *Chara*_matKF2, and *Chara*_matKR2 (Table 3).

DNA extraction, amplification, and sequencing

Total genomic DNA was extracted from *Chara* material following the standard protocols at the Canadian Centre for DNA Barcoding (CCDB, Ivanova et al.~~,~~ 2008, Ivanova et al.~~,~~ 2011). The amplification and sequencing of the three gene regions (ITS2, *matK*, and *rbc*L, using primers found in Table 3) also followed the protocols of the CCDB, as detailed and described in Kuzmina et al. (2012). Sequence chromatograms were proofed, edited, and contigs assembled using the program CodonCode Aligner version 2.0.6 (CodonCode Co, USA). Contigs were aligned using the MUSCLE multiple sequence alignment algorithm (Edgar~~,~~ 2004) as implemented in CodonCode Aligner. These initial alignments were created in order to compare contigs generated from the various specimens and aided the identification and

179 correction of base calling errors following the examination of trace files. Specific for ITS2
180 amplicons, BLAST was utilized in order to determine whether any of the sequences produced
181 were of fungal contaminants.

182

183 ***Phylogenetic analyses***

184 Of the 91 specimens of *Chara* included in this study (Table 1), phylogenetic analyses were
185 conducted using the 73 samples for which we recovered sequences at each of the three
186 markers examined (*matK*, ITS2 and *rbc*L). We did so in order to consistently compare
187 discriminatory power among the three markers, and to ensure that our conclusions have
188 maximum credibility, i.e. are supported by three independent markers. For comparison,
189 phylogenetic trees produced from individual complete data sets are given in a~~A~~ppendix (Fig.
190 S1).

191 Barcode data were quality-controlled iteratively throughout data collection to detect potential
192 contamination, misidentification, and alignment error. We produced neighbor joining trees for
193 each marker, and looked for individuals that were grossly misplaced. Voucher specimens of
194 problematic samples were re-examined resulting in the correction of misidentified taxa, or the
195 removal of the sequence from the BOLD database.

196 Sequences were aligned using Align (version 03/2007) MS Windows-based manual sequence
197 alignment editor (SequentiX - Digital DNA Processing, Klein Raden Germany) to obtain
198 DNA sequence alignments, which were then corrected manually. Segments with highly
199 variable and ambiguous regions and gaps making proper alignment impossible were excluded
200 from the analyses. A *mat*K set containing 292 positions, an ITS2 set containing 183, and a
201 *rbc*L set containing 502 positions was used. *Chara longifolia* (AY170444), *Chara foliolosa*
202 (HQ380618) and *Chara foliolosa* (HQ380452) were employed as outgroup taxa in the *matK*,
203 ITS2 and *rbc*L tree, respectively. Data sets were analyzed using the maximum likelihood
204 (ML) algorithm in MEGA version 6 (Tamura et al.~~,~~ 2013). In the ML analyses, evolutionary
205 substitution models were evaluated in MEGA version 6. The method selected the same best-
206 fitting evolutionary model (T92) for the three markers (*matK*, ITS2 and *rbc*L). ML analyses
207 were performed with 1000 bootstrap replicates in MEGA version 6 (Tamura et al.~~,~~ 2013).

208 In the tree generated using the combined *matK*-ITS2-*rbc*L data set, no outgroup was used.
209 Phylogenetic inference was based on ML, and analyses were performed with Treefinder
210 (Jobb~~,~~ 2011) with three partitions. Models and parameters proposed by Treefinder under AICc

211 criteria were as follows: *matK* (292 bases; model HKY), *rbc*L (502 bases; model HKY), ITS2
212 (183 bases; model HKY). To provide support of relationships, bootstrap analyses were
213 calculated by ML (1000 replicates) criteria with Treefinder.

214

215 ***Sequence recoverability***

216 We calculated the number of *rbc*L, *matK* and ITS2 sequences in the entire dataset, obtained
217 from both herbarium specimens and from silica-gel dried samples. To determine if herbarium
218 specimen age and sequence recovery were correlated, we counted the number of sequences
219 recovered from specimens in each year. For correlation analysis, we only used those years
220 from which we had a minimum of three samples (i.e. the years 1969, 1995, 1998 and 2007
221 were omitted from the analysis). We then used Spearman rank correlation to test for a
222 relationship between year and sequence recovery, because we expected the relationship to be
223 monotonic, but not necessarily linear.

224

225 **~~3.~~ Results**

226 ***Sequence recoverability***

227 Sequence recoverability was highest for *matK* and *rbc*L, and lowest for ITS2 (Table 4). Silica
228 gel dried samples were 100% recovered for all three loci, while recoverability for the
229 herbarium specimen ranged from 93% (*matK*) to 84% (ITS2). Age of herbarium specimens
230 and sequence recovery were not significantly correlated (Spearman rho 0.16, 0.15 and 0.57
231 for *matK*, ITS2 and *rbc*L, respectively, for the correlation between sampling year and
232 sequence recovery; all $p > 0.05$).

233

234 ***Taxonomic uncertainties***

235 We corrected one misidentification in our dataset, where a *C. contraria* with unusually long
236 spine cells had mistakenly been identified as *C. polyacantha* (field ID MB 70; Table 1). This
237 sample clearly clustered to other *C. contraria* individuals, and this result was consistent
238 among all three analysed loci.
239 In addition, the three specimens of *C. vulgaris* from Argentina were found to cluster to two
240 different groups. While two accessions, determined by their elongated stipulodes as *C.*
241 *vulgaris* var. *vulgaris* f. *calveraensis* (field IDs 47 and 48) clustered within a large group

242 containing 9 different taxa (Figs. 2 and 3), the third sample (field ID 44) formed its own
243 branch (Figs. 2 and 3).

244

245 ***Consistency between barcode haplotypes and morphological species boundaries***

246 Each of the three investigated loci separated the taxa into six main groups (Fig. 2), and the
247 same six clusters were recovered in the tree produced from the analysis of the concatenated
248 data matrix (plastid and nuclear sequences; Fig. 3). The first group is a large cluster
249 containing 55 individuals (labelled cluster I; Fig. 3) which have traditionally been assigned to
250 nine different taxa (*C. intermedia, C. hispida, C. horrida, C. baltica, C. polyacantha, C. rudis,*
251 *C. aculeolata, C. corfuensis*, *C. vulgaris* var. *vulgaris* f. *calveraensis)*. ITS2 and *matK*
252 supported this cluster with bootstrap values of 93 and 97%, respectively, while the *rbc*L
253 cluster was poorly supported (36%). The only locus that separated a subgroup containing two
254 individuals of *C. vulgaris* var. *vulgaris* f. *calveraensis* within this large cluster of nine
255 different taxa was *matK* (Fig. 2b). None of the loci discriminated between the other eight taxa
256 in this cluster. A second cluster (labelled cluster II; Fig. 2, 3) containing five individuals of *C.*
257 *contraria* was recovered with bootstrap support of ≥ 67%; ITS2 was the only locus that
258 separated a subgroup containing two samples from Austria and Germany (labelled *C.*
259 *contraria* group b; Fig. 2a and Fig. 3) from the other three individuals which are from Canada
260 and Norway (labelled *C. contraria* group a; Fig. 2a and Fig. 3). Clusters III and IV contain
261 one individual each of *C.* c.f. *vulgaris* and *C. galioides,* respectively. A fifth group (labelled
262 cluster V; Fig. 2, 3) contains three individuals of *C. tomentosa* with bootstrap support of ≥
263 98%. Finally, cluster VI containing eight individuals, four of which belong to *C. virgata* and
264 four to *C. strigosa*, was supported by bootstrap values of 99% (Fig. 2, 3). None of the three
265 loci discriminated between *C. virgata* and *C. strigosa* (Fig. 2).

266 Tree topology for the three loci was similar, though not identical. In each of the trees, *C.*
267 *contraria* (in case of *rbc*L, together with *C.* c.f. *vulgaris* and *C. galioides*) was the group most
268 closely related to the large cluster I containing nine taxa, while *C. strigosa/virgata* (in case of
269 *rbc*L, together with *C. tomentosa*) was most distant (Fig. 2). In total, ITS2, *rbc*L and *matK*
270 were consistent in their fundamental results. ITS2 and *matK* each discriminated one subgroup
271 (Ib and IIb, respectively; Fig. 2) which the other loci did not recover. Apart from that,
272 discriminatory power of the three investigated loci did not differ. The concatenated tree
273 resulting from the analysis of the combined plastid and nuclear sequences resolved the same
274 six main groups as the individual analyses. Apart from the two groups consisting of one

275 individual each (*C.* c.f. *vulgaris*, *C. galioides*), the remaining four groups were each supported
276 by bootstrap values of ≥ 93% (Fig. 3). Tree topology was consistent with the results from the
277 separate analyses. The subgroup containing two individuals of *C. vulgaris* var. *vulgaris* f.
278 *calveraensis* differentiated by *matK* (Fig. 2 b), and the subgroup containing two samples of *C.*
279 *contraria* from Austria and Germany differentiated by ITS2 (Fig. 2a) were both represented in
280 the concatenated tree (Fig. 3), such that overall resolution of the concatenated tree was
281 slightly higher than the trees resulting from each individual analysis.

282

283 **~~4.~~ Discussion**

284 ***Sequence recoverability***

285 A multitude of factors influences recoverability of DNA barcodes, ranging from careless
286 preparation of samples, sample contamination and age of samples to unsatisfactory primer
287 design or inadequate amplicon length. For our samples, sequence recoverability for the
288 freshly collected and silica gel dried samples was 100% for all three loci, while recoverability
289 for the herbarium specimen ranged from 93% (*matK*) to 84% (ITS2). These values are
290 considerably higher than in some other analyses for plants (de Vere et al.~~,~~ 2012) and algae
291 (Kucera and Saunders~~,~~ 2012; Pérez et al.~~,~~ 2014), but correspond well with results on arctic
292 vascular plants (Kuzmina et al.~~,~~ 2012). They may simply be explained by the younger age of
293 the herbarium specimens we used (our oldest sample was from 1969, while de Vere et al.
294 (2012) managed to barcode a sample from 1868).

295 However, while matK performed considerably poorer than *rbc*L in de Vere et al. (2012) and
296 Kuzmina et al. (2012), the performance of these two plastid markers was approximately equal
297 for our data. The poor performance of *matK* in analyses of large datasets such as de Vere et al.
298 (2012) and Kuzmina et al. (2012) is mainly due to the fact that no universal primers exist for
299 more distantly related taxa, unlike *rbc*L for which universal primers are much more accessible
300 (Hollingsworth et al.~~,~~ 2011; de Vere et al.~~,~~ 2012). The good performance of *matK* may
301 therefore be taken as a sign that the *Chara* specimens we analysed are very closely related,
302 resulting in the ease of developing primers specific to the genus that would amplify *matK* for
303 all *Chara* specimens used. A poorer recoverability of ITS compared to *rbc*L and *matK* has
304 been described before (Hollingsworth~~,~~ 2011) and is commonly ascribed to the existence of
305 paralogous copies within individuals, which can prevent readable sequences from being
306 obtained (Hollingsworth et al.~~,~~ 2011). Our data are consistent with this finding (even with
307 developing ITS2 specific primers for the genus *Chara*). In summary, our results indicate that

for analyzing closely related species such as those in the genus *Chara*, *matK* and *rbc*L both perform reasonably well in terms of sequence recoverability, while ITS2 shows a somewhat lower recoverability.

In addition, a negative relationship between sequencing success and age of herbarium specimens has in some instances been documented for plants (de Vere et al.~~,~~ 2012; Saarela et al.~~,~~ 2013), while other studies showed no such association (Kuzmina et al.~~,~~ 2012). For our data, the correlation was not significant. However, visual inspection of the data indicated that sequence recoverability likely would have decreased with increasing age of herbarium samples if our data series had been longer (i.e. older samples were included). Altogether, we have shown that i) herbarium specimens are useful as a source of material for aquatic plants, like *Chara*, in spite of the expected slower drying timeframe compared to the often less "watery" land plants, ii) material sampled from up to 12 years old herbarium specimens is readily usable to obtain DNA and amplify barcode markers, in particular *rbc*L and *matK*, but iii) recoverability of sequences may decrease for older specimens.

Discriminatory power of ITS2, matK and rbcL

The choice of *rbcL+matK* as a core barcode for plants was based on the straightforward recovery of the *rbcL* region, plus the high discriminatory power of the *matK* region (CBOL~~,~~ 2009; Hollingsworth et al.~~,~~ 2011). Nevertheless, a greater discriminatory power of the entire nrITS compared to plastid regions has often been shown at low taxonomic levels (China Plant BOL Group~~,~~ 2011; Hollingsworth~~,~~ 2011). In our data, all three investigated regions discriminated the same six main groups, but ITS2 and *matK* each discriminated one additional subgroup which the other regions did not recover. This is consistent with the well-established fact that *rbc*L has lower discriminatory power (Hollingsworth et al.~~,~~ 2011), while *matK* is one of the most rapidly evolving coding sections of the plastid genome (Hilu and Liang~~,~~ 1997), and that the ITS region generally has great discriminatory power (China Plant BOL Group~~,~~ 2011). The topological agreement that was found in this study between all three regions (including plastid and nuclear derived markers) resulting in the recovery of six main groups is consistent with results of Sakayama et al. (2004a) on *Nitella* (which is, after *Chara*, the second most species rich-genus of the Charales), who also found that the phylogeny derived from nuclear ribosomal DNA (including ITS2) was congruent with chloroplast gene phylogeny (including *rbc*L). With respect to discriminatory power, our data confirm that ITS2 and *matK* may perform equally well, while *rbc*L has lower performance. At the same time,

341 ITS2 had lower sequence recoverability (see above). The choice of *rbc*L+*matK* as a core
342 barcode for plants seems therefore suitable also for *Chara*.

343

344 ***Taxonomic uncertainties***

345 DNA barcoding aims to facilitate species identification through substituting morphological
346 traits by standardized portions of the genome. In our dataset, one individual of *C. contraria*
347 had mistakenly been identified as *C. polyacantha*. Each of the three markers ITS2, *rbc*L and
348 *matK* clearly assigned this sample to *C. contraria*. Re-examination of the voucher specimen
349 revealed that this individual had unusually long spine cells. Thus, our results support the view
350 of Krause (1997) that "spiny" individuals of *C. contraria* are to be regarded as infraspecific
351 morphotypes of *C. contraria* (*C. contraria* var. *hispidula*). Our results contradict the view of
352 Wood and Imahori (1965), who placed the "forma *hispidula*" and "forma *contraria*" on equal
353 ranks within *C. vulgaris* var. *vulgaris*. *C. contraria* var. *hispidula* can be differentiated from
354 *C. polyacantha* by having solitary spine cells (*C. polyacantha*: generally fasciculate), and by
355 its smaller internode diameter (*C. contraria*: less than 1 mm, *C. polyacantha*: more than 1
356 mm; Krause, 1997).

357 Based on ITS2 results, *C. contraria* was divided into two sub-groups: "group a" containing
358 three samples from Canada and Norway, and "group b" containing two samples from Austria
359 and Germany (Figs. 2a, 3). While one of the samples in group b (field ID MB 70, Table 1)
360 was morphologically different from the other four individuals by having elongated spine cells
361 (= *C. contraria* var. *hispidula*), the other individuals showed no obviously distinct
362 morphological characteristics. Consequently, the two subgroups apparently did not represent
363 different morphotypes, but may possibly be regarded as geographically separated groups
364 (Northern circumpolar (Canada and Norway) and Central-European (Austria and Germany)).
365 More data are necessary to test this hypothesis, however.

366 The taxonomic concept applied by Wood and Imahori (1965) is based on the assumption that
367 a close phylogenetic relationship exists among the members of each subsection. However,
368 members of the subsections Chara and Grovesia did not form monophyletic groups (Fig. 3),
369 thus contradicting Wood and Imahori's (1965) assumption. This notwithstanding, all samples
370 belonging to the subsection Hartmania (Table 1) indeed clustered together (cluster I, Fig. 3),
371 which supports Wood and Imahori's assumption of a close phylogenetic relationship of the
372 taxa in this group. However, both individuals of *C. vulgaris* var. *vulgaris* f. *calveraensis* (field

373 ID 47, 48; Table 1) also clustered to the Hartmania section (Fig. 3). This result was consistent
374 among all three analysed loci (Fig. 2). Wood and Imahori (1965) separate the subsection
375 Chara (into which they placed *C. vulgaris* var. *vulgaris* f. *calveraensis*) from the subsection
376 Hartmania by their spine cells: individuals of the subsection Chara have solitary or rarely
377 geminate spine cells, while individuals of the subsection Hartmania have predominantly
378 fasciculate spine cells. Since re-examination of the voucher specimens gave no indication of a
379 possible misidentification, we suggest that *C. vulgaris* var. *vulgaris* f. *calveraensis* belongs to
380 the subsection Hartmania instead of the subsection Chara. This further indicates that the
381 number of spine cells (solitary, geminate or fasciculate), i.e. the criterion which was used by
382 Wood and Imahori (1965) to separate the subsection Hartmania from the subsection Chara,
383 may not be useful for inferring phylogenetic relationships.

384 The third sample of *C. c.f. vulgaris* (field ID 44; Table 1) formed its own branch (Figs. 2, 3).
385 *C. vulgaris* is characterized by a diplostichous (twice as many cortex rows as the number of
386 branchlets per whorl) and aulacanthous cortex (the primary cortex cells supporting the spines
387 are thinner than the secondary rows), but both these characteristics are also true for *C.*
388 *vulgaris* var. *vulgaris* f. *calveraensis*, which, as shown above, is genetically different from *C.*
389 *vulgaris* (Fig. 3). However, individuals of *C. vulgaris* var. *vulgaris* f. *calveraensis* are, in
390 accordance with all other taxa in the subsection Hartmania, rather stout and large plants
391 (Wood and Imahori (1965) describe the taxon as “moderately stout”, axis diameter to 900
392 µm). In contrast, *C. vulgaris* is, in accordance with most other varieties of *C. vulgaris* sensu
393 Wood and Imahori (1965), generally quite slender and small (Wood and Imahori (1965)
394 describe the taxon as “moderately slender”, axis diameter circa 500 µm). Clearly, more
395 molecular data are needed from *C. vulgaris* before answers can be given. Nevertheless, we
396 recommend that the internode diameter (corresponding to the often used description of a
397 *Chara* specimen as being “large and stout” as opposed to being “slender and quite small”)
398 should be paid more attention. This is surprising because plant size may be expected to be
399 influenced by the environment. Nevertheless, also the above mentioned misidentification of
400 *C. contraria* as *C. polyacantha* could have been avoided if internode diameter had been taken
401 into account.

402

403 ***Consistency between barcode haplotypes and morphological species boundaries***

404 Following a traditional morphological approach, the tree produced using the concatenated
405 data matrix (Fig. 3) consisted of individuals from 12 different taxa. In contrast, molecular data

revealed six main groups, in addition to the *C. vulgaris* var. *vulgaris* f. *calveraensis* subgroup and a morphologically un-differentiated subgroup within *C. contraria*. Differences between genomic and morphological species resolution are well-documented, and the existence of cryptic species (i.e. species that are morphologically indistinguishable but show genetic differences; e.g. Kucera and Saunders 2012) as well as the opposite, i.e. the existence of morphotypes that cannot be separated by barcoding (e.g. Seberg and Petersen 2009; China plant BOL group 2011; Kuzmina et al. 2012) have often been shown. In our data, there occurred two groups of genetically unresolved species. The first unresolved group (cluster Ia; Fig. 3) consisted of *C. intermedia, C. hispida, C. horrida, C. baltica, C. polyacantha, C. rudis, C. aculeolata,* and *C. corfuensis*. These eight taxa shared identical barcode sequences on the 977 positions of the concatenated sequence tree. This is remarkable, because other algal groups show considerably higher intra-specific variation (e.g. the marine green macroalgal genus *Caulerpa* J.V. Lamouroux shows intra-specific variation of the *rbc*L region; Belton et al., 2014). In spite of identical barcode sequences on the 977 positions we analyzed, these taxa are morphologically discriminated from each other by spine cells (single versus fasciculate), cortication (tylachanthous versus aulacanthous), stipulodes (two rows versus multiple rows), bract cells (elongated versus not elongated) and habitat (fresh water versus brackish water). Our results indicate that these morphological traits may not reflect major differences in DNA sequences and consequently may be of low taxonomic value for species discrimination in *Chara*. These results are consistent with earlier results obtained by AFLP (Amplified Fragment Length Polymorphism), a genetic fingerprinting technique that may have higher species resolution than barcoding (Roy et al. 2010). Boegle et al. (2010a) concluded, in an extension of results obtained by Mannschreck (2003) and Boegle et al. (2007), that the species complex around *C. intermedia* and *C. baltica* formed a continuum. In addition, there occurred no differences in AFLP fingerprints between *C. baltica* and *C. horrida* (Boegle et al. 2010b). Likewise, Urbaniak and Combik (2013) were unable to consistently differentiate individuals of *C. intermedia*, *C. baltica*, *C. polyacantha*, *C. rudis* and *C. hispida* by AFLP. Mannschreck (2003) and Boegle et al. (2007, 2010a, b) indeed were able to differentiate *C. hispida* by AFLP. However, *C. hispida* clustered in all instances next to *C. intermedia*, *C. baltica* and *C. horrida*, the difference between *C. hispida* and the other taxa was small and they together formed a monophyletic group supported by high bootstrap values. Our dataset on this species group was from a larger geographic area than the samples used in the above mentioned AFLP studies, spanning seven countries in Europe, from Norway in the North to Greece in the South, and from Poland in the East to Spain in the West.

Nevertheless, barcodes of all taxa in cluster Ia were identical on the 977 positions used for the concatenated tree. A subgroup consisting of *C. vulgaris* var. *vulgaris* f. *calveraensis* from Argentina (cluster Ib, Fig. 3) was separated. This taxon differed in 1 out of the 292 basepairs within *matK*, while the 502 *rbc*L and 183 ITS2 sequence sites were identical to the taxa found in cluster Ia. Such small differences are well within the accepted intra-specific variation of other algal species (Belton et al.~~,~~ 2014; Leliaert et al.~~,~~ 2014). In summary, the differences in barcode sequences of samples in cluster I were very small, even though samples were from two different continents. They point towards a very close phylogenetic relationship among these taxa, and lend support to Wood and Imahori's (1965) view of lumping these taxa into one species that has considerable morphological variation.

The second group of unresolved species was formed by *C. virgata* and *C. strigosa* (cluster VI, Fig. 3). Barcodes of these species were identical across the 977 positions of the concatenated sequence matrix, in spite of conspicuous morphological differences that exist in spine cells and stipulodes (*C. virgata*: only the upper row of stipulodes is well developed, spine cells are rudimentary; *C. strigosa*: two well-developed rows of stipulodes, spine cells are elongated and fasciculate). These results are in accordance with earlier results obtained by AFLP (Mannschreck~~,~~ 2003), and indicate that the length and number of spine cells, as well as the length of stipulodes may be of low taxonomic value in *Chara*. There is a general consensus that algal species may be viewed as separately evolving metapopulation lines (Leliaert et al.~~,~~ 2014). Since *C. virgata* typically occurs in Calcium poor habitats, while *C. strigosa* typically occurs in Calcium rich habitats (Rey-Boissezon and Auderset Joye~~,~~ 2015), these taxa may well "evolve separately" in spite of their genetic similarity. Clearly, more data are needed to clarify the taxonomic status of these taxa. However, our results indicate a very close phylogenetic relationship between *C. virgata* and *C. strigosa.*

In clades where speciation has been very recent, or where rates of mutation are slow, barcode sequences may be shared among related taxa (Hollingsworth et al.~~,~~ 2011). In such cases, the commonly used barcode markers may be too conservative, while loci that are associated with traits that are under selection may be more informative (Leliaert et al.~~,~~ 2014). Heritable phenotypic modifications in the absence of differences in DNA barcodes can also be caused by the environment via e.g. epigenetic variation, such as DNA methylation (Cubas et al.~~,~~ 1999; Zhang et al.~~,~~ 2013). Verhoeven et al. (2010) showed that stress, e.g. chemical induction of herbivore and pathogen defenses, can trigger considerable variation in methylation of plant DNA. Consequently, habitat salinity (*C. baltica* and *C. horrida* typically occur in brackish

473 water habitats, while the other taxa in cluster I typically occur in freshwater habitats) or Ca-
474 concentration (*C. virgata*: Calcium poor habitats, *C. strigosa*: Calcium rich habitats; Rey-
475 Boissezon and Auderset Joye, 2015) may also be related to epigenetic variation. A third
476 explanation for phenotypic modifications in the absence of differences in DNA barcodes may
477 be polyploidy (Schranz and Osborn, 2004). However, recent evidence suggests that epigenetic
478 rather than genetic factors may explain phenotypic divergence between plant populations of
479 different ploidy (Rois et al., 2013).

480

481 **5. Conclusions**

482 The morphological characters that are traditionally used to differentiate between taxa found in
483 the two genetically unresolved clusters in this study (Fig. 3) are the number and length of
484 spine cells, stipulodes, and bract cells. We suggest that these morphological traits are of lower
485 taxonomic relevance than hitherto assumed. This is supported by (1) the observation that the
486 "spiny" *C. contraria* var. *hispidula* was not differentiated from the "normal" *C. contraria*
487 either, (2) no differences were recovered in AFLP fingerprints among varieties of *C. aspera*
488 that morphologically differ with respect to number and length of spine cells (poorly
489 developed, single, fasciculate; Mannschreck, 2003; O'Reilly et al., 2007), and (3) the length
490 of stipulodes and bract cells did not coincide with genetic differences in *C. braunii* either
491 (Kato et al., 2008). Similar observations exist for the genus *Nitella* (Charales), where the form
492 and cell number of dactyls (terminal cells in a branchlet ray), a morphological trait used for
493 species differentiation, were variable within the clades (Sakayama et al., 2004b). Our results
494 are also in accordance with results of Perez et al. (2014) on the genus *Tolypella*, the third most
495 common genus within the Characeae. Perez et al. (2014) found that some species shared
496 identical sequences despite radically different growth forms, and indicated that some authors
497 tend to over-emphasize slight morphological differences to delineate species.

498 Morphological traits such as the length and number of spine cells or stipulodes are readily
499 visible in *Chara*, resulting in their prominent use for species discrimination in this genus.
500 Also, one may sometimes easily be misled. For example, *C. vulgaris* var. *vulgaris* f.
501 *calveraensis* was discriminated as a subgroup by *matK*, and this taxon indeed is
502 morphologically different from the other eight taxa in this group (cluster I, Fig. 3) by its
503 elongated and irregular stipulodes. However, the genetic difference may also be explained by
504 geographic separation (*C. vulgaris* var. *vulgaris* f. *calveraensis* was from Argentina, while the
505 other eight taxa in this group were from Europe). Taken together, increasing evidence has

accumulated that *Chara* taxa which exclusively differ in the number and length of spine cells, stipulodes, and bract cells are genetically closely related, and may be regarded as varieties rather than species. In contrast, oospore traits were shown to be useful for species differentiation in *Nitella* (Sakayama et al. 2004b), and this may well work also for *Chara*. Indeed, taxa included in cluster I cannot be differentiated by oospore traits (Blume et al. 2009), while other species are different (Holzhausen et al. 2015), thus supporting the results summarized in Fig. 3.

The unresolved cluster Ia contains eight *Chara* taxa from Europe which share identical barcode sequences. Many, but not all of these taxa are assigned to various IUCN Red List categories (e.g. Sjøtun et al. 2010; Auderset Joye and Schwarzer 2012). The same is true for cluster VI, which consists of *C. virgata* and *C. strigosa*. While the former often is regarded as quite common, *C. strigosa* is often red listed (e.g. Sjøtun et al. 2010; Auderset Joye and Schwarzer 2012). While the IUCN criteria for Red Lists are open for inclusion of subspecies and varieties, provided an assessment of the full species is also given (IUCN 2014), the conservation status of *Chara* species clearly requires renewed attention.

Acknowledgements

Anders Langangen and Michael Boegle are gratefully acknowledged for producing many of the analyzed herbarium samples. The CHARA project was financially supported by Artsdatabanken and by NIVA through the strategic institute initiative “biodiversity”.

References

Auderset Joye, D. & Rey-Boissezon, A. 2015. Will charophyte species increase or decrease their distribution in a changing climate? *Aquat Bot* doi:10.1016/j.aquabot.2014.05.003.

Auderset Joye, D. & Schwarzer, A. 2012: Rote Liste Armleuchteralgen. Gefährdete Arten der Schweiz, Stand 2010. Bundesamt für Umwelt, Bern und Laboratoire d’écologie et de biologie aquatique (LEBA) der Universität Genf. Umwelt-Vollzug Nr. 1213:72 pp.

Baastrup-Spohr, L., Iversen, L.L., Dahl-Nielsen, J. & Sand-Jensen, K. 2013. Seventy years of changes in the abundance of Danish charophytes. *Freshwater Biol* 58:1682-1693.

Belton, G.S., van Reine, W.F.Prud'homme, Huisman, J.M., Draisma, S.G.A., Gurgel, C.F.D. 2014. Resolving phenotypic plasticity and species designation in the morphologically

539 challenging *Caulerpa racemosa-peltata* complex (Caulerpaceae, Chlorophyta). *J~~.~~ Phycol~~.~~*
540 50:32-54.
541 Blindow, I. 1992. Decline of charophytes during eutrophication – comparison with
542 angiosperms. *Freshwater Biol~~.~~* 28:9-14.
543 Blume, M., Blindow, I., Dahlke, S. & Vedder, F. 2009. Oospore variation in closely related
544 *Chara* taxa. *J~~.~~ Phycol~~.~~* 45:995-1002.
545 Boegle, M.G., Schneider, S., Mannschreck, B. & Melzer, A. 2007. Differentiation of *Chara*
546 *intermedia* and *C. baltica* compared to *C. hispida* based on morphology and amplified
547 fragment length polymorphism. *Hydrobiologia* 586:155-166.
548 Boegle, M.G., Schneider, S.C., Schubert, H. & Melzer, A. 2010a. *Chara baltica* Bruzelius
549 1824 and *Chara intermedia* A. Braun 1859 - distinct species or habitat specific
550 modifications? *Aquat~~.~~ Bot~~.~~* 93:195-201.
551 Boegle, M.G., Schneider, S.C., Melzer, A. & Schubert, H. 2010b. Distinguishing *Chara*
552 *baltica*, *C. horrida* and *C. liljebladii* – conflicting results from analysis of morphology and
553 genetics. *Charophytes* 2:53-58.
554 CBOL Plant Working Group, 2009. A DNA Barcode for land plants. *P Natl Acad Sci USA*
555 ~~*PNAS*~~ 106:12794-12797.
556 China Plant BOL Group, 2011. Comparative analysis of a large dataset indicates that internal
557 transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. *P*
558 *Natl Acad Sci USA* ~~*PNAS*~~ 108:19641-19646.
559 Cubas, P., Vincent, C., & Coen, E. 1999. An epigenetic mutation responsible for natural
560 variation in floral symmetry. *Nature* 401:157-161.
561 de Vere, N., Rich, T.C.G., Ford, C.R., et al., 2012. DNA Barcoding the Native Flowering
562 Plants and Conifers of Wales. *PLOS ONE* 7: article number e37945.
563 Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high
564 throughput. *Nucleic Acids Res~~earch~~* 32:1792–1797.
565 Hilu, K.W. & Liang, H. 1997. The *matK* gene: sequence variation and application in plant
566 systematics. *Am~~.~~ J~~.~~ Bot~~.~~* 84:830–839.
567 Hollingsworth, P.M. 2011. Refining the DNA barcode for land plants. *P Natl Acad Sci USA*
568 ~~*PNAS*~~ 108:19451–19452.
569 Hollingsworth, P.M., Graham, S.W. & Little, D.P. 2011. Choosing and Using a Plant DNA
570 Barcode. *PLOS ONE* 6: article number e19254.
571 Holzhausen, A., Nowak, P., Niedrig, C., Feike, M., Schubert, H. 2015. Morphometry of
572 *Chara aspera, C. canescens, C. baltica* var. *baltica*, *C. baltica* var. *liljebladii* and *C.*

573 *intermedia* oospores: local variation versus taxonomic differences. *Aquat. Bot.* doi:
574 10.1016/j.aquabot.2014.07.001.
575 IUCN (International Union for Conservation of Nature and Natural Resources) 2014.
576 Guidelines for Using the IUCN Red List Categories and Criteria, version 11.
577 http://jr.iucnredlist.org/documents/RedListGuidelines.pdf.
578 Ivanova, N.V., Fazekas, A., & Hebert, P.D.N. 2008. Semi-automated, membrane based
579 protocol for DNA isolation from plants. *Plant Mol. Biol. Rep.* 26:186-198.
580 Ivanova, N., Kuzmina, M. & Fazekas, A.J. 2011. CCDB Protocols (2011). Manual Protocol
581 Employing Centrifugation: Glass Fiber Plate DNA Extraction Protocol for Plants, Fungi,
582 Echinoderms and Mollusks. http://www.ccdb.ca/docs/CCDB_DNA_Extraction-Plants.pdf.
583 Jobb, G. *Treefinder version of March 2011*. 2011. Munich, Germany. Distributed by the
584 author at www.Treefinder.de.
585 Karol, K., McCourt, R., Cimino, M. & Delwiche, C.F. 2001. The closest living relatives of
586 land plants. *Science* 294:2351–2353.
587 Kato, S., Sakayama, H., Sano, S., Kasai, F., Watanabe, M.M., Tanaka, J. & Nozaki, H. 2008.
588 Morphological variation and intraspecific phylogeny of the ubiquitous species *Chara*
589 *braunii* (Charales, Charophyceae) in Japan. *Phycologia* 47:191-202.
590 Krause, W. 1997. Charales (Charophyceae). In: Ettl, H., Gärtner, G., Heynig, H. &
591 Mollenhauer, D. (eds.). Süßwasserflora von Mitteleuropa, Band 18. Fischer, Jena.
592 Kress, W.J., Erickson, D.L. 2007. A Two-Locus Global DNA Barcode for Land Plants: The
593 Coding rbcL Gene Complements the Non-Coding trnH-psbA Spacer Region. *PLOS ONE*
594 2: e508.
595 Kucera, H. & Saunders, G.W. 2012. A survey of Bangiales (Rhodophyta) based on multiple
596 molecular markers reveals cryptic diversity. *J. Phycol* 48:869-882.
597 Kufel, L., Biardzka, E. & Strzalek, M. 2013. Calcium carbonate incrustation and phosphorus
598 fractions in five charophyte species. *Aquat. Bot.* 109:54-57.
599 Kuzmina, M.L., Johnson, K.L., Barron, H.R. & Hebert, P.D.N. 2012. Identification of the
600 vascular plants of Churchill, Manitoba, using a DNA barcode library. *BMC Ecology* 12:
601 article number 25.
602 Leliaert, F., Verbruggen, H., Vanormelingen, P., Steen, F., Lopez-Bautista, J.M., Zuccarello,
603 G.C., De Clerck, O. 2014. DNA- based species delimitation in algae. *Eur. J. Phycol.*
604 49:179-196.

Formatted: Font: Italic

605 Levin, R.A. Wagner, W.L., Hoch, P.C., Nepokroeff, M., Pires, J.C., Zimmer, E.A., Sytsma,
606 K.J. 2003. Family-level relationships of Onagraceae based on chloroplast rbcL and ndhF
607 data. *Am J Bot* 90:107-115.

608 Mannschreck, B. 2003. Genetische und morphologische Differnzierung ausgewählter Arten
609 der Gattung *Chara* (genetic and morphologic characterization of selected *Chara* species).
610 Ph.D. dissertation. Technical University Munich, Germany. Shaker Verlag.

611 O'Reilly, C.L., Cowan, R.S., & Hawkins, J.A. 2007. Amplified fragment length
612 polymorphism genetic fingerprinting challenges the taxonomic status of the near-endemic
613 species *Chara curta* Nolte ex Kütz. (Characeae). *Bot J Linn Soc* 155:467–476.

614 Pérez, W., Hall, J.D., McCourt, R., Karol, K.G. 2014. Phylogeny of North American
615 *Tolypella* (Charophyceae, Charophyta) based on plastid DNA sequences with a description
616 of *Tolypella ramosissima* sp. nov. *J Phycol* 50:776-789.

617 Rey-Boissezon, A. & Auderset Joye, D. 2015. Habitat requirements of charophytes - evidence
618 of species discrimination through distribution analysis. *Aquat Bot* doi:
619 10.1016/j.aquabot.2014.05.007.

620 Rois, A.S., Lopez, C.M.R., Cortinhas, A., Erben, M., Espirito-Santo, D., Wilkinson, M.J. &
621 Caperta, A.D. 2013. Epigenetic rather than genetic factors may explain phenotypic
622 divergence between coastal populations of diploid and tetraploid *Limonium* spp.
623 (Plumbaginaceae) in Portugal. *BMC Plant Biol* 13: article number 205.

624 Roy, S., Tyagi, A., Shukla, V., Kumar, A., Singh, U.M., Chaudhary, L.B., Datt, B., Bag, S.K.,
625 Singh, P.K., Nair, N.K., Husain, T., & Tuli, R. 2010. Universal Plant DNA Barcode Loci
626 May Not Work in Complex Groups: A Case Study with Indian Berberis Species. *PLOS*
627 *ONE* 5: article number e13674.

628 Saarela, J.M., Sokoloff, P.C., Gillespie, L.J., Consaul, L.L. & Bull, R.D. 2013. DNA
629 Barcoding the Canadian Arctic Flora: Core Plastid Barcodes (rbcL plus matK) for 490
630 Vascular Plant Species. *PLOS ONE* 8: article number e77982.

631 Sakayama, H., Hara, Y., Arai, S., Sato, H., & Nozaki, H. 2004a. Phylogenetic analyses of
632 *Nitella* subgenus Tieffallenia (Charales, Charophyceae) using nuclear ribosomal DNA
633 internal transcribed spacer sequences. *Phycologia* 43:672-681

634 Sakayama, H., Hara, Y. & Nozaki, H. 2004b. Taxonomic re-examination of six species of
635 *Nitella* (Charales, Charophyceae) from Asia, and phylogenetic relationships within the
636 genus based on *rbc*L and *atp*B gene sequences. *Phycologia* 43:91-104.

637 Sarkinen, T., Staats, M., Richardson, J.E., Cowan, R.S. & Bakker, F.T. 2012. How to Open
638 the Treasure Chest? Optimising DNA Extraction from Herbarium Specimens. *PLOS ONE*
639 7: article number e43808.
640 Schneider, S. & Melzer, A. 2003. The Trophic Index of Macrophytes (TIM) – a New Tool for
641 Indicating the Trophic State of Running Waters. *Int Rev Hydrobiol* 88:49-67.
642 Schneider, S.C., Garcia, A., Martín-Closas, C. & Chivas, A.R. 2015a. The role of charophytes
643 (Charales) in past and present environments: an overview. *Aquat Bot* doi:
644 10.1016/j.aquabot.2014.10.001.
645 Schneider, S.C., Pichler, D.E., Andersen, T. & Melzer, A. 2015b. Light acclimation in
646 submerged macrophytes: the roles of plant elongation, pigmentation and branch orientation
647 differ among *Chara* species. *Aquat Bot* doi: 10.1016/j.aquabot.2014.05.002.
648 Schranz, M.E. & Osborn, T.C. 2004. De novo variation in life-history traits and responses to
649 growth conditions of resynthesized polyploid Brassica napus (Brassicaceae). *Am J Bot*
650 91:174-183.
651 Seberg, O. & Petersen, G. 2009. How many loci does it take to DNA barcode a *Crocus*?
652 *PLOS ONE* 4: article number e4598.
653 Sjøtun, K., Fredriksen, S., Heggøy, E., Husa, V., Langangen, A., Lindstrøm, E-A., Moy, F.,
654 Rueness, J. & Åsen, P.A. 2010. Cyanophyta, Rhodophyta, Chlorophyta, Ochrophyta. In:
655 Kålås, J.A., Viken, Å., Henriksen, S. & Skjelseth, S. (eds.). 2010. The Norwegian Red List
656 for Species. Norwegian Biodiversity Information Centre, Norway.
657 Stelzer, D., Schneider, S. & Melzer, A. 2005. Macrophyte based assessment of lakes – a
658 contribution to the implementation of the European Water Framework Directive in
659 Germany. *Int Rev Hydrobiol* 90:223-237.
660 Tamura, K., Stecher, G., Peterson, D., Filipski, A. & and Kumar, S. 2013. MEGA6:
661 Molecular Evolutionary Genetics Analysis Version 6.0. *Mol Biol Evol* 30:2725-2729.
662 Timme, R.E., Bachvaroff, T.R. & Delwiche, C.F. 2012. Broad phylogenomic sampling and
663 the sister lineage of land plants. *PLOS ONE* 7: article number e29696.
664 Urbaniak, J. & Combik, M. 2013. Genetic and morphological data fail to differentiate *Chara*
665 *intermedia* from *C. baltica*, or *C. polyacantha* and *C. rudis* from *C. hispida*. *Eur J Phycol*
666 48:253-259.
667 Verhoeven, K. J. F., Jansen, J. J., van Dijk, P. J. & Biere, A. 2010. Stress-induced DNA
668 methylation changes and their heritability in asexual dandelions. *New Phytol*. 185:1108-
669 1118.

670 Vogel, J., Borner, T. & Hess, W.R. 1999. Comparative analysis of splicing of the complete set
671 of chloroplast group II introns in three higher plant mutants. *Nucleic Acids Res*. 27:3866-
672 3874.
673 Wodniok, S., Brinkmann, H., Glockner, G., Heidel, A., Philippe, H., Melkonian, M. &
674 Becker, B. 2011. Origin of land plants: Do conjugating green algae hold the key? *BMC*
675 *Evol. Biol.* 11:104.
676 Wood, R.D. & Imahori, K. 1965. A Revision of the Characeae. In: First Part: Monograph of
677 the Characeae. J. Cramer Verlag, Weinheim.
678 Zhang, Y.-Y., Fischer, M., Colot, V. & Bossdorf, O. 2013. Epigenetic variation creates
679 potential for evolution of plant phenotypic plasticity. *New Phytol.* 197:314-322.

Formatted: Font: Italic

680 **Figure captions**

681

682 **Fig. 1.** Schematic drawing of the uppermost part of a *Chara* specimen; the cortex can be i)
683 haplostichous (number of cortex cell rows corresponds to the number of branchlets),
684 diplostichous (twice as many cortex cell rows as the number of branchlets), or triplostichous
685 (three times as many cortex cell rows as the number of branchlets), and ii) aulacanthous
686 (secondary cortex cell rows more prominent, spines on thinner cortex cells), tylacanthous
687 (primary cortex cell rows more prominent, spines on thicker cortex cells), or isostichous
688 (primary and secondary cortex cells equally prominent). Fig 1a shows an example of a
689 diplostichous aulacanthous cortex with fasciculate spines (e.g. *C. hispida*), Fig. 1b illustrates
690 an example of a diplostichous tylacanthous cortex with single spines (e.g. *C. baltica*).

691

692

693 **Fig. 2.** Bootstrapped condensed ML (maximum likelihood) tree of a) ITS2, b) *matK* and c)
694 *rbc*L sequences of 73 *Chara* samples from which all three loci were successfully recovered;
695 bootstrap values above 50 % are shown in the tree. The clusters belong to the following
696 subsections according to Wood and Imahori (1965): I = Hartmania, II = Chara, III = Chara,
697 IV = Grovesia, V = Chara, VI = Grovesia. To improve readability, outgroups are not shown.

698

699

700 **Fig. 3.** Concatenated ML (maximum likelihood) tree of ITS2, *matK* and *rbc*L sequences of 73
701 *Chara* samples. Bootstrap values above 50 % are shown in the tree. The bar indicates 1%
702 sequence divergence. *= for sample ID see Table 1. The clusters belong to the following
703 subsections according to Wood and Imahori (1965): I = Hartmania, II = Chara, III = Chara,
704 IV = Grovesia, V = Chara, VI = Grovesia.

705

706

Table 1. List of 91 *Chara* individuals used in the present study. All samples were taken from herbarium collections, except those marked as "silica gel". NHM = Natural History Museum, University of Oslo; NIVA = Norwegian Institute for Water Research

Identification	author	subsection sensu Wood and Imahori (1965)	comment	BOLD Sample ID	Field ID	coll. year	country	silica gel (1=yes)	matK re-covered (1=yes)	ITS2 re-covered (1=yes)	rbcL re-covered (1=yes)	voucher specimen
Chara aculeolata	Kütz. in Rchb. 1832	Hartmania		CHARA_A_00038	T38	2007	Norway		1	1	1	NHM
Chara aculeolata				CHARA_A_00003	T3	2012	Norway		1	1	1	NHM
Chara aculeolata				CHARA_A_00037	T37	2010	Norway		1	1	1	NHM
Chara aculeolata				CHARA_A_00039	T39	2010	Norway		1	1	1	NHM
Chara aspera	Willd. 1809	Grovesia		CHARA_A_00049	T49	2011	Norway		1	0	1	NHM
Chara aspera				CHARA_A_00074	MB10	2000	Germany		1	0	1	NIVA
Chara aspera				CHARA_A_00075	MB13	2000	Germany		1	0	0	NIVA
Chara aspera				CHARA_A_00077	MB23	2005	Sweden		1	0	1	NIVA
Chara aspera				CHARA_A_00085	MB 67	2005	UK		1	0	1	NIVA
Chara baltica	Bruzelius 1824	Hartmania		CHARA_A_00082	MB39	2004	Greece		0	1	1	NIVA
Chara baltica				CHARA_A_00078	MB34	2004	France		1	1	1	NIVA
Chara baltica				CHARA_A_00035	T35	2010	Norway		1	1	1	NHM
Chara baltica				CHARA_A_00036	T36	2010	Norway		1	1	1	NHM
Chara baltica				CHARA_A_00084	MB47	2004	Sweden		1	0	0	NIVA
Chara baltica				CHARA_A_00083	MB43	2002	Sweden		0	1	1	NIVA
Chara baltica				CHARA_A_00081	MB37	2005	Germany		1	1	1	NIVA
Chara baueri	A. Br. 1847	Braunia		CHARA_A_00052	S29	2011	Germany		1	0	1	NIVA
Chara contraria	A. Br. ex Kütz. 1845 s. str.	Chara		CHARA_A_00092	MB 83	2000	Germany		1	1	1	NIVA
Chara contraria				CHARA_A_00050	T50	2011	Norway		1	1	1	NHM
Chara contraria				CHARA_A_00054	S54	2006	Canada		1	1	1	NIVA
Chara contraria			var. *hispidula*	CHARA_A_00087	MB 70	2001	Austria		1	1	1	NIVA
Chara contraria				CHARA_A_00051	T51	2009	Norway		1	1	1	NHM
Chara corfuensis	(J. Gr. Ex Fil.) R.D.W. 1965	Hartmania		CHARA_A_00055	S55	2006	Greece		1	1	1	NIVA
Chara fibrosa	Ag. ex Bruz. 1824	Agardhia		CHARA_A_00053	S51	2006	Canada		1	0	1	NIVA
Chara galioides	De Candolle 1813	Grovesia		CHARA_A_00091	MB 81	2001	France		1	1	1	NIVA
Chara galioides				CHARA_A_00090	MB 80	2001	France		1	0	0	NIVA
Chara hispida	(L.) Hartm. 1820	Hartmania		CHARA_A_00001	T1	2005	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00002	T2	2003	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00005	T5	2012	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00006	T6	2012	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00009	T9	2012	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00010	T10	2012	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00011	T11	2002	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00012	T12	1998	Norway		0	1	0	NHM
Chara hispida				CHARA_A_00013	T13	2011	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00014	T14	2011	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00015	T15	1995	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00016	T16	2003	Norway		1	1	0	NHM
Chara hispida				CHARA_A_00017	T17	2010	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00018	T18	2009	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00019	T19	2010	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00020	T20	2010	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00021	T21	2002	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00070	49	2012	Germany	1	1	1	1	NIVA
Chara hispida				CHARA_A_00086	MB 68	2005	Germany		1	1	1	NIVA
Chara hispida				CHARA_A_00095	MB 87	2001	Germany		1	1	1	NIVA
Chara horrida	Wahlst. 1862	Hartmania		CHARA_A_00079	MB35	2005	Sweden		1	1	1	NIVA
Chara intermedia	A. Br. in Br., Rab. and Stiz. 1859	Hartmania		CHARA_A_00094	MB 86	2001	Germany		1	1	1	NIVA
Chara intermedia				CHARA_A_00072	MB2	2004	Germany		1	1	1	NIVA
Chara intermedia				CHARA_A_00093	MB 85	2003	Sweden		1	1	1	NIVA
Chara intermedia				CHARA_A_00063	35	2012	Poland	1	1	1	1	NIVA
Chara polyacantha	A. Br. in Br., Rab. and Stiz. 1859	Hartmania		CHARA_A_00073	MB3	2004	Germany		1	1	1	NIVA
Chara polyacantha				CHARA_A_00057	5	2008	Spain	1	1	1	1	NIVA
Chara polyacantha				CHARA_A_00022	T22	2008	Norway		1	1	1	NHM
Chara polyacantha				CHARA_A_00023	T23	2009	Norway		0	0	0	NHM
Chara polyacantha				CHARA_A_00064	37	2012	Poland	1	1	1	1	NIVA
Chara polyacantha				CHARA_A_00065	38	2012	Poland	1	1	1	1	NIVA
Chara polyacantha				CHARA_A_00080	MB36	2005	Sweden		1	1	1	NIVA
Chara rudis	A. Br. in Leonhardi 1882	Hartmania		CHARA_A_00004	T4	2012	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00028	T28	2008	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00061	28	2010	Norway	1	1	1	1	NIVA
Chara rudis				CHARA_A_00024	T24	2010	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00007	T7	2012	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00027	T27	2009	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00029	T29	2008	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00025	T25	2010	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00034	T34	2011	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00033	T33	2011	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00032	T32	2011	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00031	T31	2011	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00026	T26	2010	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00030	T30	2008	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00008	T8	2012	Norway		1	1	1	NHM
Chara strigosa	A. Braun 1847	Grovesia		CHARA_A_00089	MB 72	2003	Germany		1	1	1	NIVA
Chara strigosa				CHARA_A_00047	T47	2010	Norway		1	1	1	NHM
Chara strigosa				CHARA_A_00048	T48	2010	Norway		1	1	1	NHM
Chara strigosa				CHARA_A_00088	MB 71	2002	Germany		0	1	1	NIVA
Chara strigosa				CHARA_A_00046	T46	2011	Norway		1	1	1	NHM
Chara tomentosa	L. 1753	Chara		CHARA_A_00041	T41	2011	Norway		1	1	1	NHM
Chara tomentosa				CHARA_A_00040	T40	2010	Norway		1	1	1	NHM
Chara tomentosa				CHARA_A_00042	T42	1969	Norway		0	0	0	NHM
Chara tomentosa				CHARA_A_00076	MB18	2005	Sweden		1	1	1	NIVA
Chara virgata	Kütz. 1834	Grovesia		CHARA_A_00071	50	2012	Germany	1	1	1	1	NIVA
Chara virgata				CHARA_A_00066	39	2012	Finland	1	1	1	1	NIVA
Chara virgata				CHARA_A_00056	S57	2012	Norway		1	1	1	NIVA
Chara virgata				CHARA_A_00044	T44	2010	Norway		1	0	1	NHM
Chara virgata				CHARA_A_00045	T45	2008	Norway		1	1	1	NHM
Chara virgata				CHARA_A_00043	T43	2011	Norway		1	0	1	NHM
Chara vulgaris	R.D.W. 1965	Chara	var. *vulgaris* f. *calveraensis*	CHARA_A_00068	47	2012	Argentina	1	1	1	1	NIVA
Chara vulgaris	R.D.W. 1965		var. *vulgaris* f. *calveraensis*	CHARA_A_00069	48	2012	Argentina	1	1	1	1	NIVA
Chara vulgaris	L. 1753		c.f.	CHARA_A_00067	44	2012	Argentina	1	1	1	1	NIVA

712 **Table 2.** *Chara* sequences used to design *Chara*-specific *matK* and ITS2 primers

713

Marker	Species	GenBank Accession
ITS2	*C. foliolosa*	HQ380620
	C. hydropitys	HQ380626
	C. haitensis	HQ380624
	C. rusbyana	HQ380627
	C. zeylanica	HQ380634
matK	*C. connivens*	AY170442
	C. globularis	AY170443
	C. longifolia	AY170444
	C. polycantha	AY170445
	C. vulgaris	108773196

714

715

716 **Table 3.** Primers used in this study

717

Primer	Sequences	Reference
Chara_ITSF2	CCCCCTTCGATTTTGAAGTT	This study
Chara_ITSR2	ACATCCCCGATTGCCAAC	This study
Chara_matKF2	GAACGAATCCGTGATAAAAGC	This study
Chara_matKR2	CTTCGGCCTTTCAAAAAGAA	This study
rbcLa-F	ATGTCACCACAAACAGAGACTAAAGC	Levin et al., 2003
rbcLa-R	GTAAAATCAAGTCCACCRCG	Kress and Erickson, 2007

718

719

720 **Table 4.** Number of recovered *matK*, ITS2 and *rbc*L sequences for 91 *Chara* individuals

721

	matK	ITS2	rbcL
total			
number of samples	91	91	91
number of recovered sequences	85	78	84
sequence recovery (%)	*93*	*86*	*92*
silica gel			
number of samples	11	11	11
number of recovered sequences	11	11	11
sequence recovery (%)	*100*	*100*	*100*
herbarium			
number of samples	80	80	80
number of recovered sequences	74	67	73
sequence recovery (%)	*93*	*84*	*91*

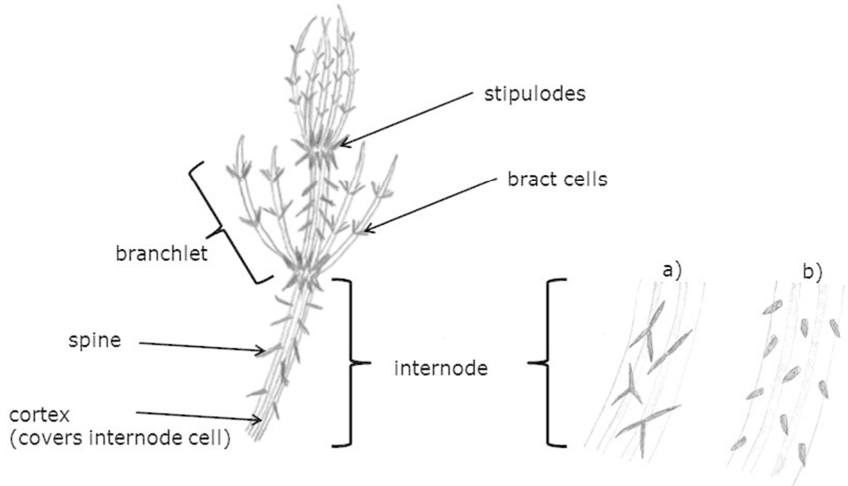
722

723

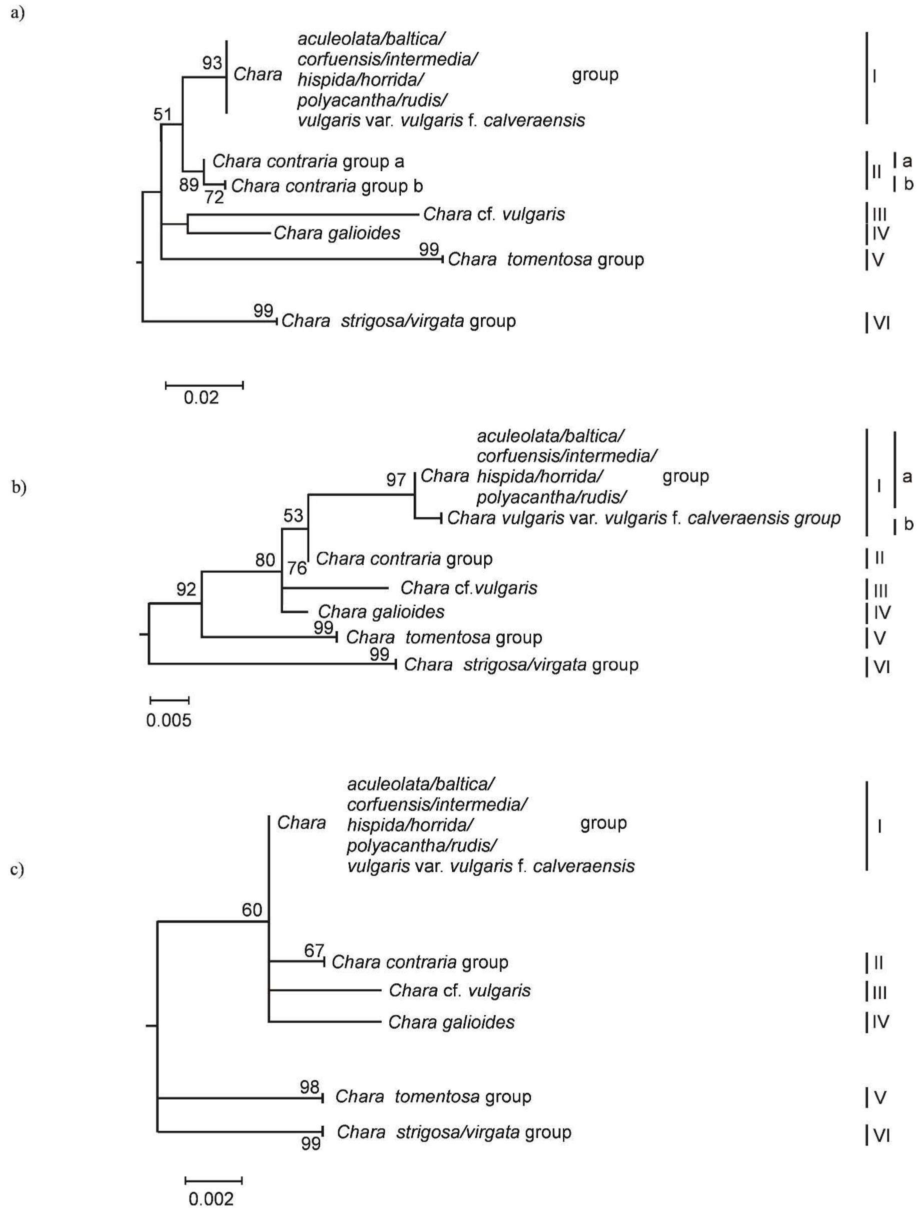
724

725

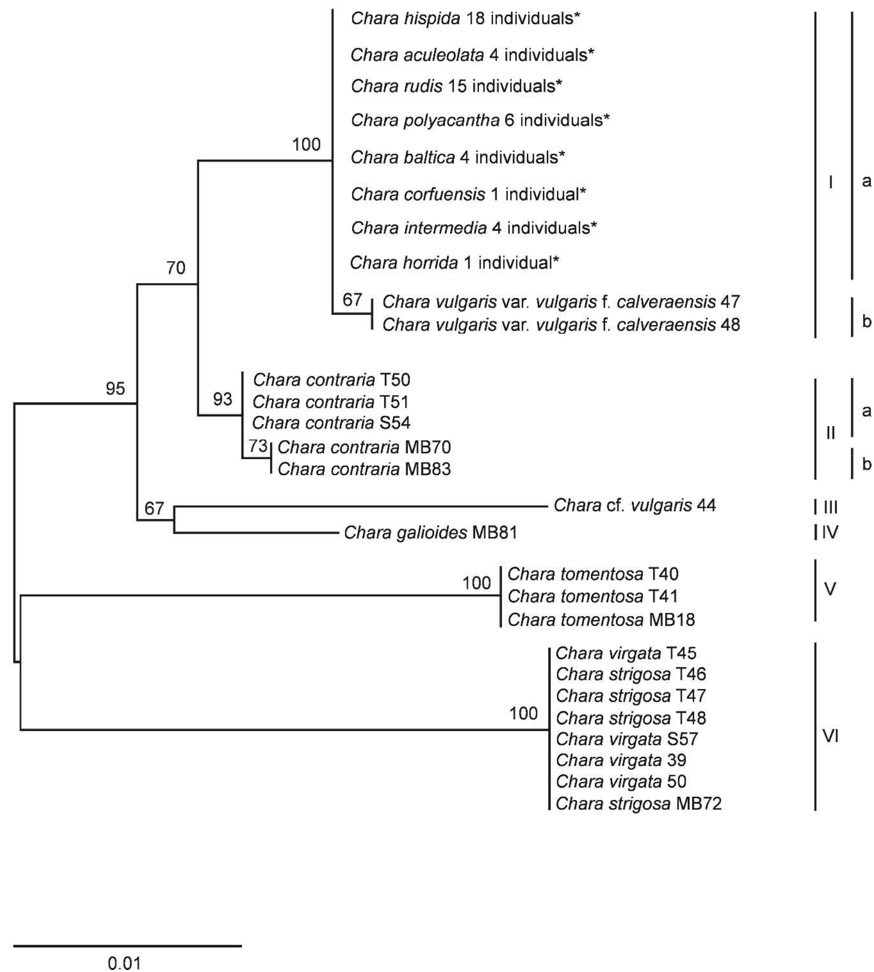
726 ~~Appendix A1~~
727 **Supplementary information**
728
729 **Legend to Figure S1**
730 Bootstrapped condensed ML (maximum likelihood) trees of a) ITS2 (78 *Chara* samples, 261
731 positions), b) *matK* (85 *Chara* samples, 565 positions) and c) *rbc*L (84 *Chara* samples, 552
732 positions). In the ML analyses, evolutionary substitution models were evaluated in MEGA
733 version 6 (Tamura et al.~~,~~ 2013). The method selected T92 +G as the best-fitting evolutionary
734 model for *matK*, and T92 for ITS2 and *rbc*L, respectively. Phylogenetic trees were
735 constructed separately for each marker using the ML algorithm in MEGA version 6.
736 Bootstrap analyses were calculated by ML (1000 replicates) and values above 50 % are
737 shown in the trees. Scale bars indicate 1% (ITS2 and *matK*) and 0.2% (*rbc*L) sequence
738 divergence, respectively.



254x190mm (96 x 96 DPI)



Bootstrapped condensed ML (maximum likelihood) tree of a) ITS2, b) matK and c) rbcL sequences of 73 Chara samples from which all three loci were successfully recovered; bootstrap values above 50 % are shown in the tree. The clusters belong to the following subsections according to Wood and Imahori (1965): I = Hartmania, II = Chara, III = Chara, IV = Grovesia, V = Chara, VI = Grovesia. To improve readability, outgroups are not shown.



Concatenated ML (maximum likelihood) tree of ITS2, matK and rbcL sequences of 73 Chara samples. Bootstrap values above 50 % are shown in the tree. The bar indicates 1% sequence divergence. *= for sample ID see Table 1. The clusters belong to the following subsections according to Wood and Imahori (1965): I = Hartmania, II = Chara, III = Chara, IV = Grovesia, V = Chara, VI = Grovesia.

Table 1. List of 91 *Chara* individuals used in the present study. All samples were taken from herbarium collections, except those marked as "silica gel". NHM = Natural History Museum, University of Oslo; NIVA = Norwegian Institute for Water Research

Identification	author	subsection sensu Wood and Imahori (1965)	comment	BOLD Sample ID	Field ID	coll. year	country	silica gel (1=yes)	matK re-covered (1=yes)	ITS2 re-covered (1=yes)	rbcL re-covered (1=yes)	voucher specimen
Chara aculeolata	Kütz. in Rchb. 1832	Hartmania		CHARA_A_00038	T38	2007	Norway		1	1	1	NHM
Chara aculeolata				CHARA_A_00003	T3	2012	Norway		1	1	1	NHM
Chara aculeolata				CHARA_A_00037	T37	2010	Norway		1	1	1	NHM
Chara aculeolata				CHARA_A_00039	T39	2010	Norway		1	1	1	NHM
Chara aspera	Willd. 1809	Grovesia		CHARA_A_00049	T49	2011	Norway		1	0	1	NHM
Chara aspera				CHARA_A_00074	MB10	2000	Germany		1	0	1	NIVA
Chara aspera				CHARA_A_00075	MB13	2000	Germany		1	0	0	NIVA
Chara aspera				CHARA_A_00077	MB23	2005	Sweden		1	0	1	NIVA
Chara aspera				CHARA_A_00085	MB 67	2005	UK		1	0	1	NIVA
Chara baltica	Bruzelius 1824	Hartmania		CHARA_A_00082	MB39	2004	Greece		0	1	1	NIVA
Chara baltica				CHARA_A_00078	MB34	2004	France		1	1	1	NIVA
Chara baltica				CHARA_A_00035	T35	2010	Norway		1	1	1	NHM
Chara baltica				CHARA_A_00036	T36	2010	Norway		1	1	1	NHM
Chara baltica				CHARA_A_00084	MB47	2004	Sweden		1	0	0	NIVA
Chara baltica				CHARA_A_00083	MB43	2002	Sweden		0	1	1	NIVA
Chara baltica				CHARA_A_00081	MB37	2005	Germany		1	1	1	NIVA
Chara baueri	A. Br. 1847	Braunia		CHARA_A_00052	S29	2011	Germany		1	0	1	NIVA
Chara contraria	A. Br. ex Kütz. 1845 s. str.	Chara		CHARA_A_00092	MB 83	2000	Germany		1	1	1	NIVA
Chara contraria				CHARA_A_00050	T50	2011	Norway		1	1	1	NHM
Chara contraria				CHARA_A_00054	S54	2006	Canada		1	1	1	NIVA
Chara contraria			var. *hispidula*	CHARA_A_00087	MB 70	2001	Austria		1	1	1	NIVA
Chara contraria				CHARA_A_00051	T51	2009	Norway		1	1	1	NHM
Chara corfuensis	(J. Gr. Ex Fil.) R.D.W. 1965	Hartmania		CHARA_A_00055	S55	2006	Greece		1	1	1	NIVA
Chara fibrosa	Ag. ex Bruz. 1824	Agardhia		CHARA_A_00053	S51	2006	Canada		1	0	1	NIVA
Chara galioides	De Candolle 1813	Grovesia		CHARA_A_00091	MB 81	2001	France		1	1	1	NIVA
Chara galioides				CHARA_A_00090	MB 80	2001	France		1	0	0	NIVA
Chara hispida	(L.) Hartm. 1820	Hartmania		CHARA_A_00001	T1	2005	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00002	T2	2003	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00005	T5	2012	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00006	T6	2012	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00009	T9	2012	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00010	T10	2012	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00011	T11	2002	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00012	T12	1998	Norway		0	1	0	NHM
Chara hispida				CHARA_A_00013	T13	2011	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00014	T14	2011	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00015	T15	1995	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00016	T16	2003	Norway		1	1	0	NHM
Chara hispida				CHARA_A_00017	T17	2010	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00018	T18	2009	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00019	T19	2010	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00020	T20	2010	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00021	T21	2002	Norway		1	1	1	NHM
Chara hispida				CHARA_A_00070	49	2012	Germany	1	1	1	1	NIVA
Chara hispida				CHARA_A_00086	MB 68	2005	Germany		1	1	1	NIVA
Chara hispida				CHARA_A_00095	MB 87	2001	Germany		1	1	1	NIVA
Chara horrida	Wahlst. 1862	Hartmania		CHARA_A_00079	MB35	2005	Sweden		1	1	1	NIVA
Chara intermedia	A. Br. in Br., Rab. and Stiz. 1859	Hartmania		CHARA_A_00094	MB 86	2001	Germany		1	1	1	NIVA
Chara intermedia				CHARA_A_00072	MB2	2004	Germany		1	1	1	NIVA
Chara intermedia				CHARA_A_00093	MB 85	2003	Sweden		1	1	1	NIVA
Chara intermedia				CHARA_A_00063	35	2012	Poland	1	1	1	1	NIVA
Chara polyacantha	A. Br. in Br., Rab. and Stiz. 1859	Hartmania		CHARA_A_00073	MB3	2004	Germany		1	1	1	NIVA
Chara polyacantha				CHARA_A_00057	5	2008	Spain	1	1	1	1	NIVA
Chara polyacantha				CHARA_A_00022	T22	2008	Norway		1	1	1	NHM
Chara polyacantha				CHARA_A_00023	T23	2009	Norway		0	0	0	NHM
Chara polyacantha				CHARA_A_00064	37	2012	Poland	1	1	1	1	NIVA
Chara polyacantha				CHARA_A_00065	38	2012	Poland	1	1	1	1	NIVA
Chara polyacantha				CHARA_A_00080	MB36	2005	Sweden		1	1	1	NIVA
Chara rudis	A. Br. in Leonhardi 1882	Hartmania		CHARA_A_00004	T4	2012	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00028	T28	2008	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00061	28	2010	Norway	1	1	1	1	NIVA
Chara rudis				CHARA_A_00024	T24	2010	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00007	T7	2012	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00027	T27	2009	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00029	T29	2008	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00025	T25	2010	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00034	T34	2011	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00033	T33	2011	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00032	T32	2011	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00031	T31	2011	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00026	T26	2010	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00030	T30	2008	Norway		1	1	1	NHM
Chara rudis				CHARA_A_00008	T8	2012	Norway		1	1	1	NHM
Chara strigosa	A. Braun 1847	Grovesia		CHARA_A_00089	MB 72	2003	Germany		1	1	1	NIVA
Chara strigosa				CHARA_A_00047	T47	2010	Norway		1	1	1	NHM
Chara strigosa				CHARA_A_00048	T48	2010	Norway		1	1	1	NHM
Chara strigosa				CHARA_A_00088	MB 71	2002	Germany		0	1	1	NIVA
Chara strigosa				CHARA_A_00046	T46	2011	Norway		1	1	1	NHM
Chara tomentosa	L. 1753	Chara		CHARA_A_00041	T41	2011	Norway		1	1	1	NHM
Chara tomentosa				CHARA_A_00040	T40	2010	Norway		1	1	1	NHM
Chara tomentosa				CHARA_A_00042	T42	1969	Norway		0	0	0	NHM
Chara tomentosa				CHARA_A_00076	MB18	2005	Sweden		1	1	1	NIVA
Chara virgata	Kütz. 1834	Grovesia		CHARA_A_00071	50	2012	Germany	1	1	1	1	NIVA
Chara virgata				CHARA_A_00066	39	2012	Finland	1	1	1	1	NIVA
Chara virgata				CHARA_A_00056	S57	2012	Norway		1	1	1	NIVA
Chara virgata				CHARA_A_00044	T44	2010	Norway		1	0	1	NHM
Chara virgata				CHARA_A_00045	T45	2008	Norway		1	1	1	NHM
Chara virgata				CHARA_A_00043	T43	2011	Norway		1	0	1	NHM
Chara vulgaris	R.D.W. 1965	Chara	var. *vulgaris* f. *calveraensis*	CHARA_A_00068	47	2012	Argentina	1	1	1	1	NIVA
Chara vulgaris	R.D.W. 1965		var. *vulgaris* f. *calveraensis*	CHARA_A_00069	48	2012	Argentina	1	1	1	1	NIVA
Chara vulgaris	L. 1753		c.f.	CHARA_A_00067	44	2012	Argentina	1	1	1	1	NIVA

Table 2. *Chara* sequences used to design *Chara*-specific *matK* and ITS2 primers

Marker	Species	GenBank Accession
ITS2	*C. foliolosa*	HQ380620
	C. hydropitys	HQ380626
	C. haitensis	HQ380624
	C. rusbyana	HQ380627
	C. zeylanica	HQ380634
matK	*C. connivens*	AY170442
	C. globularis	AY170443
	C. longifolia	AY170444
	C. polycantha	AY170445
	C. vulgaris	108773196

Table 3. Primers used in this study

Primer	Sequences	Reference
Chara_ITSF2	CCCCCTTCGATTTTGAAGTT	This study
Chara_ITSR2	ACATCCCCGATTGCCAAC	This study
Chara_matKF2	GAACGAATCCGTGATAAAAGC	This study
Chara_matKR2	CTTCGGCCTTTCAAAAAGAA	This study
rbcLa-F	ATGTCACCACAAACAGAGACTAAAGC	Levin et al., 2003
rbcLa-R	GTAAAATCAAGTCCACCRCG	Kress and Erickson, 2007

Table 4. Number of recovered *matK*, ITS2 and *rbc*L sequences for 91 *Chara* individuals

	matK	ITS2	rbcL
total			
number of samples	91	91	91
number of recovered sequences	85	78	84
sequence recovery (%)	*93*	*86*	*92*
silica gel			
number of samples	11	11	11
number of recovered sequences	11	11	11
sequence recovery (%)	*100*	*100*	*100*
herbarium			
number of samples	80	80	80
number of recovered sequences	74	67	73
sequence recovery (%)	*93*	*84*	*91*