J. Phycol. **51,** 367–380 (2015)

DOI: 10.1111/jpy.12282

DNA BARCODING THE GENUS *CHARA*: MOLECULAR EVIDENCE RECOVERS FEWER TAXA THAN THE CLASSICAL MORPHOLOGICAL APPROACH[1]

Susanne C. Schneider[2]

Norwegian Institute for Water Research, Gaustadalleen 21, Oslo 0349, Norway

Anuar Rodrigues

Biodiversity Institute of Ontario, University of Guelph, 50 Stone Road East, Guelph, ON, Canada

Therese Fosholt Moe, and Andreas Ballot

Norwegian Institute for Water Research, Gaustadalleen 21, Oslo 0349, Norway

Charophytes (Charales) are benthic algae with a complex morphology. They are vulnerable to ecosystem changes, such as eutrophication, and are red-listed in many countries. Accurate identification of *Chara* species is critical for understanding their diversity and for documenting changes in species distribution. Species delineation is, however, complicated, because of high phenotypic plasticity. We used barcodes of the ITS2, *matK* and *rbc*L regions to test if the distribution of barcode haplotypes among individuals is consistent with species boundaries as they are currently understood. The study included freshly collected and herbarium material of 91 specimens from 10 European countries, Canada and Argentina. Results showed that herbarium specimens are useful as a source of material for genetic analyses for aquatic plants like *Chara*. *rbc*L and *matK* had highest sequence recoverability, but *rbc*L had a somewhat lower discriminatory power than ITS2 and *matK*. The tree resulting from the concatenated data matrix grouped the samples into six main groups contrary to a traditional morphological approach that consisted of 14 different taxa. A large unresolved group consisted of *C. intermedia*, *C. hispida*, *C. horrida*, *C. baltica*, *C. polyacantha*, *C. rudis*, *C. aculeolata*, and *C. corfuensis*. A second unresolved group consisted of *C. virgata* and *C. strigosa*. The taxa within each of the unresolved groups shared identical barcode sequences on the 977 positions of the concatenated data matrix. The morphological differences of taxa within both unresolved groups include the number and length of spine cells, stipulodes, and bract cells. We suggest that these morphological traits have less taxonomic relevance than hitherto assumed.

Key index words: **algae; barcode; Charales charophyte; ITS; matK; plant; rbcL; taxonomy**

***Abbreviations*: AFLP, Amplified Fragment Length Polymorphism; CCDB, Canadian Centre for DNA Barcoding; ITS2, nuclear ribosomal internal transcribed spacer region 2; *matK*, plastid-encoded protein-coding gene; *rbc*L, plastid-encoded gene coding for the large subunit of RUBISCO**

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, that is, 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 Joye 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 and Melzer 2003) and lakes (Stelzer et al. 2005), such that accurate species identification is of high practical relevance for ecosystem assessment and management.

[1]Received 22 May 2014. Accepted 25 November 2014.
[2]Author for correspondence: e-mail susi.schneider@niva.no.
Editorial Responsibility: O. De Clerck (Associate Editor)

Species delineation of charophytes is commonly based on morphological traits of the plant thallus. This is, however, more complicated than it might seem, because (i) there is considerable overlap in morphological characteristics used to discriminate species (Boegle et al. 2007), and (ii) phenotypic plasticity in charophytes may be environmentally induced, for example, 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 amplified fragment length polymorphism (AFLP), 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 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 because of 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 analyzed 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 (i) design ITS2 and *matK* primers for the genus *Chara*, (ii) solve taxonomic problems, (iii) 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, (iv) test if results obtained from ITS2, *matK* and *rbc*L are consistent with each other, or if one marker discriminates better than the others, and (v) test if the distribution of barcode haplotypes among individuals is consistent with species boundaries as they are currently understood.

MATERIALS AND METHODS

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

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

Many *Chara* taxa have been variously recognized as species, varieties, or forms, and there is little consensus about appropriate rank among different flora treatments. The two most widely applied taxonomic concepts are those of Wood and Imahori (1965) and Krause (1997). While 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.

TABLE 1. List of 91 *Chara* individuals used in the present study.

Identification	Author	Subsection sensu Wood and Imahori (1965)	Comment	BOLD Sample ID	Field ID	Coll. year	Country	Silica gel (1 = yes)	matK recovered (1 = yes)	ITS2 recovered (1 = yes)	rbcL recovered (1 = yes)	Voucher specimen
Chara aculeolata	Kützing in Reichenbach 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	Willdenow 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. Braun 1847	Braunia		CHARA_A_00052	S29	2011	Germany		1	0	1	NIVA
Chara contraria	A. Braun ex Kützing 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. Groves Ex Filarszky) R.D.W. 1965	Hartmania		CHARA_A_00055	S55	2006	Greece		1	1	1	NIVA
Chara fibrosa	Agardh ex Bruzelius 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	(Linnaeus) Hartman 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

(continued)

TABLE 1. (continued)

Identification	Author	Subsection sensu Wood and Imahori (1965)	Comment	BOLD Sample ID	Field ID	Coll. year	Country	Silica gel (1 = yes)	matK recovered (1 = yes)	ITS2 recovered (1 = yes)	rbcL recovered (1 = yes)	Voucher specimen
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	Wahlstedt 1862	Hartmania		CHARA_A_00079	MB35	2005	Sweden		1	1	1	NIVA
Chara intermedia	A. Braun in Braun Rabenhorst and Stizenberger 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. Braun in Braun Rabenhorst and Stizenberger 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. Braun 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	Linnaeus 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	Kützing 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

(continued)

TABLE 1. (continued)

Identification	Author	Subsection sensu Wood and Imahori (1965)	Comment	BOLD Sample ID	Field ID	Coll. year	Country	Silica gel (1 = yes)	matK recovered (1 = yes)	ITS2 recovered (1 = yes)	rbcL recovered (1 = yes)	Voucher specimen
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.Wood 1965	Chara	var. *vulgaris* f. *calveraensis*	CHARA_A_00068	47	2012	Argentina	1	1	1	1	NIVA
Chara vulgaris	R.D.Wood 1965		var. *vulgaris* f. *calveraensis*	CHARA_A_00069	48	2012	Argentina	1	1	1	1	NIVA
Chara vulgaris	Linnaeus 1753		c.f.	CHARA_A_00067	44	2012	Argentina	1	1	1	1	NIVA

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.

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 to be consistent with Krause's (1997) taxonomic concept; and (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, 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, Centerville, MA, USA). Contigs were aligned using the MUSCLE multiple sequence alignment algorithm (Edgar 2004) as implemented in CodonCode Aligner. These initial alignments were created to compare contigs generated from the various specimens and aided the identification and correction of base calling errors following the examination of trace files. Specific for ITS2 amplicons, BLAST was utilized in order to determine whether any of the sequences produced were of fungal contaminants.

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

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

Sequences were aligned using Align (version 03/2007) MS Windows-based manual sequence alignment editor (SequentiX - Digital DNA Processing, Klein Raden Germany) to obtain DNA sequence alignments, which were then corrected manually. Segments with highly variable and ambiguous regions and gaps making proper alignment impossible were excluded from the analyses. A *mat*K set containing 292 positions, an ITS2 set containing 183, and a *rbc*L set containing 502 positions was used. *Chara longifolia* (AY170444), *Chara*

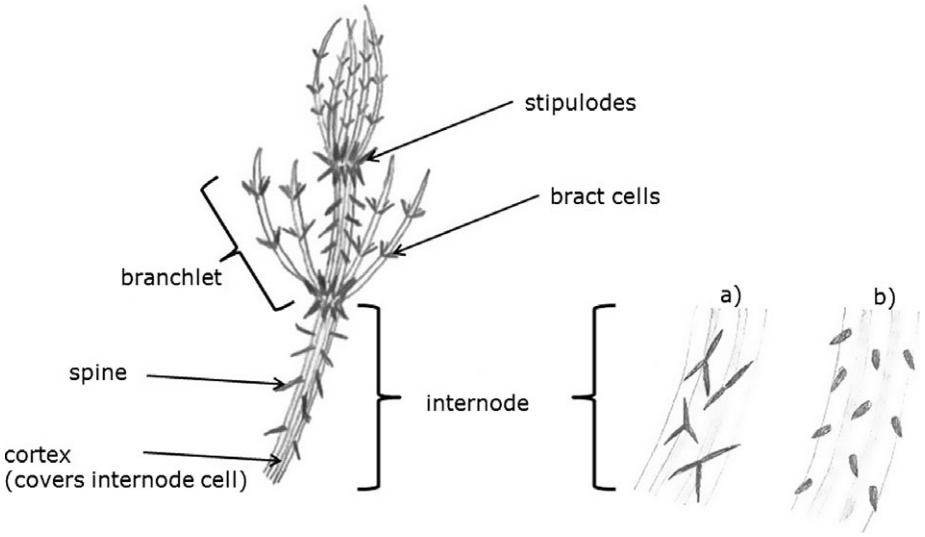

FIG. 1. Schematic drawing of the uppermost part of a *Chara* specimen; the cortex can be (i) haplostichous (number of cortex cell rows corresponds to the number of branchlets), diplostichous (twice as many cortex cell rows as the number of branchlets), or triplostichous (three times as many cortex cell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

foliolosa (HQ380618) and *Chara foliolosa* (HQ380452) were employed as outgroup taxa in the *matK*, ITS2 and *rbc*L tree, respectively. Data sets were analyzed using the maximum likelihood (ML) algorithm in MEGA version 6 (Tamura et al. 2013). In the ML analyses, evolutionary substitution models were evaluated in MEGA version 6. The method selected the same best-fitting evolutionary model (T92) for the three markers (*matK*, ITS2 and *rbc*L). ML analyses were performed with 1,000 bootstrap replicates in MEGA version 6 (Tamura et al. 2013).

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

In the tree generated using the combined *matK*-ITS2-*rbc*L data set, no outgroup was used. Phylogenetic inference was based on ML, and analyses were performed with Treefinder (Jobb 2011) with three partitions. Models and parameters proposed by Treefinder under AICc criteria were as follows: *matK* (292 bases; model HKY), *rbc*L (502 bases; model HKY), ITS2 (183 bases; model HKY). To provide support of relationships, bootstrap analyses were calculated by ML (1,000 replicates) criteria with Treefinder.

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

RESULTS

Sequence recoverability. Sequence recoverability was highest for *matK* and *rbc*L, and lowest for ITS2

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

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

Taxonomic uncertainties. We corrected one misidentification in our dataset, where a *C. contraria* with unusually long spine cells had mistakenly been identified as *C. polyacantha* (field ID MB 70; Table 1). This sample clearly clustered to other *C. contraria* individuals, and this result was consistent among all three analyzed loci.

In addition, the three specimens of *C. vulgaris* from Argentina were found to cluster to two different groups. While two accessions, determined by their elongated stipulodes as *C. vulgaris* var. *vulgaris* f. *calveraensis* (field IDs 47 and 48) clustered within a large group containing nine different taxa (Figs. 2 and 3), the third sample (field ID 44) formed its own branch (Figs. 2 and 3).

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

Tree topology for the three loci was similar, though not identical. In each of the trees, *C. contraria* (in case of *rbc*L, together with *C.* c.f. *vulgaris* and *C. galioides*) was the group most closely related to the large cluster I containing nine taxa, while *C. strigosa/virgata* (in case of *rbc*L, together with *C. tomentosa*) was most distant (Fig. 2). In total, ITS2, *rbc*L and *matK* were consistent in their fundamental results. ITS2 and *matK* each discriminated one subgroup (Ib and IIb, respectively; Fig. 2) which the other loci did not recover. Apart from that, discriminatory power of the three investigated loci did not differ. The concatenated tree resulting from the analysis of the combined plastid and nuclear sequences resolved the same six main groups as the individual analyses. Apart from the two groups consisting of one individual each (*C.* c.f. *vulgaris*, *C. galioides*), the remaining four groups were each supported by bootstrap values of ≥93% (Fig. 3). Tree topology was consistent with the results from the separate analyses. The subgroup containing two individuals of *C. vulgaris* var. *vulgaris* f. *calveraensis* differentiated by *matK* (Fig. 2b), and the subgroup containing two samples of *C. contraria* from Austria and Germany differentiated by ITS2 (Fig. 2a) were both represented in the concatenated tree (Fig. 3), such that overall resolution of the concatenated tree was slightly higher than the trees resulting from each individual analysis.

DISCUSSION

Sequence recoverability. A multitude of factors influences recoverability of DNA barcodes, ranging from careless preparation of samples, sample contamination, and age of samples to unsatisfactory primer design or inadequate amplicon length. For our samples, sequence recoverability for the freshly collected and silica gel dried samples was 100% for all three loci, while recoverability for the herbarium specimen ranged from 93% (*matK*) to 84% (ITS2). These values are considerably higher than in some other analyses for plants (de Vere et al. 2012) and

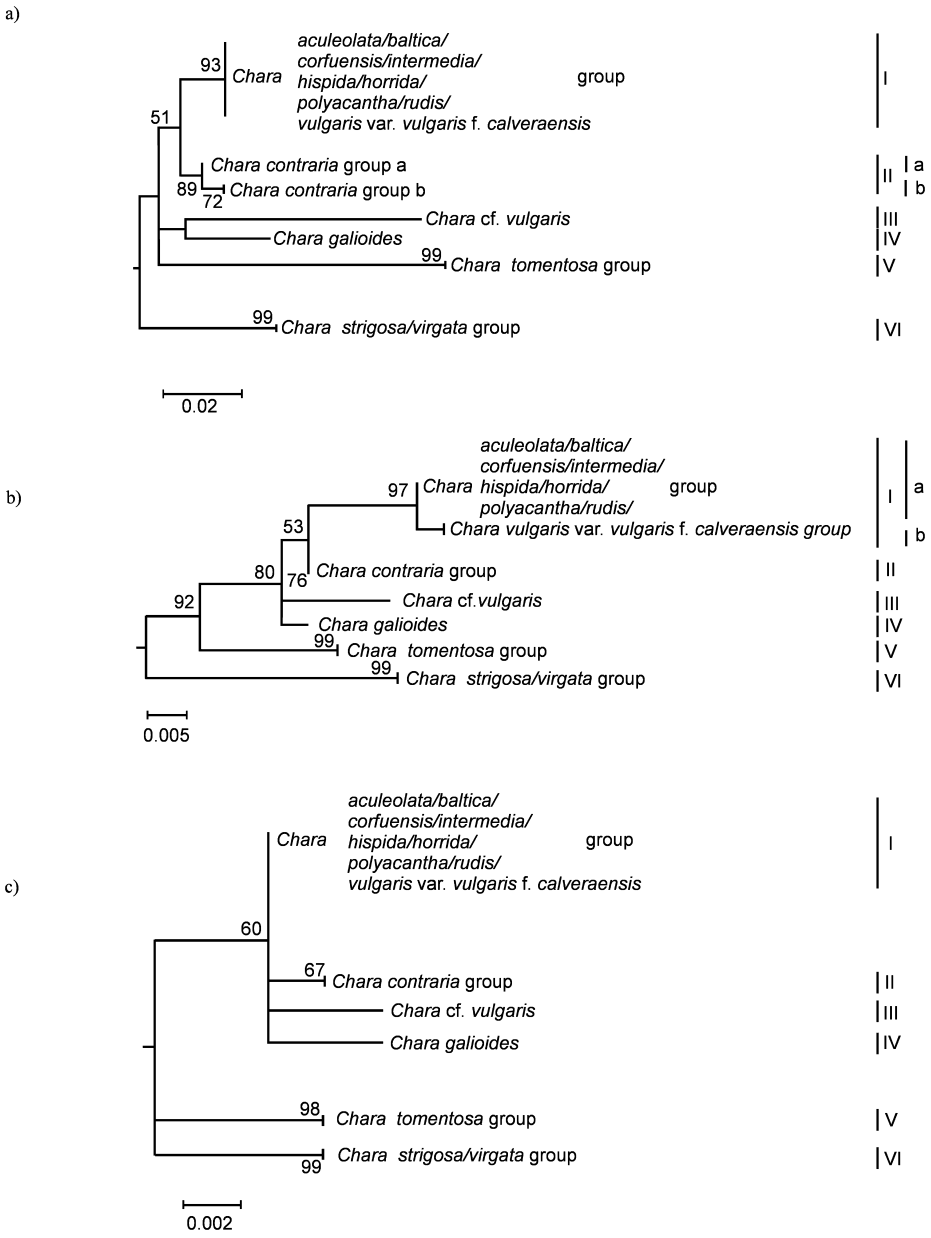

FIG. 2. Bootstrapped condensed maximum likelihood (ML) tree of (a) ITS2, (b) *matK* and (c) *rbc*L sequences of 73 *Chara* samples from which all three loci were successfully recovered; bootstrap values above 50% are shown in the tree. The clusters belong to the following subsections according to Wood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

algae (Kucera and Saunders 2012, Pérez et al. 2014), but correspond well with results on arctic vascular plants (Kuzmina et al. 2012). They may simply be explained by the younger age of the herbarium specimens we used (our oldest sample was from 1969, while de Vere et al. (2012) managed to barcode a sample from 1868).

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

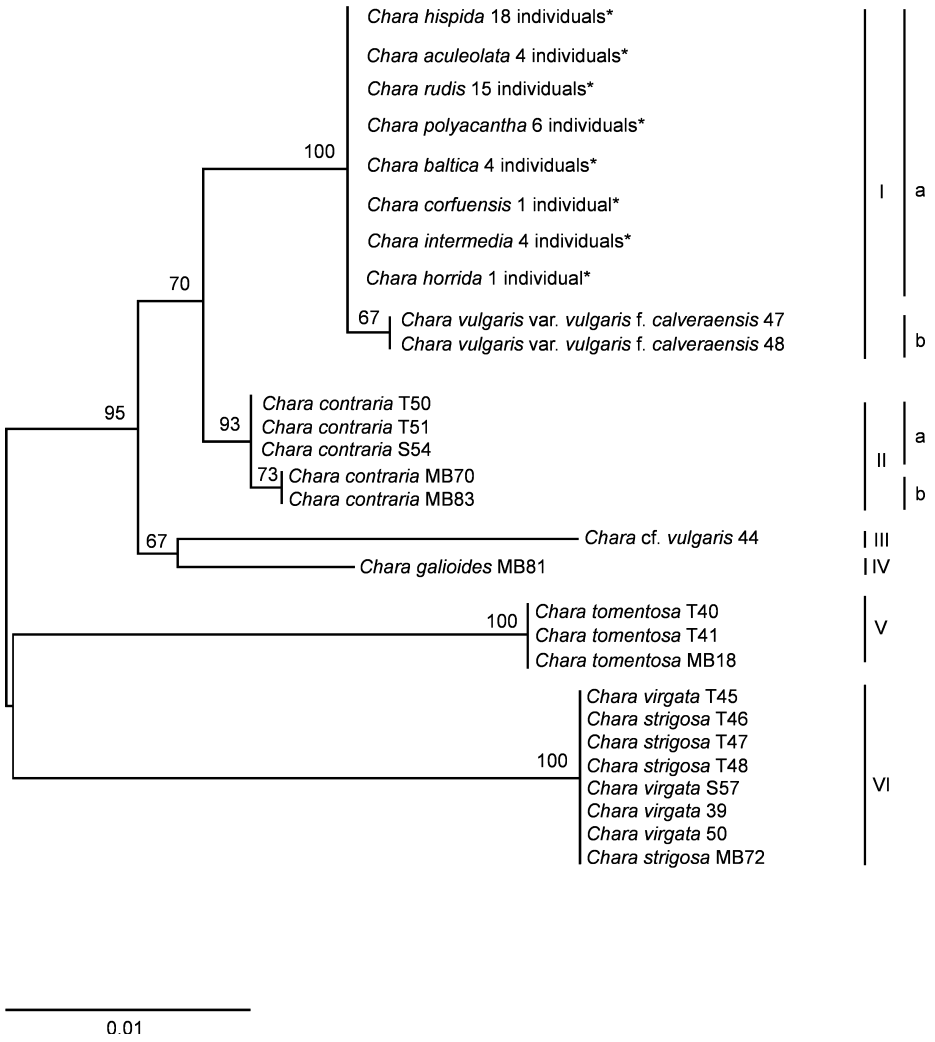

FIG. 3. Concatenated maximum likelihood (ML) tree of ITS2, *matK* and *rbc*L sequences of 73 *Chara* samples. Bootstrap values above 50% are shown in the tree. The bar indicates 1% sequence divergence. *= for sample ID see Table 1. The clusters belong to the following subsections according to Wood and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

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 with 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 Plant Working Group 2009, Hollingsworth et al. 2011). Nevertheless, a greater discriminatory power of the entire nrITS compared with 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, ITS2 had lower sequence recoverability (see above). The choice of *rbc*L+*matK* as a core barcode for plants seems therefore suitable also for *Chara*.

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

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

The taxonomic concept applied by Wood and Imahori (1965) is based on the assumption that a close phylogenetic relationship exists among the members of each subsection. However, members of the subsections Chara and Grovesia did not form monophyletic groups (Fig. 3), thus contradicting Wood and Imahori's (1965) assumption. This notwithstanding, all samples belonging to the subsection Hartmania (Table 1) indeed clustered together (cluster I, Fig. 3), which supports Wood and Imahori's assumption of a close phylogenetic relationship of the taxa in this group. However, both individuals of *C. vulgaris* var. *vulgaris* f. *calveraensis* (field ID 47, 48; Table 1) also clustered to the Hartmania section (Fig. 3). This result was consistent among all three analyzed loci (Fig. 2). Wood and Imahori (1965) separate the subsection Chara (into which they placed *C. vulgaris* var. *vulgaris* f. *calveraensis*) from the subsection Hartmania by their spine cells: individuals of the subsection Chara have solitary or rarely geminate spine cells, while individuals of the subsection Hartmania have predominantly fasciculate spine cells. Since re-examination of the voucher specimens gave no indication of a possible misidentification, we suggest that *C. vulgaris* var. *vulgaris* f. *calveraensis* belongs to the subsection Hartmania instead of the subsection Chara. This further indicates that the number of spine cells (solitary, geminate or fasciculate, i.e., the criterion which was used by Wood and Imahori (1965) to separate the subsection Hartmania from the subsection Chara) may not be useful for inferring phylogenetic relationships.

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

Consistency between barcode haplotypes and morphological species boundaries. Following a traditional morphological approach, the tree produced using the concatenated data matrix (Fig. 3) consisted of indi-

viduals from 14 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 vs. fasciculate), cortication (tylachanthous vs. aulacanthous), stipulodes (two rows vs. multiple rows), bract cells (elongated vs. not elongated), and habitat (fresh water vs. 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, 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 one 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, although 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 epigenetic variation, such as DNA methylation (Cubas et al. 1999, Zhang et al. 2013). Verhoeven et al. (2010) showed that stress, for example 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 water habitats, while the other taxa in cluster I typically occur in freshwater habitats) or Ca-concentration (*C. virgata*: Calcium poor habitats, *C. strigosa*: Calcium rich habitats; Rey-Boissezon and Auderset Joye 2015) may also be related to epigenetic variation. A third explanation for phenotypic modifications in the absence of differences in DNA barcodes may be polyploidy (Schranz and Osborn 2004). However, recent evidence suggests that epigenetic rather than genetic factors may explain phenotypic divergence between plant populations of different ploidy (Rois et al. 2013).

CONCLUSIONS

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

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

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."

Auderset Joye, D. & Rey-Boissezon, A. 2015. Will charophyte species increase or decrease their distribution in a changing climate? *Aquat. Bot.* 120:73–83. doi:10.1016/j.aquabot.2014.05.003.

Auderset Joye, D. & Schwarzer, A. 2012. Rote Liste Armleuchteralgen. Gefährdete Arten der Schweiz, Stand 2010. Umwelt-Vollzug Nr. 1213. Bundesamt für Umwelt, Bern und Laboratoire d'écologie et de biologie aquatique (LEBA) der Universität Genf, Bern, 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.* 58:1682–93.

Belton, G. S., van Reine, W. F. P., Huisman, J. M., Draisma, S. G. A. & Gurgel, C. F. D. 2014. Resolving phenotypic plasticity and species designation in the morphologically challenging *Caulerpa racemosa-peltata* complex (Caulerpaceae, Chlorophyta). *J. Phycol.* 50:32–54.

Blindow, I. 1992. Decline of charophytes during eutrophication – comparison with angiosperms. *Freshwater Biol.* 28:9–14.

Blume, M., Blindow, I., Dahlke, S. & Vedder, F. 2009. Oospore variation in closely related *Chara* taxa. *J. Phycol.* 45:995–1002.

Boegle, M. G., Schneider, S., Mannschreck, B. & Melzer, A. 2007. Differentiation of *Chara intermedia* and *C. baltica* compared to *C. hispida* based on morphology and amplified fragment length polymorphism. *Hydrobiologia* 586:155–66.

Boegle, M. G., Schneider, S. C., Melzer, A. & Schubert, H. 2010b. Distinguishing *Chara baltica*, *C. horrida* and *C. liljebladii* – conflicting results from analysis of morphology and genetics. *Charophytes* 2:53–8.

Boegle, M. G., Schneider, S. C., Schubert, H. & Melzer, A. 2010a. *Chara baltica* Bruzelius 1824 and *Chara intermedia* A. Braun 1859 - distinct species or habitat specific modifications? *Aquat. Bot.* 93:195–201.

CBOL Plant Working Group 2009. A DNA Barcode for land plants. *Proc. Natl. Acad. Sci. USA* 106:12794–7.

China Plant BOL Group 2011. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. *Proc. Natl. Acad. Sci. USA* 108:19641–6.

Cubas, P., Vincent, C. & Coen, E. 1999. An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* 401:157–61.

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792–7.

Hilu, K. W. & Liang, H. 1997. The *matK* gene: sequence variation and application in plant systematics. *Am. J. Bot.* 84:830–9.

Hollingsworth, P. M. 2011. Refining the DNA barcode for land plants. *Proc. Natl. Acad. Sci. USA* 108:19451–2.

Hollingsworth, P. M., Graham, S. W. & Little, D. P. 2011. Choosing and using a plant DNA barcode. *PLOS ONE* 6:e19254.

Holzhausen, A., Nowak, P., Niedrig, C., Feike, M. & Schubert, H. 2015. Morphometry of *Chara aspera*, *C. canescens*, *C. baltica* var. *baltica*, *C. baltica* var. *liljebladii* and *C. intermedia* oospores: local variation versus taxonomic differences. *Aquat. Bot.* 120:60–6. doi:10.1016/j.aquabot.2014.07.001.

IUCN (International Union for Conservation of Nature and Natural Resources) 2014. Guidelines for Using the IUCN Red List Categories and Criteria, version 11. Available at: http://jr.iucnredlist.org/documents/RedListGuidelines.pdf (last accessed February 3, 2014).

Ivanova, N. V., Fazekas, A. & Hebert, P. D. N. 2008. Semi-automated, membrane based protocol for DNA isolation from plants. *Plant Mol. Biol. Rep.* 26:186–98.

Ivanova, N., Kuzmina, M. & Fazekas, A. J. 2011. CCDB Protocols (2011). Manual Protocol Employing Centrifugation: Glass Fiber Plate DNA Extraction Protocol for Plants, Fungi, Echinoderms and Mollusks. Available at: http://www.ccdb.ca/docs/CCDB_DNA_Extraction-Plants.pdf (last accessed August 15, 2012).

Jobb, G. 2011. *Treefinder version of March 2011*. Munich, Germany. Distributed by the author at www.Treefinder.de.

Karol, K., McCourt, R., Cimino, M. & Delwiche, C. F. 2001. The closest living relatives of land plants. *Science* 294:2351–3.

Kato, S., Sakayama, H., Sano, S., Kasai, F., Watanabe, M. M., Tanaka, J. & Nozaki, H. 2008. Morphological variation and intraspecific phylogeny of the ubiquitous species *Chara braunii* (Charales, Charophyceae) in Japan. *Phycologia* 47:191–202.

Krause, W. 1997. Charales (Charophyceae). *In* Ettl, H., Gärtner, G., Heynig, H. & Mollenhauer, D. [Eds.]. *Süßwasserflora von Mitteleuropa*, Band 18. Fischer, Jena, pp. 1–202.

Kress, W. J. & Erickson, D. L. 2007. A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. *PLoS ONE* 2:e508.

Kucera, H. & Saunders, G. W. 2012. A survey of Bangiales (Rhodophyta) based on multiple molecular markers reveals cryptic diversity. *J. Phycol.* 48:869–82.

Kufel, L., Biardzka, E. & Strzalek, M. 2013. Calcium carbonate incrustation and phosphorus fractions in five charophyte species. *Aquat. Bot.* 109:54–7.

Kuzmina, M. L., Johnson, K. L., Barron, H. R. & Hebert, P. D. N. 2012. Identification of the vascular plants of Churchill, Manitoba, using a DNA barcode library. *BMC Ecol.* 12:25.

Leliaert, F., Verbruggen, H., Vanormelingen, P., Steen, F., Lopez-Bautista, J. M., Zuccarello, G. C. & De Clerck, O. 2014. DNA- based species delimitation in algae. *Eur. J. Phycol.* 49:179–96.

Levin, R. A., Wagner, W. L., Hoch, P. C., Nepokroeff, M., Pires, J. C., Zimmer, E. A. & Sytsma, K. J. 2003. Family-level relationships of Onagraceae based on chloroplast rbcL and ndhF data. *Am. J. Bot.* 90:107–15.

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

O'Reilly, C. L., Cowan, R. S. & Hawkins, J. A. 2007. Amplified fragment length polymorphism genetic fingerprinting challenges the taxonomic status of the near-endemic species *Chara curta* Nolte ex Kütz. (Characeae). *Bot. J. Linn. Soc.* 155:467–76.

Pérez, W., Hall, J. D., McCourt, R. & Karol, K. G. 2014. Phylogeny of North American *Tolypella* (Charophyceae, Charophyta) based on plastid DNA sequences with a description of *Tolypella ramosissima* sp. nov. *J. Phycol.* 50:776–89.

Rey-Boissezon, A. & Auderset Joye, D. 2015. Habitat requirements of charophytes - evidence of species discrimination through distribution analysis. *Aquat. Bot.* 120:84–91. doi:10.1016/j.aquabot.2014.05.007.

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

Roy, S., Tyagi, A., Shukla, V., Kumar, A., Singh, U. M., Chaudhary, L. B., Datt, B. et al. 2010. Universal plant DNA barcode loci may not work in complex groups: a case study with Indian berberis species. *PLoS ONE* 5:e13674.

Saarela, J. M., Sokoloff, P. C., Gillespie, L. J., Consaul, L. L. & Bull, R. D. 2013. DNA barcoding the Canadian Arctic flora: core plastid barcodes (rbcL plus matK) for 490 vascular plant species. *PLoS ONE* 8:e77982.

Sakayama, H., Hara, Y., Arai, S., Sato, H. & Nozaki, H. 2004a. Phylogenetic analyses of *Nitella* subgenus Tieffallenia (Charales, Charophyceae) using nuclear ribosomal DNA internal transcribed spacer sequences. *Phycologia* 43:672–81.

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

Sarkinen, T., Staats, M., Richardson, J. E., Cowan, R. S. & Bakker, F. T. 2012. How to open the treasure chest? Optimising DNA extraction from herbarium specimens *PLoS ONE* 7: e43808.

Schneider, S. C., Garcia, A., Martín-Closas, C. & Chivas, A. R. 2015a. The role of charophytes (Charales) in past and present environments: an overview. *Aquat. Bot.* 120:2–6. doi:10.1016/j.aquabot.2014.10.001.

Schneider, S. & Melzer, A. 2003. The trophic index of macrophytes (TIM) – a new tool for indicating the trophic state of running waters. *Int. Rev. Hydrobiol.* 88:49–67.

Schneider, S. C., Pichler, D. E., Andersen, T. & Melzer, A. 2015b. Light acclimation in submerged macrophytes: the roles of plant elongation, pigmentation and branch orientation differ among *Chara* species. *Aquat. Bot.* 120:121–128. doi:10.1016/j.aquabot.2014.05.002.

Schranz, M. E. & Osborn, T. C. 2004. De novo variation in life-history traits and responses to growth conditions of resynthesized polyploid Brassica napus (Brassicaceae). *Am. J. Bot.* 91:174–83.

Seberg, O. & Petersen, G. 2009. How many loci does it take to DNA barcode a *Crocus*? *PLoS ONE* 4:article number e4598.

Sjøtun, K., Fredriksen, S., Heggøy, E., Husa, V., Langangen, A., Lindstrøm, E. A., Moy, F., Rueness, J. & Åsen, P. A. 2010. Cyanophyta, Rhodophyta, Chlorophyta, Ochrophyta. *In* Kålås, J. A., Viken, Å., Henriksen, S. & Skjelseth, S. [Eds.]. *The Norwegian Red List for Species*. Norwegian Biodiversity Information Centre, Norway, pp. 79–86.

Stelzer, D., Schneider, S. & Melzer, A. 2005. Macrophyte based assessment of lakes – a contribution to the implementation of the European Water Framework Directive in Germany. *Int. Rev. Hydrobiol.* 90:223–37.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol. Biol. Evol.* 30:2725–9.

Timme, R. E., Bachvaroff, T. R. & Delwiche, C. F. 2012. Broad phylogenomic sampling and the sister lineage of land plants. *PLoS ONE* 7:e29696.

Urbaniak, J. & Combik, M. 2013. Genetic and morphological data fail to differentiate *Chara intermedia* from *C. baltica*, or *C. polyacantha* and *C. rudis* from *C. hispida*. *Eur. J. Phycol.* 48:253–9.

de Vere, N., Rich, T. C. G., Ford, C. R., Trinder, S. A., Long, C., Moore, C. W., Satterthwaite, D., et al. 2012. DNA barcoding the native flowering plants and conifers of Wales. *PLoS ONE* 7:e37945.

Verhoeven, K. J. F., Jansen, J. J., van Dijk, P. J. & Biere, A. 2010. Stress-induced DNA methylation changes and their heritability in asexual dandelions. *New Phytol.* 185: 1108–18.

Vogel, J., Borner, T. & Hess, W. R. 1999. Comparative analysis of splicing of the complete set of chloroplast group II introns in three higher plant mutants. *Nucleic Acids Res.* 27:3866–74.

Wodniok, S., Brinkmann, H., Glockner, G., Heidel, A., Philippe, H., Melkonian, M. & Becker, B. 2011. Origin of land plants: do conjugating green algae hold the key? *BMC Evol. Biol.* 11:104.

Wood, R. D. & Imahori, K. 1965. *A Revision of the Characeae. First Part: Monograph of the Characeae.* J. Cramer Verlag, Weinheim.

Zhang, Y. Y., Fischer, M., Colot, V. & Bossdorf, O. 2013. Epigenetic variation creates potential for evolution of plant phenotypic plasticity. *New Phytol.* 197:314–22.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web site:

Figure S1. Bootstrapped condensed ML (maximum likelihood) trees of (a) ITS2 (78 *Chara* samples, 261 positions), (b) *matK* (85 *Chara* samples, 565 positions) and (c) *rbc*L (84 *Chara* samples, 552 positions).