Cryptogamie, Algologie, 2010, 31 (4): 487-528

An evaluation of *rbcL*, *tufA*, UPA, LSU and ITS as DNA barcode markers for the marine green macroalgae

Gary W. SAUNDERS[*] *& Hana KUCERA*

Centre for Environmental and Molecular Algal Research, Department of Biology, University of New Brunswick, Fredericton, NB E3B 5A3 Canada

Abstract — The universality and species discriminatory power of the plastid rubisco large subunit (*rbc*L) (considering 5' and 3' fragments independently), elongation factor *tufA*, and universal amplicon (UPA), and the nuclear D2/D3 region of the large ribosomal subunit (LSU) and the internal transcribed spacer of the ribosomal cistron (ITS) were evaluated for their utility as DNA barcode markers for green macroalgae. Excepting low success for ITS, all of these markers failed for the Cladophoraceae. For the remaining taxa, the 3' region of the *rbc*L (*rbc*L-3P) and *tufA* had the largest barcode gaps (difference between maximum intra- and minimum inter-specific divergence). Unfortunately, moderate amplification success (80 % excluding Cladophoraceae) caused, at least in part, by the presence of introns within the *rbc*L-3P for some taxa reduced the utility of this marker as a universal barcode system. The *tufA* marker, on the other hand, had strong amplification success (95% excluding the Cladophoraceae) and no introns were uncovered. We thus recommend that *tufA* be adopted as the standard marker for the routine barcoding of green marine macroalgae (excluding the Cladophoraceae). During this survey we discovered cryptic species in *Acrosiphonia*, *Monostroma*, and *Ulva* indicating that significant taxonomic work remains for green macroalgae.

Chlorophyta / DNA barcoding / green algae / ITS / LSU / *rbc*L / *tufA* / UP

INTRODUCTION

DNA barcoding, identifying organisms based on comparisons of short, standardized DNA sequences as agreed upon by the Consortium for the Barcode of Life (CBOL; http://barcoding.si.edu/), has been championed as a revolutionary system for the identification and discovery of all of the world's eukaryotic organisms (Hebert *et al.,* 2003a; Hebert *et al.,* 2003b). The benefits of such an endeavour are multifold. First, developing a database of sequences against which unknown biological materials can be compared greatly increases the speed at which routine identification can be carried out. Across all taxa, DNA barcoding removes the reliance on morphological characters traditionally used to discriminate species. This is of particular importance for taxonomic groups that have few characters on which to base identification and for which expert taxonomists are required for even routine morphological identifications. Second, during diversity surveys, specimens may be flagged as potentially new species, alerting the taxonomist that further study is required, and ultimately leading to the discovery of new records and/or description of new species. The case of

* Corresponding author: gws@unb.ca

marine macroalgae is an excellent example of the value and strength of DNA barcoding; seaweeds are notoriously difficult to identify due to simple morphologies, phenotypic plasticity and convergent evolution (see Saunders, 2005). DNA barcoding has proven a phenomenal tool that has aided in species identification, discovery of cryptic species, or new records for red (Rhodophyta) and brown (Phaeophyceae) seaweeds (Saunders, 2005; Robba *et al.*, 2006; Kucera & Saunders, 2008; Saunders, 2008; McDevit & Saunders, 2009; Saunders, 2009; Walker *et al.*, 2009; Le Gall & Saunders, 2010).

Whereas DNA barcoding has been advanced for the red and brown seaweeds, the marine green macroalgae (most belong to one of the Bryopsidophyceae, Trebouxiophyceae or Ulvophyceae) are the final group of marine macroalgae for which a DNA barcode marker remains to be developed. There are several requirements when choosing an appropriate marker for DNA barcoding. First, the genetic variability of the marker must be adequate for species level resolution. To this end, species discrimination is considered successful when specimens from a single species cluster together in distance analysis (Hebert *et al.*, 2003a) and the largest intraspecific divergence is less than the smallest interspecific divergence – this difference termed the "barcoding gap" (Meier *et al.*, 2008). Second, and of equal importance, the marker should be universally recoverable across taxa, meaning there is high PCR and sequencing success, ideally with a near-universal set of PCR primers. Given current technologies, it is also of benefit for the DNA barcode marker to be short enough to be sequenced in a single read (<700bp) (see Hollingsworth *et al.,* 2009b). Markers in which intra-individual heterogeneity is common, or which can exhibit microsatellites or single nucleotide runs (particularly of thymine or adenine) are not preferred because these cause problems with direct sequencing of PCR product (Gribble & Anderson, 2007). The ideal marker consists of a highly variable region, which provides for species discrimination, flanked by highly conserved regions from which primers can be designed.

During the development of DNA barcoding, several markers have been proposed for different phyla. In animals, red and brown algae, the 5' end of the cytochrome c oxidase 1 gene (COI-5P) provides resolution at the species level among most groups tested, and has come to be accepted as the barcode marker and is now being applied to biodiversity and taxonomic studies across the globe (e.g., Hebert *et al.*, 2004; Saunders, 2005; Ferri *et al.*, 2009; McDevit & Saunders, 2009). For the green macroalgae, COI-5P was our first marker of interest; however, preliminary investigations failed to generate successful amplification, despite extensive primer testing (Kucera & Saunders, unpublished data). Our failure in designing primers for the successful amplification of green algal COI-5P may be in part attributable to the lack of available mitochondrial sequences; only two published ulvophycean mitochondrial genomes were available at the time of study (Pombert *et al.*, 2004; Pombert *et al.*, 2006). However, the presence of introns within the cytochrome c oxidase 1 of ulvophycean taxa (Watanabe *et al.*, 1998; Pombert *et al.*, 2004; Pombert *et al.*, 2006) may be the largest obstacle to developing the COI-5P as a DNA barcode marker for marine green macroalgae. These difficulties make COI-5P unsuitable for DNA barcoding in green macroalgae, leading us to ultimately abandon it as a potential marker (also see Hall *et al.* (2010) in this volume).

Similarly, difficulties in applying the COI-5P have been encountered in land plants and fungi. In fungi, the presence of introns (Seifert *et al.*, 2007) and lack of species level resolution (Geiser *et al.*, 2007) meant that the COI-5P could not be applied universally. The internal transcribed spacer region (ITS) had been

used for fungal identification for at least a decade before the advent of DNA barcoding in 2003, and is now recommended as the marker for barcoding fungi (see Seifert, 2009). Finding a marker that is both universal and which provides sufficient species discrimination was much more difficult in land plants. After four years of extensive debate, the Plant Working Group members of CBOL have settled on two plastid markers, the rubisco large subunit (*rbc*L) and *matK,* as the DNA barcodes for plants (Hollingsworth *et al.*, 2009b). Since 2005, the debate over which marker (and eventually which combination of makers) would be used for land plant barcoding has produced a wealth of data on a variety of markers and some interesting conclusions. Firstly, a trade-off between universality and species discrimination has affected the search for the plant barcode as there has been no single marker studied to date that can achieve both of these criteria as well as COI-5P does for animals, and therefore, typically two or more markers have been recommended to be used in concert. Secondly, the optimal marker (or marker combination) tends to vary based on the taxonomic group studied (Hollingsworth *et al.*, 2009b).

We selected several markers to investigate as potential DNA barcode markers for the marine green macroalgae. The *rbc*L was an obvious choice for testing since its utility as a DNA barcode has been established among plant groups (Hollingsworth *et al.*, 2009b) and because it has formed the basis of several taxonomic and phylogenetic studies in marine green macroalgae. As an example, in the genus *Ulva* the *rbc*L has been employed extensively to resolve taxonomic issues (e.g., Hayden & Waaland, 2002; Hayden *et al.*, 2003; Hayden & Waaland, 2004). Unfortunately, the presence of introns in the *rbc*L of some marine green macroalgae (Hanyuda *et al.*, 2000) may negatively affect the universality of *rbc*L as a barcode marker since the ability to amplify and sequence large fragments with a single bidirectional read will be hampered. Given the fact that green algae are particularly prone to acquiring intron sequences (Bhattacharya *et al.*, 1996; Haugen *et al.*, 2005), and that the extent to which introns are present among their *rbc*L genes is unknown, the wider applicability of this marker as a barcode needs to be evaluated empirically.

The plastid elongation factor *tufA* has similarly been used for phylogenetic analyses and to discriminate among green algal species and is known to lack introns for the variety of green taxa tested to date (Fama *et al.*, 2002; O' Kelly *et al.*, 2004; De Clerck *et al.*, 2008; Verbruggen *et al.*, 2009; Zuccarello *et al.*, 2009; Händeler *et al.*, 2010).

The universal plastid amplicon has been proposed as an alternative DNA barcoding marker for photosynthetic eukaryotes (Sherwood & Presting, 2007; Sherwood *et al.*, 2008), but has had mixed results in different taxonomic groups. For some red algal lineages it has worked as well as COI-5P to distinguish species (Sherwood *et al.*, 2008; Clarkston & Saunders, 2010) while in other groups it was not sufficiently variable to identify species (Hamsher *et al.*, *in press*; Kucera & Saunders, unpublished data). The major advantage with the UPA is its universality. A single primer pair can reliably recover sequences from a broad taxonomic range including green, red and brown marine macroalgae, diatoms, and even cyanobacteria (Sherwood & Presting, 2007).

The D2/D3 region of the LSU is known to be variable at the species level in some lineages and has been used to identify and screen for new species in taxonomic groups including animals, plants and algae (e.g., Saunders & Lehmkuhl, 2005; Subbotin *et al.*, 2005; Hajieghrari *et al.*, 2007; Heraty *et al.*, 2007). Primers modified from those developed for the red algae (Harper & Saunders, 2001) are anchored in conserved regions and are expected to be broadly universal

among green algae having amplified target sequence from a broad range of taxa in our laboratory.

As well as serving as the DNA barcode marker for fungi, the internal transcribed spacer region of the ribosomal cistron has been used extensively for investigations of phylogeny, molecular ecology and evolution in a broad array of taxa including marine green macroalgae (e.g., Bakker *et al.*, 1995; Hayden *et al.*, 2003; Hayden & Waaland, 2004). White *et al.*,'s (1990) now classic paper has guided primer design for amplification of the ITS from many lineages including the green algae (e.g., Bakker *et al.*, 1992). Despite the fact that there are potential problems (such as intraindividual heterogeneity or difficulties with multiple sequence alignments due to high levels of variation) with the ITS (Alvarez & Wendel, 2003), it can also provide insightful results at the species level (Feliner & Rossello, 2007).

The aim of this study was to conduct a preliminary test of the universality and genetic variability of molecular markers for DNA barcoding of marine green macroalgae. We focused our testing on three plastid markers (divided into four putative barcodes) and two nuclear markers. From the plastid: 559 nucleotides of the 5' end of the *rbc*L (*rbc*L-5P); approximately 735 nucleotides of the 3' end of the *rbc*L (*rbc*L-3P); 807 nucleotides of the *tufA*; and the universal plastid amplicon (UPA) (Presting, 2006; Sherwood & Presting, 2007). From the nucleus: the D2/D3 variable domains of the large ribosomal subunit (LSU-D2/D3); and the internal transcribed spacer (ITS) region of the ribosomal cistron.

MATERIALS AND METHODS

Specimen collection and DNA extraction

For the initial screening of potential barcode markers, 99 samples were chosen as a test group (referred to herein as "the test sample set") (Table 1). Samples were identified using taxonomic keys (Sears, 2002; Gabrielson *et al.*, 2006) except for several *Ulva* specimens, which were given provisional names until after sequencing (see below). The test sample set was chosen to include a broad taxonomic range within the marine green macroalgae (universality), but with an emphasis on the genus *Ulva* in order to evaluate intra- versus interspecific variation for this genus (discriminatory power). In the genus *Ulva*, samples were explicitly chosen to include a wide geographic distribution and variety of morphologies for recognized species. Following collection, each sample was pressed on herbarium paper as a voucher, while a subsample was dried in silica gel for DNA extraction. The silica-dried portion was ground in liquid nitrogen using a mortar and pestle. Extraction of total genomic DNA was carried out using the protocol from Saunders (1993) with modifications (Saunders, 2008).

Primer design and selection

Because DNA barcode markers should be short enough to be sequenced bidirectionally with a single primer pair, we divided the *rbc*L into 5' (*rbc*L-5P) and 3' (*rbc*L-3P) regions and treated them as separate markers. We made an alignment of available green algal sequences from Genbank and designed the

Table 1. Collection information and accession numbers for samples used in the test sample set

			Genbank Accession numbers[c]					
Species	*Voucher # BOLD ID*[a]	*Collection Site*[b]	rbc*L-5P*	rbc*L-3P*	tuf*A*	*UPA*	*LSU*	*ITS*
Bryopsidophyceae, Bryopsidales, Bryopsidaceae								
Bryopsis corticulans Setchell	GWS009004 ULVA571-09	Scotts Bay, Bamfield, BC	HQ603381	HQ603483	HQ610243	HQ603325	U/C	NA
	GWS009075 ULVA574-09	Wizard I., Bamfield, BC	HQ603380	HQ603482	HQ610242	HQ603324	U/C	U/C
Bryopsidophyceae, Bryopsidales, Codiaceae								
Codium fragile (Suringar) Hariot	GWS002780 ULVA478-09	Dixon, I., Bamfield, BC	HQ603382	U/C	HQ610245	HQ603326	U/C	NA
Codium fragile sp. 1[d]	GWS003527 ULVA488-09	Peggys Cove, NS	HQ603383	U/C	HQ610247	HQ603327	U/C	NA
Codium setchellii Gardner	GWS002933 ULVA482-09	Bradys Beach, Bamfield, BC	NA	NA	HQ610246	NA	NA	NA
Codium sp. 1GWS	GWS008955 ULVA570-09	Seapool Rock, Bamfield, BC	NA	NA	HQ610248	NA	NA	U/C
Trebouxiophyceae, Prasiolales, Derbesiaceae								
Derbesia marina (Lyngbye) Solier	GWS008836 ULVA567-09	Beaver Harbour, NB	U/C	HQ603485	HQ610249	HQ603328	NA	U/C
Trebouxiophyceae, Prasiolales, Prasiolaceae								
Prasiola meridionalis Setchell et Gardner[e]	GWS002865 ULVA481-09	Wizard I., Bamfield, BC	HQ603386	HQ603501	HQ610266	HQ603330	U/C	NA

[a] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[b] Abbreviations for Provinces/States as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; QC, Quebec, Canada.
[c] Abbreviations for sample outcomes: NA, no amplification; U/C, unreadable/contaminated; IN, presence of an intron prevented acquisition of sequence for this region.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.
[e] Based on all of our genetic results these two *Prasiola* sp. are conspecific.

Table 1. Collection information and accession numbers for samples used in the test sample set *(cont'd)*

			Genbank Accession numbers[c]					
Species	*Voucher # BOLD ID*[a]	*Collection Site*[b]	rbc*L-5P*	rbc*L-3P*	tuf*A*	*UPA*	*LSU*	*ITS*
Prasiola stipitata Suhr ex Jessen[e]	GWS003898 ULVA502-09	Lepreau, NB	HQ603385	HQ603500	HQ610265	U/C	U/C	NA
Ulvophyceae, Cladophorales, Cladophoraceae								
Chaetomorpha brachygona Harvey	GWS007053 ULVA550-09	St. Paul, Bonne Bay, NL	NA	U/C	NA	U/C	NA	U/C
	GWS008847 ULVA569-09	Beaver Harbour, NB	NA	U/C	NA	U/C	U/C	NA
Chaetomorpha melagonium (F. Weber et D. Mohr) Kützing	GWS003521 ULVA487-09	Cape St. Marys, NS	NA	NA	NA	NA	NA	NA
	GWS005272 ULVA514-09	Churchill Northern Studies Centre, MB	NA	NA	NA	NA	NA	U/C
	GWS006949 ULVA546-09	Peggys Cove, NS	NA	NA	NA	NA	NA	NA
Chaetomorpha piquotiana Montagne ex. Kützing	GWS003617 ULVA490-09	Cape Neddick, ME	NA	NA	NA	NA	NA	U/C
Chaetomorpha sp.	GWS006227 ULVA528-09	Kouchibouguac National Park, NB	NA	NA	NA	NA	NA	NA
Cladophora albida (Nees) Kützing	GWS006228 ULVA529-09	Kouchibouguac National Park, NB	U/C	U/C	NA	U/C	NA	BOLD

[a] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[b] Abbreviations for Provinces/States as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; QC, Quebec, Canada.
[c] Abbreviations for sample outcomes: NA, no amplification; U/C, unreadable/contaminated; IN, presence of an intron prevented acquisition of sequence for this region.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.
[e] Based on all of our genetic results these two *Prasiola* sp. are conspecific.

Table 1. Collection information and accession numbers for samples used in the test sample set *(cont'd)*

			Genbank Accession numbers[c]					
Species	*Voucher # BOLD ID*[a]	*Collection Site*[b]	rbc*L-5P*	rbc*L-3P*	tuf*A*	*UPA*	*LSU*	*ITS*
Cladophora columbiana Collins in Setchell et Gardner	GWS002814 ULVA479-09	Dixon, I., Bamfield, BC	NA	NA	NA	NA	NA	NA
Cladophora ruchingeri (C. Agardh) Kützing	GWS005860 ULVA525-09	Lepreau, NB	NA	U/C	NA	U/C	U/C	BOLD
Cladophora rupestris (Linnaeus) Kützing	GWS003791 ULVA500-09	Meadow Cove, NB	NA	NA	NA	NA	NA	BOLD
Cladophora sericea (Hudson) Kützing	GWS003029 ULVA484-09	Ross Islets, Bamfield, BC	NA	NA	NA	U/C	NA	U/C
Cladophora sp.	GWS004311 ULVA505-09	Pachena Beach, Bamfield, BC	NA	NA	NA	U/C	NA	NA
	GWS006670 ULVA544-09	Island #40, Tahsis, BC	NA	NA	NA	NA	NA	NA
	GWS006968 ULVA548-09	Whycocomagh, Bras d'Or Lakes, NS	NA	NA	NA	NA	NA	NA
	GWS007351 ULVA555-09	St. Davids, NL	NA	NA	NA	NA	NA	U/C
Cladophora vagabunda (Linnaeus) C. Hoek	GWS007136 ULVA554-09	Deer Arm, Bonne Bay, NL	U/C	U/C	NA	U/C	U/C	NA

[a] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[b] Abbreviations for Provinces/States as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; QC, Quebec, Canada.
[c] Abbreviations for sample outcomes: NA, no amplification; U/C, unreadable/contaminated; IN, presence of an intron prevented acquisition of sequence for this region.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.
[e] Based on all of our genetic results these two *Prasiola* sp. are conspecific.

Table 1. Collection information and accession numbers for samples used in the test sample set *(cont'd)*

			Genbank Accession numbers[c]					
Species	*Voucher # BOLD ID*[a]	*Collection Site*[b]	rbc*L-5P*	rbc*L-3P*	tuf*A*	*UPA*	*LSU*	*ITS*
Lola lubrica (Setchell et Gardner in Gardner) Hamel et Hamel	GWS002984 ULVA483-09	Cape Beale, Bamfield, BC	NA	NA	NA	NA	NA	U/C
Rhizoclonium riparium (Roth) Harvey	GWS003717 ULVA494-09	Letete, NB	NA	NA	NA	NA	NA	NA
Ulvophyceae, Ulotrichales, Gomontiaceae								
Protomonostroma undulatum (Wittrock) K.L. Vinogradova	GWS003734 ULVA495-09	Pettes Cove, Grand Manan I., NB	HQ603389	HQ603506	HQ610271	HQ603333	HQ603270	U/C
	GWS003753 ULVA497-09	Harrington Cove, Grand Manan, NB	HQ603388	HQ603505	HQ610270	HQ603332	HQ603269	NA
	GWS005967 ULVA526-09	Long Eddy Point, Grand Manan, NB	HQ603387	HQ603504	HQ610269	HQ603331	HQ603268	BOLD
Ulvophyceae, Ulotrichales, Ulotrichales								
Acrosiphonia arcta (Dillwyn) J. Agardh	GWS003577 ULVA489-09	Lepreau, NB	HQ603373	HQ603446	HQ610204	U/C	HQ603261	BOLD
Acrosiphonia coalita (Ruprecht) Scagel et al.	GWS002760 ULVA476-09	Dixon I., Bamfield, BC	HQ603375	HQ603458	HQ610216	U/C	HQ603263	BOLD
	GWS004310 ULVA504-09	Pachena Beach, Bamfield, BC	HQ603374	HQ603457	HQ610215	HQ603318	HQ603262	BOLD

[a] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[b] Abbreviations for Provinces/States as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; QC, Quebec, Canada.
[c] Abbreviations for sample outcomes: NA, no amplification; U/C, unreadable/contaminated; IN, presence of an intron prevented acquisition of sequence for this region.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.
[e] Based on all of our genetic results these two *Prasiola* sp. are conspecific.

Table 1. Collection information and accession numbers for samples used in the test sample set *(cont'd)*

			Genbank Accession numbers[c]					
Species	*Voucher # BOLD ID*[a]	*Collection Site*[b]	rbc*L-5P*	rbc*L-3P*	tuf*A*	*UPA*	*LSU*	*ITS*
Acrosiphonia sonderi (Kützing) Kornmann	GWS003750 ULVA496-09	Harrington Cove, Grand Manan I., NB	HQ603376	HQ603461	HQ610219	HQ603319	HQ603264	BOLD
Acrosiphonia sp. 1GWS[d]	GWS006665 ULVA542-09	Island #40, Tahsis, BC	HQ603377	HQ603468	HQ610226	HQ603320	HQ603265	BOLD
	GWS008617 ULVA566-09	Comox Marina Breakwater, BC	U/C	HQ603467	HQ610225	U/C	U/C	U/C
Acrosiphonia sp. 6GWS	GWS005383 ULVA518-09	NW of Eskimo I., Churchill, MB	HQ603378	U/C	NA	HQ603321	NA	BOLD
Spongomorpha aeruginosa (Linnaeus) C. Hoek	GWS003768 ULVA498-09	Harrington Cove, Grand Manan, NB	U/C	HQ603511	HQ610276	NA	HQ603271	BOLD
Urospora penicilliformis (Roth) Areschoug sp. 1[d]	GWS008049 ULVA559-09	Riviere du Loup, QC	HQ603442	U/C	HQ610439	U/C	U/C	U/C
Urospora penicilliformis (Roth) Areschoug sp. 2[d]	GWS008147 ULVA561-09	Pachena Bay, Bamfield, BC	HQ603443	HQ603675	NA	HQ603371	HQ603316	NA
Urospora wormskioldii (Mertens ex Hornemann) Rosenvinge	GWS006374 ULVA535-09	Otter Point, Vancouver I., BC	HQ603444	HQ603676	HQ610441	HQ603372	HQ603317	U/C
Ulvophyceae, Ulvales, Kornmanniaceae								
Blidingia marginata (J. Agardh) P.J.L. Dang	GWS003689 ULVA492-09	Letete, NB	HQ603379	HQ603479	HQ610237	HQ603322	HQ603266	NA

[a] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[b] Abbreviations for Provinces/States as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; QC, Quebec, Canada.
[c] Abbreviations for sample outcomes: NA, no amplification; U/C, unreadable/contaminated; IN, presence of an intron prevented acquisition of sequence for this region.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.
[e] Based on all of our genetic results these two *Prasiola* sp. are conspecific.

Table 1. Collection information and accession numbers for samples used in the test sample set *(cont'd)*

			Genbank Accession numbers[c]					
Species	*Voucher # BOLD ID*[a]	*Collection Site*[b]	rbc*L-5P*	rbc*L-3P*	tuf*A*	*UPA*	*LSU*	*ITS*
Blidingia minima (Nägeli ex Kützing) Kylin	GWS007133 ULVA553-09	Deer Arm, Bonne Bay, NL	NA	U/C	HQ610239	HQ603323	HQ603267	NA
Blidingia sp. 5GWS	GWS005273 ULVA515-09	Churchill Northern Studies Centre, MB	NA	NA	HQ610241	NA	NA	U/C
Kornmannia leptoderma (Kjellman) Bliding	GWS004830 ULVA509-09	Ridley Island, Prince Rupert, BC	HQ603384	U/C	HQ610252	HQ603329	NA	BOLD
Ulvophyceae, Ulvales, Ulvaceae								
Ul*va californica* Wille	GWS004842 ULVA510-09	Butze Rapids, Prince Rupert, BC	HQ603391	HQ603515	HQ610280	HQ603334	HQ603273	BOLD
	GWS005072 ULVA511-09	Ridley Island, Prince Rupert, BC	HQ603390	HQ603514	HQ610279	U/C	HQ603272	U/C
Ulva compressa Linnaeus	GWS004087 ULVA503-09	Wizard I., Bamfield, BC	HQ603394	HQ603521	HQ610285	U/C	HQ603276	NA
	GWS005694 ULVA521-09	Lepreau, NB	HQ603393	HQ603520	HQ610284	U/C	HQ603275	NA
	GWS008267 ULVA562-09	Bradys Beach, Bamfield, BC	HQ603392	HQ603519	HQ610429	IN	HQ603274	NA
Ulva flexuosa Wulfen	GWS008545 ULVA564-09	Backeddy Resort, BC	HQ603395	HQ603532	HQ610296	HQ603335	HQ603277	U/C

[a] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[b] Abbreviations for Provinces/States as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; QC, Quebec, Canada.
[c] Abbreviations for sample outcomes: NA, no amplification; U/C, unreadable/contaminated; IN, presence of an intron prevented acquisition of sequence for this region.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.
[e] Based on all of our genetic results these two *Prasiola* sp. are conspecific.

Table 1. Collection information and accession numbers for samples used in the test sample set *(cont'd)*

			Genbank Accession numbers[c]					
Species	*Voucher # BOLD ID*[a]	*Collection Site*[b]	rbc*L-5P*	rbc*L-3P*	tuf*A*	*UPA*	*LSU*	*ITS*
Ulva intestinalis Linnaeus	GWS004618 ULVA508-09	Kye Bay, Vancouver I., BC	HQ603401	HQ603542	HQ610306	U/C	HQ603283	U/C
	GWS006458 ULVA538-09	Stephenson Pt., Nanaimo, BC	HQ603400	HQ603541	HQ610305	U/C	HQ603282	NA
	GWS006939 ULVA545-09	Beach Meadows, NS	HQ603399	HQ603540	HQ610304	IN	HQ603281	NA
	GWS006962 ULVA547-09	Whycocomagh, Bras d'Or, NS	HQ603398	HQ603539	HQ610303	U/C	HQ603280	NA
	GWS007959 ULVA557-09	L'Anse Bleue, Northumberland Strait, NB	HQ603397	HQ603538	HQ610302	IN	HQ603279	BOLD
	GWS008549 ULVA565-09	Courtenay Estuary, BC	HQ603396	HQ603537	HQ610301	U/C	HQ603278	NA
Ulva lactuca Linnaeus	GWS003686 ULVA491-09	Letete, NB	HQ603411	HQ603584	HQ610350	HQ603346	HQ603292	NA
	GWS003817 ULVA501-09	Beaver Harbour, NB	HQ603410	HQ603583	HQ610349	HQ603345	HQ603291	NA
	GWS005100 ULVA513-09	Ridley Island Prince Rupert, BC	HQ603409	HQ603582	HQ610348	HQ603344	HQ603290	U/C
	GWS005338 ULVA517-09	Button Bay, Churchill, MB	HQ603408	HQ603581	HQ610347	HQ603343	HQ603289	NA

[a] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[b] Abbreviations for Provinces/States as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; QC, Quebec, Canada.
[c] Abbreviations for sample outcomes: NA, no amplification; U/C, unreadable/contaminated; IN, presence of an intron prevented acquisition of sequence for this region.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.
[e] Based on all of our genetic results these two *Prasiola* sp. are conspecific.

Table 1. Collection information and accession numbers for samples used in the test sample set *(cont'd)*

			Genbank Accession numbers[c]					
Species	*Voucher # BOLD ID*[a]	*Collection Site*[b]	rbc*L-5P*	rbc*L-3P*	tuf*A*	*UPA*	*LSU*	*ITS*
	GWS005837 ULVA523-09	Scotts Bay, Bamfield, BC	HQ603407	HQ603580	HQ610346	HQ603342	HQ603288	NA
	GWS005859 ULVA524-09	Lepreau, NB	HQ603406	HQ603579	HQ610345	HQ603341	HQ603287	U/C
	GWS006258 ULVA531-09	Richebucto Cape Breakwater, NB	HQ603405	HQ603578	HQ610344	HQ603340	HQ603286	NA
	GWS006574 ULVA541-09	Nuchatliz Island, Tahsis, BC	NA	U/C	HQ610343	HQ603339	U/C	NA
	GWS006666 ULVA543-09	Island #40, Tahsis, BC	HQ603404	U/C	HQ610342	HQ603338	NA	U/C
	GWS007962 ULVA558-09	Cap des Caissie, Shediac, NB	HQ603403	HQ603577	HQ610341	HQ603337	HQ603285	NA
	GWS008295 ULVA563-09	Pachena Beach, Bamfield, BC	HQ603402	HQ603576	HQ610340	HQ603336	HQ603284	NA
Ulva linza Linnaeus	GWS006377 ULVA536-09	Otter Point, Vancouver I., BC	HQ603413	HQ603603	HQ610368	HQ603348	U/C	NA
	GWS008140 ULVA560-09	Pachena Bay, Bamfield, BC	HQ603412	HQ603602	NA	HQ603347	HQ603293	U/C
Ulva lobata (Kützing) Harvey	GWS002779 ULVA477-09	Dixon, I., Bamfield, BC	HQ603417	HQ603612	NA	HQ603352	HQ603296	U/C

[a] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[b] Abbreviations for Provinces/States as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; QC, Quebec, Canada.
[c] Abbreviations for sample outcomes: NA, no amplification; U/C, unreadable/contaminated; IN, presence of an intron prevented acquisition of sequence for this region.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.
[e] Based on all of our genetic results these two *Prasiola* sp. are conspecific.

Table 1. Collection information and accession numbers for samples used in the test sample set *(cont'd)*

			Genbank Accession numbers[c]					
Species	*Voucher # BOLD ID*[a]	*Collection Site*[b]	rbc*L-5P*	rbc*L-3P*	tuf*A*	*UPA*	*LSU*	*ITS*
	GWS002820 ULVA480-09	Dixon, I., Bamfield, BC	HQ603416	HQ603611	HQ610376	HQ603351	NA	U/C
	GWS009007 ULVA572-09	Reef N. of Isl. 76, Broken Group, BC	HQ603415	HQ603610	HQ610375	HQ603350	HQ603295	NA
	GWS009013 ULVA573-09	Gilbert Island, Broken Group, BC	HQ603414	HQ603609	HQ610374	HQ603349	HQ603294	NA
Ulva pertusa Kjellman	GWS005803 ULVA522-09	Aguilar Point, Bamfield, BC	HQ603419	HQ603616	HQ610380	HQ603354	HQ603298	NA
	GWS006489 ULVA539-09	Stephenson Pt., Nanaimo, BC	HQ603418	HQ603615	HQ610379	HQ603353	HQ603297	U/C
Ulva procera (K. Ahlner) Hayden et al.	GWS006271 ULVA532-09	Richebucto Cape Breakwater, NB	HQ603420	HQ603628	HQ610392	HQ603355	HQ603299	NA
Ulva prolifera O.F. Müller	GWS003715 ULVA493-09	Letete, NB	HQ603425	HQ603634	HQ610398	IN	HQ603304	NA
	GWS005321 ULVA516-09	Gordon Point, Churchill, MB	HQ603424	HQ603633	HQ610397	U/C	HQ603303	NA
	GWS005446 ULVA519-09	Churchill River, Churchill, MB	HQ603423	HQ603632	HQ610396	HQ603356	HQ603302	NA
	GWS005572 ULVA520-09	Gravel Pit E of Churchill, MB	HQ603422	HQ603631	HQ610395	U/C	HQ603301	U/C

[a] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[b] Abbreviations for Provinces/States as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; QC, Quebec, Canada.
[c] Abbreviations for sample outcomes: NA, no amplification; U/C, unreadable/contaminated; IN, presence of an intron prevented acquisition of sequence for this region.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.
[e] Based on all of our genetic results these two *Prasiola* sp. are conspecific.

Table 1. Collection information and accession numbers for samples used in the test sample set *(cont'd)*

			Genbank Accession numbers[c]					
Species	*Voucher # BOLD ID*[a]	*Collection Site*[b]	rbc*L-5P*	rbc*L-3P*	tuf*A*	*UPA*	*LSU*	*ITS*
	GWS007057 ULVA551-09	St. Paul, Bonne Bay, NL	HQ603421	HQ603630	HQ610394	U/C	HQ603300	NA
Ulva rigida C. Agardh	GWS006232 ULVA530-09	Kouchibouguac National Park, NB	HQ603437	HQ603664	HQ610428	U/C	NA	NA
Ulva sp. 1GWS	GWS004319 ULVA506-09	Bamfield Inlet, Bamfield, BC	HQ603438	HQ603665	U/C	HQ603367	HQ603313	NA
Ulva stenophylla Setchell et Gardner	GWS003290 ULVA485-09	Seppings I., Bamfield, BC	HQ603439	HQ603669	HQ610433	HQ603368	HQ603314	NA
Ulva torta (Mertens) Trevisan	GWS004617 ULVA507-09	Kye Bay, Vancouver I., BC	HQ603441	HQ603673	HQ610438	HQ603370	U/C	NA
	GWS006531 ULVA540-09	Tahsis Narrows, BC	HQ603440	NA	HQ610437	HQ603369	HQ603315	U/C
Ulvaria obscura (Kützing) Gayral	GWS003507 ULVA486-09	Lepreau, NB	HQ603436	HQ603651	HQ610415	HQ603366	HQ603312	NA
	GWS003786 ULVA499-09	Meadow Cove, NB	HQ603435	HQ603650	HQ610414	HQ603365	HQ603311	BOLD
	GWS005073 ULVA512-09	Ridley Island, Prince Rupert, BC	HQ603434	HQ603649	HQ610413	NA	NA	BOLD
	GWS006119 ULVA527-09	Escoumins, QC	HQ603433	HQ603648	HQ610412	HQ603364	HQ603310	NA

[a] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[b] Abbreviations for Provinces/States as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; QC, Quebec, Canada.
[c] Abbreviations for sample outcomes: NA, no amplification; U/C, unreadable/contaminated; IN, presence of an intron prevented acquisition of sequence for this region.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.
[e] Based on all of our genetic results these two *Prasiola* sp. are conspecific.

Table 1. Collection information and accession numbers for samples used in the test sample set *(cont'd)*

			Genbank Accession numbers[c]					
Species	*Voucher # BOLD ID*[a]	*Collection Site*[b]	rbc*L-5P*	rbc*L-3P*	tuf*A*	*UPA*	*LSU*	*ITS*
	GWS006315 ULVA533-09	Sidney, BC	HQ603432	HQ603647	HQ610411	HQ603363	HQ603309	NA
	GWS006316 ULVA534-09	Sidney, BC	HQ603431	HQ603646	HQ610410	HQ603362	HQ603308	NA
	GWS006404 ULVA537-09	Whiffen Spit, Vancouver I., BC	HQ603430	HQ603645	HQ610409	HQ603361	HQ603307	U/C
	GWS006999 ULVA549-09	Cape Ray, NL	HQ603429	HQ603644	HQ610408	HQ603360	HQ603306	NA
	GWS007079 ULVA552-09	Bonne Bay Station, NL	HQ603428	HQ603643	HQ610407	HQ603359	HQ603305	NA
	GWS007568 ULVA556-09	English Harbour, NL	HQ603427	HQ603642	HQ610406	HQ603358	NA	NA
	GWS008841 ULVA568-09	Beaver Harbour, NB	HQ603426	HQ603641	HQ610405	HQ603357	U/C	U/C

[a] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[b] Abbreviations for Provinces/States as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; QC, Quebec, Canada.
[c] Abbreviations for sample outcomes: NA, no amplification; U/C, unreadable/contaminated; IN, presence of an intron prevented acquisition of sequence for this region.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.
[e] Based on all of our genetic results these two *Prasiola* sp. are conspecific.

following primers: GrbLF (5' GTTAAAGATTAYCGWYTAAC 3'), G*rbcL*nF (5' GCTGGWGTAAAAGATTAYCG 3') and G*rbcL*R (5' TCACGCCAACG-CATRAASGG 3'). G*rbcL*F and G*rbcL*nF are partially overlapping forward primers at the 5' end of the *rbc*L, with G*rbcL*nF being the primer that produced better quality data. G*rbcL*R is a reverse primer anchored 560bp downstream of G*rbcL*F. Combinations of these three primers served to amplify and sequence the *rbc*L-5P. For the *rbc*L-3P, the forward primer G*rbcL*Fi (5' TCTCARCCWT-TYATGCGTTGG 3') was designed (and is a partial reverse complement of G*rbcL*R). The reverse primer used for the *rbc*L-3P was 1385R as published by Manhart (1994).

To amplify *tufA* primers tufGF4 (5' GGNGCNGCNCAAATGGAYGG 3') and tufAR (5' CCTTCNCGAATMGCRAAWCGC 3'; from Fama *et al.*, (2002)) were developed. The UPA primers were p23SrV_f1, p23SrV_r1 (Sherwood & Presting, 2007) and p23SnewR (modified from p32SrV_r1 (Clarkston & Saunders, 2010)). The primers used for the LSU-D2/D3 region were: T16N (forward; 5' AMAAGTACCRYGAGGGAAAG 3') modified from Harper and Saunders (2001) and T24 (reverse) (Harper & Saunders, 2001). For the ITS, the general primers P1 (forward) and G4 (reverse) (Harper & Saunders, 2001), modified from White *et al.*, (1990), were used. Due to low success with P1 and G4 (see results), we also tested published primers ITS1 (forward; Bakker *et al.*, (1995) modified from White (1990)), 18S1763 (forward; Hayden *et al.,* (2003) modified from Blomster *et al.*, (1998)) and JO6 (reverse; Lindstrom & Hanic (2005)) in four combinations: ITS1/JO6, ITS1/G4, 18S1763/JO6 and 18S1763/G4 on 16 samples including representatives of the genera *Blidingia*, *Bryopsis*, *Chaetomorpha*, *Cladophora*, *Codium*, *Derbesia*, *Prasiola*, *Protomonostroma*, *Rhizoclonium*, *Ulva*, *Ulvaria*, and *Urospora*.

PCR protocols and sequence acquisition

PCR amplification was carried out using the Ex Taq^{TM} DNA Polymerase (Takara, Shiga, Japan) according to the manufacturer's recommendations with a final volume of 20 µL per reaction. PCR profiles were as follows: *rbc*L-5P – an initial 2 min denaturation at 95 °C, 35 cycles of 93 °C for 1 min, 54°C annealing for 45 s, 72 °C extension for 2 min, followed by 72 °C final extension for 7 min; *rbc*L-3P – profile as for *rbc*L -5P except with an annealing temperature of 50°C; *tufA* – an initial 4 min denaturation at 94 °C, 38 cycles of 94°C for 1 min, 45°C annealing for 30 s, 72°C extension for 1 min, followed by 72°C final extension for 7 min; UPA – as published in Sherwood & Presting (2007); LSU-D2/D3 – an initial 5 min denaturation at 94°C, 38 cycles of 94°C for 30 s, 50°C annealing for 30 s, 72°C extension for 1 min, followed by 72°C final extension for 7 min; ITS – an initial 3 min denaturation at 94°C, 38 cycles of 94°C for 30 s, 54°C annealing for 40 s, 72°C extension for 1.5 min, followed by 72°C final extension for 7 min. All PCR products were held at 4°C following amplification. Amplification success was evaluated using gel electrophoresis in a 0.8% agarose gel.

PCR products were cleaned using the Exo-Sap-IT kit (USB, Cleveland OH, USA) or via the gel electrophoresis technique described by Saunders (1993). Sequencing was done using the BigDye 3.0 kit (PE Applied Biosystems; Foster City, CA, USA), employing the same primers as for PCR. Sequence trace files were generated on an Applied Biosystems 3130 XL automated sequencer (PE Applied Biosystems; Foster City, CA, USA). Sequences were edited using Sequencher version 4.2 (Gene Codes Corporation, Ann Arbor, MI, USA) and

multiple sequence alignments for the *rbc*L-5P, *rbc*L-3P, *tufA*, UPA, and LSU-D2/D3 were constructed in MacClade version 4.08 (Maddison & Maddison, 2005). The multiple sequence alignment for ITS data was assembled using the Clustalw2 (Larkin *et al*., 2007) plugin for Geneious v4.8.3 (Drummond *et al*., 2009) with a gap opening cost of 13.4 and a gap extension cost of 7.26, with free end gaps option selected.

For each marker and each sample, up to two PCR trials were attempted for each primer combination in an effort to achieve successful amplification. Successfully amplified samples proceeded to sequencing, however, if no band was detected, or if a sequence was found to be contaminated (messy or belonging to a non-green algal organism), then a second PCR attempt was made. If multiple bands were detected, and there was a strong band of the size expected for the marker of interest, the band was isolated using gel electrophoresis purification (Saunders, 1993). In cases where there were many bands and/or a weak or non-existent band of the expected size, the PCR product was discarded. In the UPA, some cases of two bands were detected — one of the expected size and one larger band. In these cases both bands were isolated and sequenced.

Comparison of markers

To evaluate universality, PCR and sequencing results were pooled into one of three descriptors of overall success as follows: i) successful sequence – amplification and sequencing of the target organism successful; ii) unreadable/contaminant sequence (U/C; Table 1) — amplification produced a single band, or a band that could be isolated, however, the resulting sequence was either messy due to contamination or read slippage, or belonged to a non-target organism; iii) sequencing not attempted (NA; Table 1) — PCR failed, or produced multiple bands of more or less equal strength or streaks. To evaluate the discriminatory power of each marker, uncorrected-p distances were calculated for each marker using PAUP* version 4.0b10 (Swofford, 2002). Distances were exported to Microsoft Excel for Mac 2004 version 11.5.5 (Microsoft Corporation, Mississauga, Ontario, Canada), transformed to percentages and the maximum and minimum values were found using the MAX and MIN formula functions. Neighbor-joining (NJ) analyses were also preformed on uncorrected distances using PAUP* for each marker to test for clustering of isolates into genetic species groups.

Species identification within the test sample set subsequent to DNA sequencing

For specimens of *Ulva* that were given a provisional name, available Genbank sequences were used to provide species names as follows. For every *Ulva* species listed on Genbank as of July 14, 2009, up to five representative sequences were downloaded. The sequences were aligned with our *rbc*L data, which were included either by concatenating the 5' and 3' sequences for each sample, or including singly whichever fragment was successfully sequenced. This dataset was subjected to neighbor-joining analysis using PAUP* and names were assigned to our specimens based on clustering with named specimens from Genbank. In cases where Genbank sequences of more than one name were found in the same group, the publications from which the sequences originated provided information as to which name is currently accepted for that group. In cases where our specimens did not cluster with any Genbank data, the specimens were given interim names (e.g., *Ulva* sp. 1GWS, etc.).

Further evaluation of the *rbc*L-3P & tufA

Following initial evaluation of all markers using the test sample set, the *rbc*L-3P and *tufA* were tested on an additional 164 samples (Table 2). The datasets (total of 263 samples; "the extended sample set") were used for a wider assessment of species resolving power for these two markers.

RESULTS AND DISCUSSION

Universality of the markers

The BOLD and Genbank accessions for each marker successfully sequenced from each sample in the test sample set are listed in Table 1. The percent of samples sequenced for each marker is shown in Figure 1 with results for the total test sample set (n=99) presented as the first bar in each case (T). For the Cladophoraceae (n=19), the only data obtained were three ITS sequences, which prompted us to present the test sample set results also as two subsets: one with the Cladophoraceae excluded (n=80; E, Fig. 1); and another for the Cladophoraceae only (n=19; C, Fig. 1).

Of the markers tested (excluding the Cladophoraceae), *tufA* had the highest universality (94%) and lowest levels of contamination (1%) and highest sequence (trace file) quality, followed by the *rbc*L markers (see Table 3 for a marker summary). The *rbc*L-5P had higher universality (90%) and lower contamination (4%) than the *rbc*L-3P (81 and 14%, respectively; Fig. 1), but it also had the lowest sequence (trace file) quality (Table 3). For samples of the genus *Codium*, *rbc*L-3P sequence could not be obtained, probably due to the presence of introns within this region of the *rbc*L (Hanyuda *et al.*, 2000). We were, however, able to sequence the *rbc*L-5P for the *Codium* samples – the introns in this species are likely confined to the 3' end of the *rbc*L. The presence of introns within other genera, such as *Bryopsis, Caulerpa* and members of the Chlorophyceae (Hanyuda *et al.*, 2000), potentially poses a significant shortfall for the universality of the *rbc*L as a DNA barcoding marker in green algae (Table 3). The *tufA* marker, on the other hand, showed no indications of containing introns for the studied samples (Table 3).

Although the UPA had relatively high PCR success (dark plus light grey bars, Fig. 1) (Table 3), 20 samples turned out to be contaminated or messy on sequencing (light grey bars, Fig. 1). It is possible that epiphytic or endophytic contaminants present in the samples were amplified preferentially or concurrently to the target organism for one of the following reasons: a) the UPA primers were a better match for the contaminants than the target organism; b) DNA extraction was more effective for the contaminants than for the target organism; c) stochastic PCR effects resulted in preferential amplification of the contaminant; or d) some combination of the previous. Markers with highly universal primers, such as the UPA, are particularly prone to these problems and effective isolation of target DNA becomes critically important. There are, however, other explanations.

Previous to this study, the only marine green macroalgal genera for which the UPA had been tested were *Codium*, *Prasiola*, *Rosenvingiella*, and *Ulva* (Sherwood & Presting, 2007). It is, therefore, possible that the primers are not a good match for other genera, though this is unlikely given the universality of this

Table 2. Collection information and accession numbers for rbcL-3P data generated for the extended sample set of specimens

				Genbank accession	
Species	*Voucher #*	*Collection Site*[a]	*BOLD ID*[b]	rbc*L-3P*	tuf*A*
Bryopsidophyceae, Bryopsidales, Bryopsidaceae					
Bryopsis sp. 1GWS	GWS005754	Fort Wetherill, RI	ULVA256-09	HQ603484	HQ610244
Bryopsidophyceae, Bryopsidales, Derbesiaceae					
Derbesia marina (Lyngbye) Solier	GWS005999	Pier #5, Narragansett, RI	ULVA229-09	HQ603486	HQ610250
Derbesia sp. 1GWS	GWS004888	Tree Knob Islands, Prince Rupert, BC	ULVA190-09	HQ603487	HQ610251
Trebouxiophyceae, Prasiolales, Prasiolaceae					
Prasiola borealis M. Reed[c]	GWS005101	Ridley Island, Prince Rupert, BC	ULVA122-09	HQ603499	HQ610264
Prasiola delicata Setchell et N.L. Gardner[c]	GWS005076	Ridley Island, Prince Rupert, BC	ULVA169-09	HQ603498	HQ610263
Prasiola meridionalis Setchell et Gardner	GWS004462	Port Renfrew, Vancouver I., BC	ULVA177-09	HQ603502	HQ610267
	GWS004831	Ridley Island, Prince Rupert, BC	ULVA167-09	HQ603503	HQ610268
Ulvophyceae, Ulotrichales, Gomontiaceae					
Monostroma grevillei (Thuret) Wittrock sp. 1[d]	GWS003588	Cape Neddick, ME	ULVA091-09	HQ603490	HQ610255
	GWS003637	Starboard Cove, ME	ULVA046-09	HQ603492	HQ610257
	GWS003763	Harrington Cove, Grand Manan I., NB	ULVA086-09	HQ603491	HQ610256
	GWS005934	Letete, NB	ULVA214-09	HQ603489	HQ610254
	GWS005974	Lepreau, NB	ULVA228-09	HQ603488	HQ610253

[a] Abbreviations for Provinces as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; PEI, Prince Edward Island, Canada; QC, Quebec, Canada; RI, Rhode Island, USA.
[b] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[c] Based on all of our genetic results these two *Prasiola* sp. are conspecific.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.

Table 2. Collection information and accession numbers for rbcL-3P data generated for the extended sample set of specimens *(cont'd)*

				Genbank accession	
Species	*Voucher #*	*Collection Site*[a]	*BOLD ID*[b]	rbc*L-3P*	tuf*A*
Monostroma grevillei (Thuret) Wittrock sp. 2[d]	GWS003626	Starboard Cove, ME	ULVA022-09	HQ603497	HQ610262
	GWS003787	Meadow Cove, NB	ULVA051-09	HQ603496	HQ610261
	GWS003847	Lepreau, NB	ULVA172-09	HQ603495	HQ610260
	GWS005975	Lepreau, NB	ULVA240-09	HQ603494	HQ610259
	GWS005976	Lepreau, NB	ULVA252-09	HQ603493	HQ610258
Protomonostroma undulatum (Wittrock) K.L. Vinogradova	GWS002668	Cape Elizabeth, near Portland, ME	ULVA013-09	HQ603510	HQ610275
	GWS003587	Cape Neddick, ME	ULVA080-09	HQ603508	HQ610273
	GWS003666	Starboard Cove, ME	ULVA093-09	HQ603507	HQ610272
	GWS003761	Harrington Cove, Grand Manan I., NB	ULVA062-09	HQ603509	HQ610274
Ulvophyceae, Ulotrichales, Ulotrichales					
Acrosiphonia arcta (Dillwyn) J. Agardh	GWS003690	Letete, NB	ULVA071-09	HQ603455	HQ610213
	GWS003692	Letete, NB	ULVA083-09	HQ603454	HQ610212
	GWS003751	Harrington Cove, Grand Manan I., NB	ULVA084-09	HQ603453	HQ610211
	GWS003755	Harrington Cove, Grand Manan I., NB	ULVA014-09	HQ603456	HQ610214
	GWS003796	Meadow Cove, NB	ULVA096-09	HQ603452	HQ610210

[a] Abbreviations for Provinces as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; PEI, Prince Edward Island, Canada; QC, Quebec, Canada; RI, Rhode Island, USA.
[b] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[c] Based on all of our genetic results these two *Prasiola* sp. are conspecific.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.

Table 2. Collection information and accession numbers for rbcL-3P data generated for the extended sample set of specimens *(cont'd)*

				Genbank accession	
Species	*Voucher #*	*Collection Site*[a]	*BOLD ID*[b]	rbc*L-3P*	tuf*A*
	GWS003827	Lepreau, NB	ULVA112-09	HQ603451	HQ610209
	GWS003839	Lepreau, NB	ULVA148-09	HQ603450	HQ610208
	GWS005760	Simpson Island, Bay of Fundy, NB	ULVA268-09	HQ603445	HQ610203
	GWS005863	Lepreau, NB	ULVA200-09	HQ603449	HQ610207
	GWS005879	Letete, NB	ULVA249-09	HQ603448	HQ610206
	GWS005963	Long Eddy Point, Grand Manan I., NB	ULVA274-09	HQ603447	HQ610205
Acrosiphonia coalita (Ruprecht) Scagel et al.	GWS002956	Seppings I., Bamfield, BC	ULVA041-09	HQ603460	HQ610218
	GWS005069	Ridley Island, Prince Rupert, BC	ULVA133-09	HQ603459	HQ610217
Acrosiphonia sonderi (Kützing) Kornmann	GWS002672	Cape Elizabeth, near Portland, ME	ULVA025-09	HQ603466	HQ610224
	GWS003756	Harrington Cove, Grand Manan I., NB	ULVA026-09	HQ603465	HQ610223
	GWS003797	Meadow Cove, NB	ULVA108-09	HQ603463	HQ610221
	GWS003819	Beaver Harbour, NB	ULVA144-09	HQ603462	HQ610220
	GWS007426	Eastport, NL	ULVA064-09	HQ603464	HQ610222
Acrosiphonia sp. 1GWS[d]	GWS002799	Dixon, I., Bamfield, BC	ULVA073-09	HQ603476	HQ610234
	GWS003449	Pachena Beach, Bamfield, BC	ULVA007-09	HQ603478	HQ610236
	GWS003451	Pachena Beach, Bamfield, BC	ULVA031-09	HQ603477	HQ610235

[a] Abbreviations for Provinces as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; PEI, Prince Edward Island, Canada; QC, Quebec, Canada; RI, Rhode Island, USA.
[b] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[c] Based on all of our genetic results these two *Prasiola* sp. are conspecific.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.

Table 2. Collection information and accession numbers for rbcL-3P data generated for the extended sample set of specimens *(cont'd)*

				Genbank accession	
Species	*Voucher #*	*Collection Site*[a]	*BOLD ID*[b]	rbc*L-3P*	tuf*A*
	GWS003983	Seppings I., Bamfield, BC	ULVA162-09	HQ603472	HQ610230
	GWS004092	Ross Islets, Bamfield, BC	ULVA151-09	HQ603474	HQ610232
	GWS004453	Port Renfrew, Vancouver I., BC	ULVA187-09	HQ603469	HQ610227
	GWS004672	Kelsey Bay, Vancouver I., BC	ULVA130-09	HQ603475	HQ610233
	GWS004719	Palmerston Reserve, Vancouver I., BC	ULVA166-09	HQ603471	HQ610229
	GWS004720	Palmerston Reserve, Vancouver I., BC	ULVA178-09	HQ603470	HQ610228
	GWS005075	Ridley Island, Prince Rupert, BC	ULVA157-09	HQ603473	HQ610231
Spongomorpha aeruginosa (Linnaeus) C. Hoek	GWS003854	Lepreau, NB	ULVA113-09	HQ603513	HQ610278
	GWS005235	Churchill, MB	ULVA182-09	HQ603512	HQ610277
Urospora penicilliformis (Roth) Areschoug sp. 1[d]	GWS002687	Samoset Resort, Rockport, ME	ULVA061-09	HQ603674	HQ610440
Ulvophyceae, Ulvales, Kornmanniaceae					
Blidingia marginata (J. Agardh) P.J.L. Dang	GWS004837	Butze Rapids, Prince Rupert, BC	ULVA179-09	HQ603480	HQ610238
Blidingia sp. 2GWS	GWS003770	Harrington Cove, Grand Manan, NB	ULVA027-09	HQ603481	HQ610240
Ulvophyceae, Ulvales, Ulvaceae					
Ulva californica Wille	GWS004221	Whiffen Spit, Vancouver I., BC	ULVA104-09	HQ603518	HQ610283
	GWS004797	Seal Cove, Prince Rupert, BC	ULVA107-09	HQ603517	HQ610282

[a] Abbreviations for Provinces as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; PEI, Prince Edward Island, Canada; QC, Quebec, Canada; RI, Rhode Island, USA.
[b] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[c] Based on all of our genetic results these two *Prasiola* sp. are conspecific.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.

Table 2. Collection information and accession numbers for rbcL-3P data generated for the extended sample set of specimens *(cont'd)*

				Genbank accession	
Species	*Voucher #*	*Collection Site*[a]	*BOLD ID*[b]	rbc*L-3P*	tuf*A*
	GWS005068	Ridley Island, Prince Rupert, BC	ULVA121-09	HQ603516	HQ610281
Ulva compressa Linnaeus	GWS003101	Pachena Beach, Bamfield, BC	ULVA077-09	HQ603530	HQ610294
	GWS003155	Pachena Beach, Bamfield, BC	ULVA088-09	HQ603529	HQ610293
	GWS003574	Lepreau, NB	ULVA032-09	HQ603531	HQ610295
	GWS003928	Dixon, I., Bamfield, BC	ULVA126-09	HQ603527	HQ610291
	GWS003929	Dixon, I., Bamfield, BC	ULVA138-09	HQ603525	HQ610289
	GWS004052	Bradys Beach, Bamfield, BC	ULVA103-09	HQ603528	HQ610292
	GWS004084	Wizard I., Bamfield, BC	ULVA127-09	HQ603526	HQ610290
	GWS004085	Wizard I., Bamfield, BC	ULVA139-09	HQ603524	HQ610288
	GWS004796	Seal Cove, Prince Rupert, BC	ULVA189-09	HQ603523	HQ610287
	GWS005862	Lepreau, NB	ULVA282-09	HQ603522	HQ610286
Ulva gigantea (Kützing) Bliding	GWS005692	Lepreau, NB	ULVA231-09	HQ603536	HQ610300
	GWS005693	Lepreau, NB	ULVA243-09	HQ603534	HQ610298
	GWS005861	Lepreau, NB	ULVA271-09	HQ603533	HQ610297
	GWS005870	Letete, NB	ULVA236-09	HQ603535	HQ610299
Ulva intestinalis Linnaeus	GWS002673	Cape Elizabeth, near Portland, ME	ULVA037-09	HQ603559	HQ610323

[a] Abbreviations for Provinces as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; PEI, Prince Edward Island, Canada; QC, Quebec, Canada; RI, Rhode Island, USA.
[b] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[c] Based on all of our genetic results these two *Prasiola* sp. are conspecific.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.

Table 2. Collection information and accession numbers for rbcL-3P data generated for the extended sample set of specimens *(cont'd)*

				Genbank accession	
Species	*Voucher #*	*Collection Site*[a]	*BOLD ID*[b]	rbc*L-3P*	tuf*A*
	GWS002815	Dixon, I., Bamfield, BC	ULVA085-09	HQ603552	HQ610316
	GWS003095	New River Beach, NB	ULVA065-09	HQ603555	HQ610319
	GWS003584	Cape Neddick, ME	ULVA044-09	HQ603558	HQ610322
	GWS003585	Cape Neddick, ME	ULVA056-09	HQ603556	HQ610320
	GWS003592	Cape Neddick, ME	ULVA045-09	HQ603557	HQ610321
	GWS003622	Starboard Cove, ME	ULVA069-09	HQ603554	HQ610318
	GWS003623	Starboard Cove, ME	ULVA081-09	HQ603553	HQ610317
	GWS003699	Letete, NB	ULVA024-09	HQ603560	HQ610324
	GWS003798	Meadow Cove, NB	ULVA120-09	HQ603550	HQ610314
	GWS003820	Beaver Harbour, NB	ULVA156-09	HQ603546	HQ610310
	GWS003857	Lepreau, NB	ULVA137-09	HQ603549	HQ610313
	GWS004455	Port Renfrew, Vancouver I., BC	ULVA117-09	HQ603551	HQ610315
	GWS004829	Ridley Island, Prince Rupert, BC	ULVA155-09	HQ603547	HQ610311
	GWS005074	Ridley Island, Prince Rupert, BC	ULVA145-09	HQ603548	HQ610312
	GWS005857	Bradys Beach, Bamfield, BC	ULVA247-09	HQ603545	HQ610309
	GWS005858	Bradys Beach, Bamfield, BC	ULVA259-09	HQ603543	HQ610307

[a] Abbreviations for Provinces as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; PEI, Prince Edward Island, Canada; QC, Quebec, Canada; RI, Rhode Island, USA.
[b] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[c] Based on all of our genetic results these two *Prasiola* sp. are conspecific.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.

Table 2. Collection information and accession numbers for rbcL-3P data generated for the extended sample set of specimens *(cont'd)*

				Genbank accession	
Species	*Voucher #*	*Collection Site*[a]	*BOLD ID*[b]	rbc*L-3P*	tuf*A*
	GWS006027	Governor Sprague Bridge 17, Narragansett, RI	ULVA253-09	HQ603544	HQ610308
Ulva lactuca Linnaeus	GWS003444	Pachena Beach, Bamfield, BC	ULVA042-09	HQ603562	HQ610326
	GWS003586	Cape Neddick, ME	ULVA068-09	HQ603574	HQ610338
	GWS003589	Cape Neddick, ME	ULVA009-09	HQ603573	HQ610337
	GWS003651	Cape Neddick, ME	ULVA058-09	HQ603561	HQ610325
	GWS003687	Letete, NB	ULVA047-09	HQ603572	HQ610336
	GWS003760	Harrington Cove, NB	ULVA050-09	HQ603571	HQ610335
	GWS003794	Meadow Cove, NB	ULVA075-09	HQ603570	HQ610334
	GWS003848	Lepreau, NB	ULVA183-09	HQ603569	HQ610333
	GWS004229	Sooke Harbour, BC	ULVA116-09	HQ603568	HQ610332
	GWS004456	Port Renfrew, BC	ULVA129-09	HQ603567	HQ610331
	GWS004717	Raft Cove, BC	ULVA142-09	HQ603566	HQ610330
	GWS005176	Prince Rupert, BC	ULVA134-09	HQ603565	HQ610329
	GWS005244	Churchill, MB	ULVA099-09	HQ603564	HQ610328
	GWS005340	Button Bay, Churchill, MB	ULVA147-09	HQ603563	HQ610327

[a] Abbreviations for Provinces as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; PEI, Prince Edward Island, Canada; QC, Quebec, Canada; RI, Rhode Island, USA.
[b] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[c] Based on all of our genetic results these two *Prasiola* sp. are conspecific.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.

Table 2. Collection information and accession numbers for rbcL-3P data generated for the extended sample set of specimens *(cont'd)*

				Genbank accession	
Species	*Voucher #*	*Collection Site*[a]	*BOLD ID*[b]	rbc*L-3P*	tuf*A*
	GWS005557	Wreck of the Ithaca, east of Churchill, MB	ULVA227-09	HQ603593	HQ610359
	GWS005613	Cape Neddick, ME	ULVA263-09	HQ603586	HQ610352
	GWS005695	Lepreau, NB	ULVA255-09	HQ603588	HQ610354
	GWS005698	Lepreau, NB	ULVA196-09	HQ603600	HQ610366
	GWS005750	Fort Wetherill, RI	ULVA244-09	HQ603590	HQ610356
	GWS005785	Seal Point, PEI	ULVA209-09	HQ603596	HQ610362
	GWS005792	Sea Cow Pond, PEI	ULVA233-09	HQ603592	HQ610358
	GWS005801	Aguilar Point, Bamfield, BC	ULVA269-09	HQ603575	HQ610339
	GWS005841	Scotts Bay, Bamfield, BC	ULVA281-09	HQ603585	HQ610351
	GWS005842	Scotts Bay, Bamfield, BC	ULVA199-09	HQ603599	HQ610365
	GWS005853	Bradys Beach, Bamfield, BC	ULVA211-09	HQ603595	HQ610361
	GWS005871	Letete, NB	ULVA248-09	HQ603589	HQ610355
	GWS005872	Letete, NB	ULVA260-09	HQ603587	HQ610353
	GWS005875	Letete, NB	ULVA201-09	HQ603598	HQ610364
	GWS005876	Letete, NB	ULVA213-09	HQ603594	HQ610360

[a] Abbreviations for Provinces as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; PEI, Prince Edward Island, Canada; QC, Quebec, Canada; RI, Rhode Island, USA.
[b] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[c] Based on all of our genetic results these two *Prasiola* sp. are conspecific.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.

Table 2. Collection information and accession numbers for rbcL-3P data generated for the extended sample set of specimens *(cont'd)*

Species	*Voucher #*	*Collection Site*[a]	*BOLD ID*[b]	*Genbank accession* rbc*L-3P*	*Genbank accession* tuf*A*
	GWS005933	Letete, NB	ULVA202-09	HQ603597	HQ610363
	GWS006041	Governor Sprague Bridge 17, Narragansett, RI	ULVA242-09	HQ603591	HQ610357
Ulva linza Linnaeus	GWS004181	Otter Pt., BC	ULVA186-09	HQ603601	HQ610367
Ulva lobata (Kützing) Harvey	GWS003250	Bamfield, BC	ULVA006-09	HQ603608	HQ610373
	GWS003445	Pachena Beach, BC	ULVA054-09	HQ603607	HQ610372
	GWS003926	Bamfield, BC	ULVA102-09	HQ603606	HQ610371
	GWS004083	Bamfield, BC	ULVA115-09	HQ603604	HQ610369
	GWS004094	Bamfield, BC	ULVA163-09	HQ603605	HQ610370
	GWS005839	Scotts Bay, Bamfield, BC	ULVA258-09	HQ603613	HQ610377
Ulva pertusa Kjellman	GWS003288	Seppings, Bamfield, BC	ULVA018-09	HQ603614	HQ610378
	GWS005800	Aguilar Point, Bamfield, BC	ULVA257-09	HQ603618	HQ610382
	GWS005802	Aguilar Point, Bamfield, BC	ULVA280-09	HQ603617	HQ610381
	GWS005804	Aguilar Point, Bamfield, BC	ULVA198-09	HQ603621	HQ610385
	GWS005834	Scotts Bay, Bamfield, BC	ULVA210-09	HQ603620	HQ610384
	GWS005836	Scotts Bay, Bamfield, BC	ULVA234-09	HQ603619	HQ610383

[a] Abbreviations for Provinces as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; PEI, Prince Edward Island, Canada; QC, Quebec, Canada; RI, Rhode Island, USA.
[b] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[c] Based on all of our genetic results these two *Prasiola* sp. are conspecific.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.

Table 2. Collection information and accession numbers for rbcL-3P data generated for the extended sample set of specimens *(cont'd)*

				Genbank accession	
Species	*Voucher #*	*Collection Site*[a]	*BOLD ID*[b]	rbc*L-3P*	tuf*A*
Ulva procera (K. Ahlner) Hayden et al.	GWS003289	Seppings, Bamfield, BC	ULVA030-09	HQ603627	HQ610391
	GWS003625	Starboard, ME	ULVA010-09	HQ603626	HQ610390
	GWS003694	Letete, NB	ULVA094-09	HQ603625	HQ610389
	GWS003752	Harrington Cove, NB	ULVA095-09	HQ603624	HQ610388
	GWS003765	Harrington Cove, NB	ULVA015-09	HQ603623	HQ610387
	GWS005177	Prince Rupert, BC	ULVA146-09	HQ603622	HQ610386
	GWS005864	Lepreau, NB	ULVA212-09	HQ603629	HQ610393
Ulva prolifera O.F. Müller	GWS003591	Cape Neddick, ME	ULVA033-09	HQ603639	HQ610403
	GWS003716	Letete, NB	ULVA072-09	HQ603638	HQ610402
	GWS003764	Harrington Cove, Grand Manan I., NB	ULVA003-09	HQ603640	HQ610404
	GWS003841	Lepreau, NB	ULVA160-09	HQ603635	HQ610399
	GWS003858	Lepreau, NB	ULVA149-09	HQ603636	HQ610400
	GWS005322	Gordon Point, Churchill, MB	ULVA123-09	HQ603637	HQ610401
Ulva sp. 5GWS	GWS005835	Scotts Bay, Bamfield, BC	ULVA222-09	HQ603668	HQ610432
	GWS005838	Scotts Bay, Bamfield, BC	ULVA246-09	HQ603667	HQ610431
	GWS005840	Scotts Bay, Bamfield, BC	ULVA270-09	HQ603666	HQ610430

[a] Abbreviations for Provinces as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; PEI, Prince Edward Island, Canada; QC, Quebec, Canada; RI, Rhode Island, USA.
[b] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[c] Based on all of our genetic results these two *Prasiola* sp. are conspecific.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.

Table 2. Collection information and accession numbers for rbcL-3P data generated for the extended sample set of specimens *(cont'd)*

				Genbank accession	
Species	*Voucher #*	*Collection Site*[a]	*BOLD ID*[b]	rbc*L-3P*	tuf*A*
Ulva stenophylla Setchell et Gardner	GWS003927	Dixon, I., Bamfield, BC	ULVA114-09	HQ603670	HQ610434
	GWS004599	Pachena Beach, Bamfield, BC	ULVA106-09	HQ603671	HQ610435
Ulva torta (Mertens) Trevisan	GWS004454	Port Renfrew, Vancouver I., BC	ULVA153-09	HQ603672	HQ610436
Ulvaria obscura (Kützing) Gayral	GWS003572	Lepreau, NB	ULVA008-09	HQ603663	HQ610427
	GWS003573	Lepreau, NB	ULVA020-09	HQ603661	HQ610425
	GWS003627	Starboard Cove, ME	ULVA034-09	HQ603659	HQ610423
	GWS003667	Starboard Cove, ME	ULVA011-09	HQ603662	HQ610426
	GWS003673	Starboard Cove, ME	ULVA023-09	HQ603660	HQ610424
	GWS003821	Beaver Harbour, NB	ULVA168-09	HQ603653	HQ610417
	GWS003826	Lepreau, NB	ULVA100-09	HQ603658	HQ610422
	GWS003834	Lepreau, NB	ULVA124-09	HQ603656	HQ610420
	GWS003855	Lepreau, NB	ULVA125-09	HQ603655	HQ610419
	GWS004457	Port Renfrew, Vancouver I., BC	ULVA141-09	HQ603654	HQ610418
	GWS004603	Kye Bay, Vancouver I., BC	ULVA118-09	HQ603657	HQ610421
	GWS005873	Letete, NB	ULVA272-09	HQ603652	HQ610416

[a] Abbreviations for Provinces as follows: BC, British Columbia, Canada; MB, Manitoba, Canada; ME, Maine, USA; NB, New Brunswick, Canada; NL, Newfoundland and Labrador, Canada; NS, Nova Scotia, Canada; PEI, Prince Edward Island, Canada; QC, Quebec, Canada; RI, Rhode Island, USA.
[b] Further information for each sample, including a list of collectors, date of collection and latitudes/longitudes, is available at the Barcode of Life Data Systems (BOLD) website: http://www.barcodinglife.com.
[c] Based on all of our genetic results these two *Prasiola* sp. are conspecific.
[d] These taxa were treated as separate cryptic species in all analyses owing to the levels of genetic diversity uncovered here.

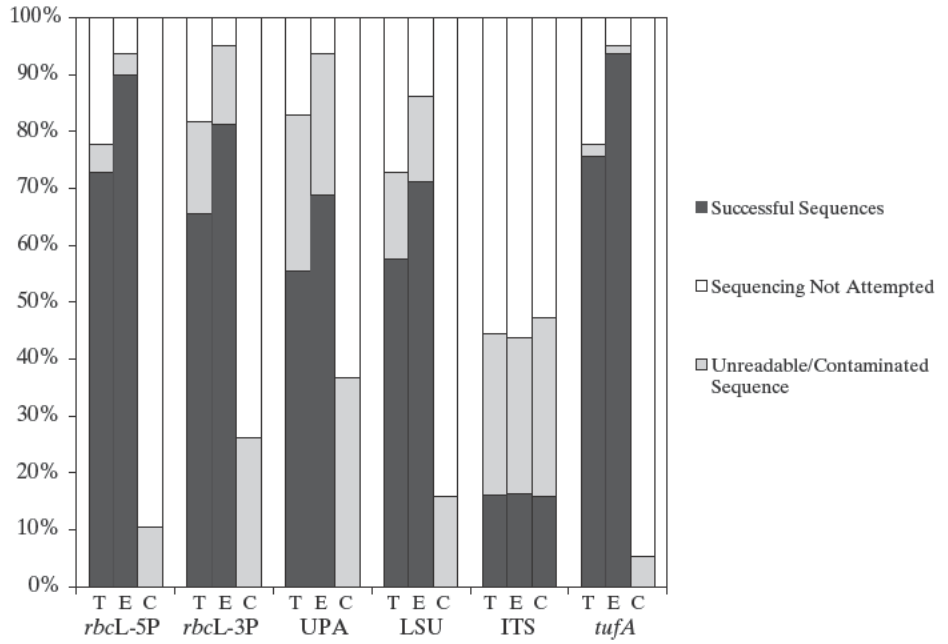


Fig. 1. Universality of each marker given as a percent of samples successfully sequenced. Dark grey indicates percent of samples for which sequence was obtained, light grey indicates sequence that either belonged to a non-target organism or was unreadable due to contamination or read slippage, while white bars indicate failure of PCR or multiple/streaked bands, with sequencing not attempted. First bar (T) – entire test sample set (n=99); second bar (E) – Cladophoraceae excluded (n=80); and third bar (C) – Cladophoraceae only (n=19).

Table 3. A summary of marker performance for the purposes of DNA barcoding

Marker [1]	*Sequence success % (F= forward; R= reverse primers used)*	*Contamination %*	*Medium or high quality trace files %* [2]	*Sequence quality rating = high %* [2]	*Putative introns detected* [3]	*Number of indels per sequence length* [4]	*Species resolution % based on (n=samples)* [4]	*Barcode gap* [4]
*rbc*L-5P	90 (F=2; R=1)	4	65.49	100	No	3/559 bp	100 (n=41)	0.18
*rbc*L-3P	81 (F=1; R=1)	14	92.54	99.57	Yes	0/742 bp	100 (n=139)	0.33
tufA	94 (F=1; R=1)	1	95.15	100	No	0/807 bp	100 (n=134)	0.39
UPA	69 (F=1; R=1)	25	66.98	81.82	Yes	0/369 bp	36 (n=27)	0
LSU D2/D3	71 (F=1; R=1)	15	90.35	100	No	1/567 bp	100 (n=36)	0

[1] Data for *rbc*L-3P and *tufA* based on the extended sample set; other markers based on the test sample set.
[2] These values are an indication of the quality of the raw sequence data and the final edited sequence, respectively, and were generated by BOLD (http://www.boldsystems.org/views/projectmenu.php?&).
[3] Introns remain a possibility for all markers and may account for incomplete sequence success (e.g., introns in the *rbc*L-5P region as defined here are reported in the literature). This value refers only to markers for which introns were unequivocally confirmed here.
[4] This parameter based on data for the genus *Ulva*, for which we have the most species diversity represented including closely related species pairs.

marker system across several kingdoms. More plausibly, it is possible that introns occur within the UPA for some species. Indeed, we observed four cases of putative introns (Tables 1, 3) within our test sample set for Ulva isolates that produced two PCR products of different sizes. When the larger product was isolated and sequenced the initial approximately 100-180 nucleotides matched *Ulva* UPA sequence in BLASTn searches and had GC contents of ~42-53% whereas the subsequent 200-500 nucleotides consisted of AT-rich regions (GC content 20-33%) lacking BLASTn matches.

The LSU-D2/D3 had approximately the same sequencing success as the UPA (Fig. 1; Table 3), but in this case spanned a lower taxonomic breadth with no sequences obtained for any samples of the genera: *Bryopsis*, *Codium*, *Derbesia*, *Kornmannia*, and *Prasiola* (Table 1). The LSU-D2/D3 primers used here are fairly general and have been employed in a variety of diatom (Hamsher *et al*., *in press*), red and brown algal studies in our lab (e.g., Lane *et al.*, 2006; Le Gall & Saunders, 2007; Phillips *et al.*, 2008; Clarkston & Saunders, 2010). We, therefore, expected these primers to have a higher success for green algae than observed here, but this cannot be guaranteed and further primer development may resolve the low levels of success encountered here. Again, given the high incidence of introns within green algae (Bhattacharya *et al.*, 1996; Haugen *et al.*, 2005), introns in the LSU are one possible explanation for our failure to recover sequence in some taxa (although none were confirmed here; Table 3). The presence and co-amplification of contaminating LSU sequences is another possible explanation for the difficulties observed here given that amplification success and levels of contaminated sequence were approximately on par with the UPA.

Sequencing success was lowest for the ITS with only 16 sequences obtained (Table 1). The majority of samples in the test sample set produced multiple or streaked PCR products and sequencing was not attempted. Similar to the LSU-D2/D3, the broad universality of the primers employed (P1 and G4) for red and brown algae (Harper & Saunders, 2001; Ross *et al.*, 2003; Kucera & Saunders, 2008) led us to expect a high level of universality in the green algae as well. It is likely that these primers do amplify the ITS for green algae, but are also amplifying contaminating organisms, leading to multiple-band or streaked PCR product. In an effort to increase specificity for green algae, we tested several published green algal primers (see Materials and Methods) on 16 samples. Unfortunately, the prevalence of multiple bands in PCR product did not decrease, and samples for which single bands were recovered produced contaminant or messy sequence in most cases. Of the 16 samples tested with four primer combinations, we were only able to acquire clean sequence for: single isolates of *Protomonostroma undulatum* and *Urospora* sp. with the primer combinations ITS1/JO6, ITS1/G4, 18S1763/JO6 and 18S1763/G4, one isolate of *Cladophora* sp. with ITS1/G4 and an *Ulvaria obscura* isolate with 18S1763/J06. There was a further problem with ITS, albeit more of an annoyance than a barrier. This region was so variable that it was not possible to align specimens from different genera making it difficult to subject ITS data to routine analyses (alignability owing to excessive indels was not a concern, at least to the genus level, for the other markers tested here; Table 3). We thus abandoned ITS as a candidate barcode marker and did not pursue it further in this study.

One general trend among all markers, except the ITS, was the failure to recover sequence from the included Cladophoraceae. We first questioned the efficacy of our DNA extraction protocol for members of this taxon; however, we excluded this explanation for three reasons: a) published molecular studies in the

Cladophoraceae use similar protocols to ours with significant success (e.g., Leliaert *et al.*, 2007); b) we tried a modified DNA extraction protocol including phenol/chloroform (as published in Saunders (1993) but without the final gel electrophoresis cleaning step) for eight Cladophoraceae samples and still failed to get *rbc*L, *tufA*, UPA or LSU-D2/D3 sequence; and c) we were able to recover ITS sequences from some samples indicating that our extraction protocol was, in fact, working. As in other cases of failure to recover sequence, primer mis-match or introns may explain why we were unable to obtain *rbc*L, *tufA*, UPA or LSU-D2/D3 sequence for the Cladophoraceae. For the chloroplast markers, attempts in other labs have similarly failed to recover plastid DNA sequence from Cladophoraceae and there are no plastid sequences for any representative of this taxon in Genbank. For our nuclear markers, we compared published LSU sequences for *Cladophora* (Leliaert *et al.*, 2007) with our primers and found that the T16N primer was mismatched with the Cladophora sequences at two sites, one of them being the terminal 3' nucleotide (G). Unfortunately, these published sequences did not overlap with the region where the T24 primer is anchored so we were unable to evaluate how well this primer would match with *Cladophora*. Primer mismatch seems the more likely explanation for lack of LSU-D2/D3 sequence success for the Cladophoraceae in the current study.

Species-discrimination power of the markers

Higher species representation for *Acrosiphonia* and *Ulva* allowed for a comparison of intra- versus interspecific distances in these genera, especially in the *rbc*L and *tufA* markers (Table 1, Fig. 2). For *rbc*L-5P, *rbc*L-3P and *tufA,* samples assigned to the same species clustered together in neighbor-joining analysis forming discrete monophyletic groups and interspecific divergences were noted among all included species (Fig. 2). In the *rbc*L-3P, a barcoding gap of 1.49% divergence and 0.68% divergence was observed for *Acrosiphonia* and *Ulva,* respectively. The *rbc*L-5P had only a 0.18% barcoding gap for *Ulva* (differences in the species successfully amplified for *rbc*L-5P from *Acrosiphonia* relative to *rbc*L-3P precluded a meaningful interspecific comparison) and showed no intraspecific divergence within *Acrosiphonia* and *Ulva* indicating lower variability for this marker. The *tufA* region had a similar barcode gap to *rbc*L-3P at 1.82% divergence and 0.52% divergence for *Acrosiphonia* and *Ulva*, respectively. Conversely, for the UPA and the LSU-D2/D3, not all species formed independent clusters owing to the absence of a barcode gap and some interspecific values as low as zero were observed (Fig. 2). For example, of the *Ulva* species, UPA failed to resolve differences among *U. californica, U. flexuosa, U. linza, U. procera, U.* sp. 1GWS*, U. stenophylla* and *U. torta* (Table 3), as well as between *Urospora* sp. 2*penicilliformis* and *Uropora wormskioldii* (not shown), while the two *Codium* spp., *C. fragile* and *C.* sp. 1*fragile,* differed by only a single substitution. On the intraspecific level, no variation at all was recorded within *Ulvaria obscura*, in contrast to *rbc*L-5P, *rbc*L-3P and *tufA* (Fig. 2). For the LSU-D2/D3, absence of variation was noted among all the species of *Acrosiphonia* while *Ulva* virtually lacked a barcode gap (i.e., intra- and interspecific variation almost overlapped; Fig. 2), although in the latter case species did resolve as monophyletic groups (Table 3) in neighbor-joining analysis (not shown).

Of the markers tested here, the *tufA* had the best combination of universality, sequence quality and discriminatory power, while *rbc*L-3P had similar levels of discriminatory power, but suffered from reduced universality (introns responsible, at least in part, for this weakness) (Table 3). We thus investigated the

Genus (n=total number of specimens	rbc*L-5P*				rbc*L-3P*				tufA				*UPA*				*LSU-D2/D3*			
	# seq.	*# spp.*	*intra*	*inter*	*# seq.*	*# spp.*	*intra*	*inter*	*# seq.*	*# spp.*	*intra*	*inter*	*# seq.*	*# spp.*	*intra*	*inter*	*# seq.*	*# spp.*	*intra*	*inter*
Acrosiphonia (n=7)	6	5	0	2.33-7.35	6	4	0-0.14	1.63-5.85	6	4	0-0.13	1.95-11.24	4	4	ID	1.08-3.26	5	4	0	0
Blidingia (n=3)	1	1	ID	ID	1	1	ID	ID	3	3	ID	10.6-14.2	2	2	ID	5.15	2	2	ID	2.18
Bryopsis (n=2)	2	1	0	ID	2	1	0	ID	2	1	0	ID	2	1	0	ID	0	0	ID	ID
Cladophora (n=9)	0	0	ID	ID	0	0	ID	ID	0	0	ID	ID	0	0	ID	ID	0	0	ID	ID
Codium (n=4)	2	2	ID	0	0	0	ID	ID	4	4	ID	1.44-14.65	2	2	ID	0.27	0	0	ID	ID
Prasiola (n=2)	2	1	0.19	ID	1	1	ID	ID	2	1	0	ID	1	1	ID	ID	0	0	ID	ID
Protomonostroma (n=3)	3	1	0	ID	3	1	0.27	ID	3	1	0-0.13	ID	3	1	0	ID	3	1	0	ID
Ulva (n=42)	41	14	0	0.18-3.42	38	14	0-0.27	0.95-4.76	39	13	0-0.39	0.91-9.1	27	11	0	0-1.5	36	13	0-0.35	0.35-3.17
Ulvaria (n=11)	11	1	0-0.18	ID	10	1	0-0.47	ID	11	1	0-0.39	ID	10	1	0	ID	8	1	0	ID
Urospora (n=3)	3	3	ID	0.18	2	2	ID	2.13	2	2	ID	4.37	2	2	ID	0	2	2	ID	0.17

Fig. 2. Summary of intra- and interspecific divergence for each marker (excepting ITS) in the test sample set for genera for which more than one specimen was attempted and at least a single sequence was successfully generated. Abbreviations as follows: # seq., total number of sequences compared; #spp., number of genetic species clusters; intra, range of intraspecific divergence (%); inter, range of interspecific divergence (%); ID, insufficient data. *Prasiola meridionalis* and *P. stipitata* were treated as conspecific in these analyses (see Table 1).

intra- versus interspecific divergences in the *tufA* and *rbc*L-3P for an additional 164 samples (Table 2). Neighbor-joining analyses (result for *tufA* presented in Fig. 3) of these sequences combined with the respective test sample set sequences (= extended sample set) produced genetic species clusters congruent with the individual test sample set analyses and relative to one another. The DNA barcoding gaps were greater for *tufA* than *rbc*L-3P (Fig. 4) and largely similar to those uncovered for the same genera in the test sample set (Fig. 2). An exception occurred for the genus *Ulva*, which had barcode gaps of 0.52 and 0.68 for *tufA* and *rbc*L, respectively, in the test sample set (Fig. 2), but only 0.39 and 0.33, respectively, in the extended analysis (Fig. 4; Table 3). Since taxon sampling was the highest in *Ulva* (n=140), the likelihood of encountering closely related species pairs was also highest (e.g., Meier *et al.*, 2008), consistent with the observation that the lowest interspecific divergences were observed in this genus. Higher sampling overall also means that isolates with high intraspecific divergences, perhaps representing isolated populations, are more likely to be encountered than for groups in which sampling was lower. Indeed, the highest intraspecific variation in *Ulva* was 0.52% (4 bp) in *Ulva procera* for *tufA* and 0.41% (3bp) in *Ulva intestinalis* for *rbc*L-3P. In the case of *tufA*, two discrete haplotypes were uncovered both occurring in the Atlantic with only one variant thus far encountered in the Pacific. An increase in taxon sampling among the other genera studied would allow a more detailed analysis of intra- versus interspecific divergence. The magnitude of the barcoding gap will not necessarily be the same among different genera, but wider ranges of both intra- and interspecific divergences are likely to be observed as more individuals and taxa are sampled.

Taxonomic observations and putative cryptic species

Based on our survey of intra- versus interspecific variation, this study has uncovered several putative cryptic species. Despite the thorough study of *Ulva* conducted in the northeast Pacific by Hayden & Waaland (2004), several genetic species groups of *Ulva* did not match published data and were more than 1.83% and 1.1-2.72% divergent in their *tufA* and *rbc*L-3P, respectively, from their nearest neighbor suggesting the occurrence of cryptic or overlooked species. These include: *Ulva* sp. 1GWS, a single collection from an estuarine habitat at the end of a long, sheltered marine inlet; and *Ulva* sp. 5GWS, consisting of three winter collections from a single site in Bamfield, British Columbia. As most of the specimens studied by Hayden & Waaland (2004) were collected during the summer months, *Ulva* sp. 5GWS is perhaps a winter annual.

Samples morphologically identified as *Acrosiphonia arcta* (Dillwyn) Gain fell into two distinct genetic species groups: "*Acrosiphonia arcta*" and "*Acrosiphonia* sp. 1GWS" (1.95 and 1.64% divergent in *tufA* and *rbc*L-3P, respectively). The "*A.* sp. 1GWS" samples were all collected from the Pacific, whereas the "*A. arcta*" samples were from the Atlantic indicating that *A. arcta* as currently considered may not inhabit both oceans.

Samples identified as *Monostroma grevillei* (Thuret) Wittrock also segregated into two clusters – *M. grevillei* sp. 1 and *M. grevillei* sp. 2 (1.16 and 1.09% divergence in the *tufA* and *rbc*L-3P, respectively). Although the floristic guide to the northwest Atlantic flora (Sears, 2002) lists a second species of *Monostroma* (*M. oxyspermum* (Kützing) Doty), this species is now considered to belong to the genus *Gayralia. Gayralia oxysperma* (Kützing) K.L. Vinogradova *ex* Scagel *et al.*, *rbc*L sequence from Genbank did not match either of our *Monostroma* genetic

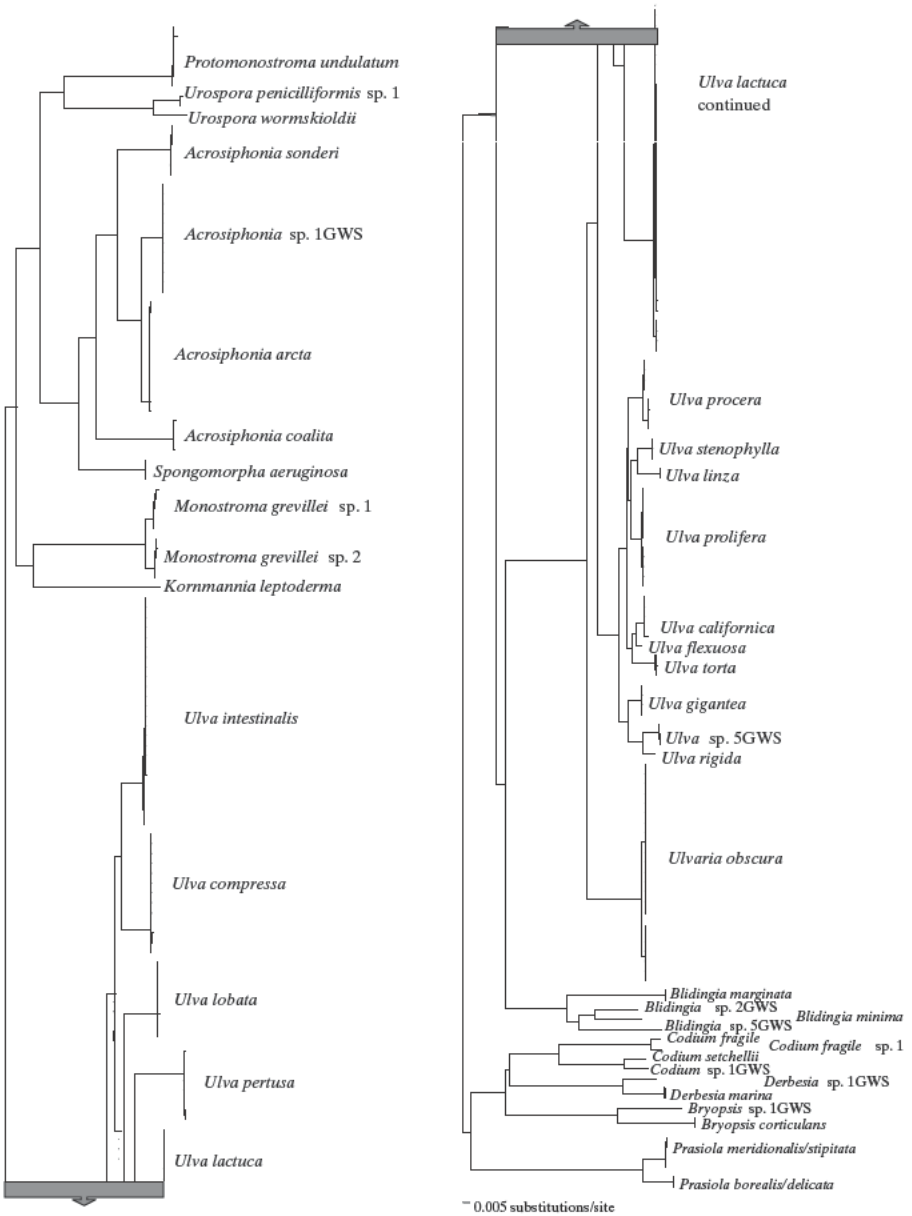


Fig. 3. Unrooted neighbor-joining phylogram for extended sample set *tufA* alignment including sequences from the test sample set plus 164 additional sequences.

species groups, supporting the presence of cryptic species within *Monostroma grevillei* in the northwest Atlantic.

Samples of *Ulvaria obscura* (Kützing) P. Gayral *ex* C. Bliding fell into two clusters (Fig. 3), which differed by 0.39 and 0.41% divergence in *tufA* and *rbc*L-3P, respectively (Figs 2 and 4). These values are approaching the upper threshold for

intraspecific variation observed here and have a biogeographic component with one cluster consisting of our Atlantic and the other our Pacific collections. These two clusters likely represent two isolated populations of *Ulvaria obscura* although additional study to assess the status of these groups is warranted.

Results for both *tufA* (Fig. 3) and *rbc*L-3P led to some interesting observations within the genus *Prasiola.* In the Canadian Atlantic, only a single species, *P. stipitata* Suhr ex Jessen, is recognized (Sears, 2002). Samples of *P. stipitata* grouped with and were identical in *tufA* and *rbc*L-3P to Pacific samples identified as *Prasiola meridionalis* Setchell *et* N.L. Gardner (based on our interpretation of Gabrielson *et al.*, (2006)). Whether these two entities are conspecific requires further alpha taxonomic investigation (see Rindi *et al.*, 2007). Another *Prasiola* genetic cluster consisted of two samples with vastly different cellular arrangements. Sample GWS005101 consists of cells arranged in groups of 4-8, and what appeared to be a fungal infection (Fig. 5) – both of these characteristics typical of *P. borealis* M. Reed (Gabrielson *et al.*, 2006). On the other hand, sample GWS005076, *Prasiola delicata* Setchell *et* N.L. Gardner, had cells arranged more regularly, not in groups of 4, and did not have an obvious fungal infection (Fig. 6). Given that these two samples had identical *tufA* and *rbc*L sequences, we hypothesize that they may belong to the same species (*P. borealis* having priority, however, see comments below) and that the fungal infection is responsible for the morphology observed in GWS005101. However, Rindi *et al.* (2007, fig. 1) associated the *P. borealis* morph with an unidentified species in the *P. crispa* (Lightfoot) Kützing cluster rather than with *P. delicata* as uncovered here (Fig. 3) suggesting that the contemporary concept of *P. borealis* may include a variety of 'fungal-infected' species (lichenization; Pérez-Ortega *et al.*, 2010). Sequence data from the type or topotype collections are thus essential to establish the true identity of the phycobiont to which the name *P. borealis* should be applied. Both of the previous hypotheses (viz. synonym for *P. stipitata/meridionalis* and *P. borealis/delicata*) are based on only a few samples and require further investigation. An alternative explanation for both of these observations is that plastid introgression has occurred between species in these two putative pairs of taxa.

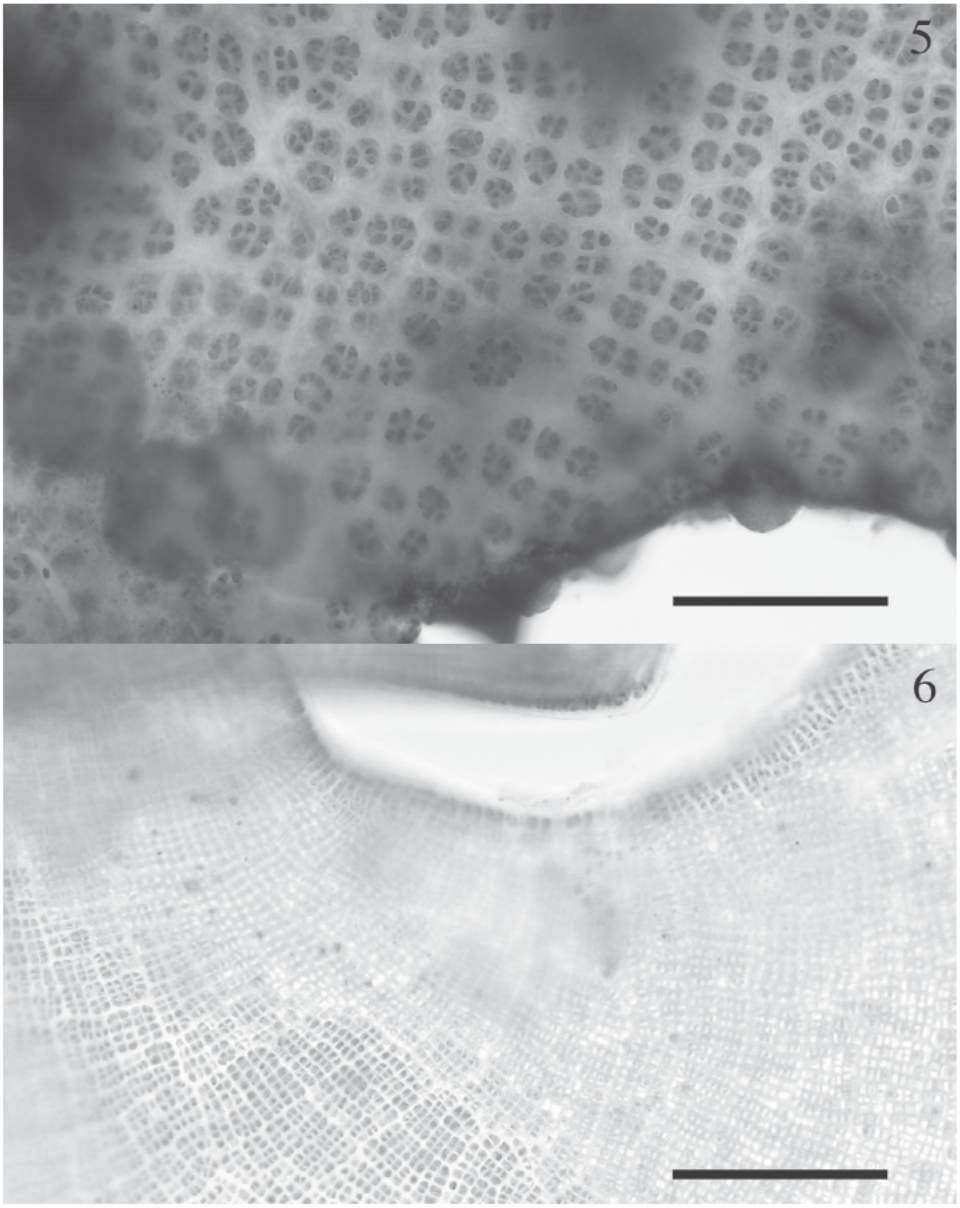
Conclusions and future research

In light of the research here, the high amplification success and sequence quality relative to other markers, and apparent lack of introns and slightly wider barcode gap relative to *rbc*L-3P, *tufA* is the best candidate for a green macroalgal barcode (Table 3; also see Händeler *et al.*, 2010). We are also advocating LSU-D2/D3 as a secondary marker for all protists (for which it is not the primary marker; see Hamsher *et al.*, *in press*). While the ability of this marker to provide reasonable species discrimination is supported here, substantial primer design (assuming that the presence of introns is not the main cause for our low success at amplifying this marker from green algae) is necessary before this marker can be applied to marine green macroalgae in this capacity. Although not as variable as the primary barcode marker recommended, it can typically assign samples to species clusters (Table 3) and will provide a tool for more global environmental surveys for which researchers are willing to sacrifice some taxonomic resolution for broader universality (Hamsher *et al.*, *in press*).

The previous conclusions obviously do not apply to the Cladophoraceae, a taxon (perhaps even to higher taxonomic levels, e.g., Cladophorales) for which further marker development is needed (note that Händeler *et al.*, (2010) indicated

Species with number of samples compared for (tufA*:* rbc*L)*	tufA				rbc*L-3P*			
	# of species on which intra; inter specific comparisons were calculated	*Intra-specific Min - Max*	*Inter-specific Min - Max*	*Gap*	*# of species on which intra; inter specific comparisons were calculated*	*Intra-specific Min-Max*	*Inter-specific Min-Max*	*Gap*
Acrosiphonia (34 ; 34)	4 ; 4	0 - 0.28	1.95 - 11.38	1.67	4 ; 4	0 - 0.27 [a]	1.64 - 6.12	1.37
Blidingia (5 ; 3)	1 ; 4	0	10.58 - 14.20	10.58	1 ; 2	0	8.61	8.61
Bryopsis (3 ; 3)	1 ; 2	0	10.57	10.57	1 ; 2	0	5.65	5.65
Codium (4 ; 0)	0 ; 4	ID	1.44 - 14.65	ID	0 ; 0	ID	ID	ID
Derbesia (3 ; 3)	1 ; 2	0.13	6 - 6.14	5.87	1 ; 2	0	3.17 - 3.20	3.17
Monostroma (10 ; 10)	2 ; 2	0	1.16 - 1.30	1.16	2 ; 2	0 - 0.27	1.09 - 1.10	0.82
Prasiola (6 ; 6)	2 ; 2 [b]	0	3.79 - 3.89	3.79 [b]	2 ; 2 [b]	0	1.36 - 1.43	1.36 [b]
Protomonostroma (7 ; 7)	1 ; 0	0 - 0.13	ID	ID	1 ; 0	0 - 0.27	ID	ID
Spongomorpha (3 ; 3)	1 ; 0	0	ID	ID	1 ; 0	0 - 0.27	ID	ID
Ulva (160 ; 160)	13 ; 15	0 - 0.52	0.91 - 10.16	0.39	13 ; 16	0 - 0.41	0.74 - 4.95	0.33
Ulvaria (23 ; 23)	1 ; 0	0 - 0.39	ID	ID	1 ; 0	0 - 0.41	ID	ID
Urospora (3 ; 3)	1 ; 2	0.26	4.37 - 4.38	4.11	0 ; 3	ID	1.12 - 2.36	ID

Fig. 4. Intra- and interspecific variation (% divergence) for the extended sample set *tufA* and *rbc*L-3P alignments (test sample set plus 164 additional samples). Abbreviation ID stands for insufficient data. [a] This value excludes *Acrosiphonia sonderi* GWS003819, which has an anomalous *rbc*L-3P sequence (does not match any of our other taxa, so it is not an obvious contaminant, but is 2.1% divergent from the other collections for *rbc*L-3P while identical in *tufA*). [b] *Prasiola meridionalis* from the Pacific and *Prasiola stipitata* from the Atlantic appear to be the same species. Also *P. borealis* and *P. delicata* seem to represent fungal infected and noninfected morphs, respectively, of a single species. Data were thus clustered to two groups, *P. meridionalis/stipitata* and *P. borealis/delicata*, for these analyses (see Tables 1 and 2).



Figs 5-6. Cellular arrangement of *Prasiola*. 5. "*Prasiola borealis*". 6. "*Prasiola delicata*". Scale bars = 200 µm.

that *tufA* data had been acquired for this members of this taxon, but none was included in their paper so this result awaits confirmation). Subtle modifications to the LSU-D2/D3 primers should easily facilitate amplification of the secondary marker, but identification of a primary marker may prove a considerable challenge. Alternatively, as in land plants, a combination of markers may need to

be used to achieve a balance of universality and species discrimination (see Newmaster *et al.*, 2006; Hollingsworth *et al.*, 2009a; Hollingsworth *et al.*, 2009b).

Acknowledgements. We would like to thank Andrew Blakney, Bridgette Clarkston, Susan Clayden, Natasha Chisti, M. Corey, Jeremy De Waard, Kyatt Dixon, Bob Hooper, Chris Lane, Line Le Gall, Dan McDevit, Tanya Moore, Kathryn Roy, Jose Maria Utge Buil, and Rodney Withall for contributions to sample collection. We are grateful to the Bamfield Marine Sciences Centre for hosting a large component of the fieldwork for this study. For technical assistance in the laboratory, we thank Dan McDevit and Ross Campbell. Heroen Verbruggen is gratefully acknowledged for recommending that we trial *tufA* as a possible green algal marker. This research was supported through funding to the Canadian Barcode of Life Network from Genome Canada (through the Ontario Genomics Institute), the Natural Sciences and Engineering Research Council of Canada and other sponsors listed at www.bolnet.ca. The Canada Research Chair Program, the Canada Foundation for Innovation and the New Brunswick Innovation Foundation provided additional support.

REFERENCES

ALVAREZ I. & WENDEL J.F., 2003 — Ribosomal ITS sequences and plant phylogenetic inference. *Molecular phylogenetics and evolution* 29: 417-434.

BAKKER F.T., OLSEN J.L. & STAM W.T., 1995 — Evolution of nuclear rDNA ITS sequences in the *Cladophora albida/sericea* clade (Chlorophyta). *Journal of molecular evolution* 40: 640-651.

BAKKER F.T., OLSEN J.L., STAM W.T. & VANDENHOEK C., 1992 — Nuclear ribosomal DNA internal transcribed spacer regions (ITS1 and ITS2) define discrete biogeographic groups in *Cladophora albida* (Chlorophyta). *Journal of phycology* 28: 839-845.

BHATTACHARYA D., FRIEDL T. & DAMBERGER S., 1996 — Nuclear-encoded rDNA group I introns: Origin and phylogenetic relationships of insertion site lineages in the green algae. *Molecular biology and evolution* 13: 978-989.

BLOMSTER J., MAGGS C.A. & STANHOPE M.J., 1998 — Molecular and morphological analysis of *Enteromorpha intestinalis* and *E. compressa* (Chlorophyta) in the British Isles. *Journal of phycology* 34: 319-340.

CLARKSTON B.E. & SAUNDERS G.W., 2010 — A comparison of two DNA barcode markers for species discrimination in the red algal family Kallymeniaceae (Gigartinales) with a description of *Euthora timburtonii*. *Botany* 88: 119-131.

DE CLERCK O., VERBRUGGEN H., HUISMAN J.M., FAYE E.J., LELIAERT F., SCHILS T. & COPPEJANS E., 2008 — Systematics and biogeography of the genus *Pseudocodium* (Bryopsidales, Chlorophyta), including the description of *P. natalense* sp. nov. from South Africa. *Phycologia* 47: 225-235.

DRUMMOND A., ASHTON B., CHEUNG M., HELED J., KEARSE M., MOIR R., STONES-HAVAS S., THIERER T. & WILSON A., 2009 — Geneious. v4.8. Avaliable from: http://www.geneious.com

FAMA P., WYSOR B., KOOISTRA W. & ZUCCARELLO G.C., 2002 — Molecular phylogeny of the genus *Caulerpa* (Caulerpales, Chlorophyta) inferred from chloroplast tufA gene. *Journal of phycology* 38: 1040-1050.

FELINER G.N. & ROSSELLO J.A., 2007 — Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Molecular phylogenetics and evolution* 44: 911-919.

FERRI E., BARBUTO M., BAIN O., GALIMBERTI A., UNI S., GUERRERO R., FERTE H., BANDI C., MARTIN C. & CASIRAGHI M., 2009 — Integrated taxonomy: traditional approach and DNA barcoding for the identification of filarioid worms and related parasites (Nematoda). *Frontiers in Zoology* 6: 1.

GABRIELSON P.W., WIDDOWSON T.B. & LINDSTROM S.C., 2006 — *Keys to the Seaweeds and Seagrasses of Southeast Alaska, British Columbia, Washington and Oregon.* Vancouver, University of British Columbia, 209 p.

GEISER D.M., KLICH M.A., FRISVAD J.C., PETERSON S.W., VARGA J. & SAMSON R.A., 2007 — The current status of species recognition and identification in *Aspergillus*. *Studies in mycology* 59: 1-10.

GRIBBLE K.E. & ANDERSON D.M., 2007 — High intraindividual, intraspecific, and interspecific variability in large-subunit ribosomal DNA in the heterotrophic dinoflagellates *Protoperidinium*, *Diplopsalis*, and *Preperidinium* (Dinophyceae). *Phycologia* 46: 315-324.

HAJIEGHRARI B., TORABI-GIGLOU M. & WAEYENBERGE L., 2007 — Comparative d2/d3 LSU–rDNA sequence study of some Iranian *Pratylenchus loosi* populations. *African journal of biotechnology.* 6: 2458-2466.

HALL J.D., FUČÍKOVÁ K., LO, C., LEWIS L.A. & KAROL K. G., 2010 — Assessing proposed DNA barcodes in the green algae. *Cryptogamie, Algologie* 31(4): 529-555.

HAMSHER S.E., EVANS K.M., MANN D.G., POULÍČKOVÁ A. & SAUNDERS G.W. — Barcoding diatoms: exploring alternatives to COI-5P. *Protist* in press.

HÄNDELER K., WÄGELE H., WAHRMUND U., RÜDINGER M. & KNOOP V., 2010 — Slugs' last meals: molecular identification of sequestered chloroplasts from different algal origins in Sacoglossa (Opisthobranchia, Gastropoda). *Molecular ecology resources* 10: 968-978.

HANYUDA T., ARAI S. & UEDA K., 2000 — Variability in the *rbcL* introns of Caulerpalean algae (Chlorophyta, Ulvophyceae). *Journal of plant research* 113: 403-413.

HARPER J.T. & SAUNDERS G.W., 2001 — The application of sequences of the ribosomal cistron to the systematics and classification of the florideophyte red algae (Florideophyceae, Rhodophyta). *Les cahiers de biologie marine* 42: 25-38.

HAUGEN P., SIMON D.M. & BHATTACHARYA D., 2005 — The natural history of group I introns. *Trends in genetics* 21: 111-119.

HAYDEN H.S. & WAALAND J.R., 2002 — Phylogenetic systematics of the Ulvaceae (Ulvales, Ulvophyceae) using chloroplast and nuclear DNA sequences. *Journal of phycology* 38: 1200-1212.

HAYDEN H.S., BLOMSTER J., MAGGS C.A., SILVA P.C., STANHOPE M.J. & WAALAND J.R., 2003 — Linnaeus was right all along: *Ulva* and *Enteromorpha* are not distinct genera. *European journal of phycology* 38: 277-294.

HAYDEN H.S. & WAALAND J.R., 2004 — A molecular systematic study of *Ulva* (Ulvaceae, Ulvales) from the northeast Pacific. *Phycologia* 43: 364-382.

HEBERT P.D.N., CYWINSKA A., BALL S.L. & DEWAARD J.R., 2003a — Biological identifications through DNA barcodes. *Proceedings of the Royal society of London B Biology* 270: 313-321.

HEBERT P.D.N., RATNASINGHAM S. & DEWAARD J.R., 2003b — Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal society of London B Biology* 270: S96-99.

HEBERT P.D.N., STOECKLE M.Y., ZEMLAK T.S. & FRANCIS C.M., 2004 — Identification of birds through DNA barcodes. *PLoS Biology* 2: 1657-1663.

HERATY J.M., WOOLLEY J.B., HOPPER K.R., HAWKS D.L., KIM J.W. & BUFFINGTON M., 2007 — Molecular phylogenetics and reproductive incompatibility in a complex of cryptic species of aphid parasitoids. *Molecular phylogenetics and evoltion* 45: 480-493.

HOLLINGSWORTH M.L., CLARK A.A., FORREST L.L., RICHARDSON J., PENNINGTON R.T., LONG D.G., COWAN R., CHASE M.W., GAUDEUL, M. & HOLLINGSWORTH P.M., 2009a — Selecting barcoding loci for plants: evaluation of seven candidate loci with species-level sampling in three divergent groups of land plants. *Molecular ecology resources* 9: 439-457.

HOLLINGSWORTH P.M., FORREST L.L., SPOUGE J.L., HAJIBABAEI M., RATNASINGHAM S., VAN DER BANK M., CHASE M.W., COWAN R.S., ERICKSON D.L., FAZEKAS A. J., GRAHAM S.W., JAMES K.E., KIM K. J., KRESS W.J., SCHNEIDER H., VAN ALPHENSTAHL J., BARRETT S. C.H., VAN DEN BERG C., BOGARIN D., BURGESS K.S., CAMERON K. M., CARINE M., CHACON J., CLARK A., CLARKSON J.J., CONRAD F., DEVEY D. S., FORD C.S., HEDDERSON T.A.J., HOLLINGSWORTH M. L., HUSBAND B.C., KELLY L.J., KESANAKURTI P.R., KIM J.S., KIM Y. D., LAHAYE R., LEE H.L., LONG D.G., MADRINAN S., MAURIN O., MEUSNIER I., NEWMASTER S.G., PARK C.W., PERCY D.M., PETERSEN G., RICHARDSON J. E., SALAZAR G. A., SAVOLAINEN V., SEBERG O., WILKINSON M.J., YI D.K., LITTLE D.P., 2009b — A DNA barcode for land plants. *Proceedings of the National Academy of Sciences U.S.A.* 106: 12794-12797.

KUCERA H. & SAUNDERS G.W., 2008 — Assigning morphological variants of *Fucus* (Fucales, Phaeophyceae) in Canadian waters to recognized species using DNA barcoding. *Botany* 86: 1065-1079.

LANE C.E., MAYES C., DRUEHL L.D. & SAUNDERS G.W., 2006 — A multi-gene molecular investigation of the kelp (Laminariales, Phaeophyceae) supports substantial taxonomic re-organization. *Journal of phycology* 42: 493-512.

LARKIN M.A., BLACKSHIELDS G., BROWN N.P., CHENNA R., MCGETTINGAN P.A., MCWILLIAMS H., VALENTIN F., WALLACE I.M., WILM A., LOPEZ R.,

THOMPSON J.D., GIBSON T.J. & HIGGINS D.G., 2007 — Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948.

LE GALL L. & SAUNDERS G.W., 2007 — A nuclear phylogeny of the Florideophyceae (Rhodophyta) inferred from combined EF2, small subunit and large subunit ribosomal DNA: establishing the new red algal subclass Corallinophycidae. *Molecular phylogenetics and evolution* 43: 1118-1130.

LE GALL L. & SAUNDERS G.W., 2010 — DNA barcoding is a powerful tool to uncover algal diversity: a case study of the Phyllophoraceae (Gigartinales, Rhodophyta) in the Canadian flora. *Journal of phycology* 46: 374-389.

LELIAERT F., DE CLERCK O., VERBRUGGEN H., BOEDEKER C. & COPPEJANS E., 2007 — Molecular phylogeny of the Siphonocladales (Chlorophyta : Cladophorophyceae). *Molecular phylogenetics and evolution* 44: 1237-1256.

LINDSTROM S.C. & HANIC L.A., 2005 — The phylogeny of North American *Urospora* (Ulotrichales, Chlorophyta) based on sequence analysis of nuclear ribosomal genes, introns and spacers. *Phycologia* 44: 194-201.

MADDISON W.D. & MADDISON D.R., 2005 — MacClade v4.08. Sinauer Associates, Sunderland, Massachusetts.

MANHART J.R., 1994 — Phylogenetic Analysis of Green Plant rbcL Sequences. *Molecular phylogenetics and evolution* 3: 114-127.

MCDEVIT D.C. & SAUNDERS G.W., 2009 — On the utility of DNA barcoding for species differentiation among brown macroalgae (Phaeophyceae) including a novel extraction protocol. *Phycological research* 57: 131-141.

MEIER R., ZHANG G. & ALI F., 2008 — The use of mean instead of smallest interspecific distances exaggerates the size of the "Barcoding Gap" and leads to misidentification. *Systematic biology* 57: 809-813.

NEWMASTER S.G., FAZEKAS A.J. & RAGUPATHY S., 2006 — DNA barcoding in land plants: evaluation of *rbcL* in a multigene tiered approach. *Canadian journal of botany* 84: 335-341.

O' KELLY C.J., WYSOR B. & BELLOWS W.K., 2004 — Gene sequence diversity and the phylogenetic position of algae assigned to the genera *Phaeophila* and *Ochlochaete* (Ulvophyceae, Chlorophyta). *Journal of phycology* 40: 789-799.

PÉREZ-ORTEGA S., DE LOS RÍOS A., CRESPO A. & SANCHO L.G., 2010 — Symbiotic lifestyle and phylogenetic relationships of the bionts of *Mastodia tessellata* (Ascomycota, incertae sedis). *American journal of botany* 97: 738-752.

PHILLIPS N., BURROWES R., ROUSSEAU F., DE REVIERS B. & SAUNDERS G.W., 2008 — Resolving evolutionary relationships among the brown algae using chloroplast and nuclear genes. *Journal of phycology* 44: 394-405.

POMBERT J.F., BEAUCHAMP P., OTIS C., LEMIEUX C. & TURMEL M., 2006 — The complete mitochondrial DNA sequence of the green alga *Oltmannsiellopsis viridis*: evolutionary trends of the mitochondrial genome in the Ulvophyceae. *Current genetics* 50: 137-147.

POMBERT J.F., OTIS C., LEMIEUX, C. & TURMEL M., 2004 — The complete mitochondrial DNA sequence of the green alga *Pseudendoclonium akinetum* (Ulvophyceae) highlights distinctive evolutionary trends in the chlorophyta and suggests a sister-group relationship between the Ulvophyceae and Chlorophyceae. *Molecular biology and evolution* 21: 922-935.

PRESTING G.G., 2006 — Identification of conserved regions in the plastid genome: implications for DNA barcoding and biological function. *Canadian journal of botany* 84: 1434-1443.

RINDI F., MCIVOR L., SHERWOOD A.R., FRIEDL T., GUIRY M.D. & SHEATH R.G., 2007 — Molecular phylogeny of the green algal order Prasioales (Trebouxiophyceae, Chlorophyta). *Journal of phyclogy* 43: 811-822.

ROBBA L., RUSSELL S.J., BARKER G.L. & BRODIE J., 2006 — Assessing the use of the mitochondrial cox1 marker for use in DNA barcoding of red algae (Rhodophyta). *American journal of botany* 93: 1101-1108.

ROSS P.I., DONALDSON S.L. & SAUNDERS G.W., 2003 — A molecular investigation of *Mazzaella* (Gigartinales, Rhodophyta) morphologically intermediate between *Mazzaella linearis* and *M. splendens*. *Botanica marina* 46: 202-213.

SAUNDERS G.W., 1993 — Gel purification of red algal genomic DNA – an inexpensive and rapid method for the isolation of polymerase chain reaction-friendly DNA. *Journal of phycology* 29: 251-254.

SAUNDERS G.W., 2005 — Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. *Philosophical transactions of the royal society of London B Biological sciences* 360: 1879-1888.

SAUNDERS G.W. & LEHMKUHL V.K., 2005 — Molecular divergence and morphological diversity among four cryptic species of *Plocamium* (Plocamiales, Florideophyceae) in northern Europe. *European journal of phycology* 40: 293-312.

SAUNDERS G.W., 2008 — A DNA barcode examination of the red algal family Dumontiaceae in Canadian waters reveals substantial cryptic species diversity. 1. The foliose *Dilsea-Neodilsea* complex and *Weeksia*. *Botany* 86: 773-789.

SAUNDERS G.W., 2009 — Routine DNA barcoding of Canadian Gracilariales (Rhodophyta) reveals the invasive species *Gracilaria vermiculophylla* in British Columbia. *Molecular ecology resources* 9: 140-150.

SEARS J.R., 2002 — *NEAS Keys to the Benthic Marine Algae of the Northeastern Coast of North America from Long Island Sound to the Strait of Belle Isle.* Dartmouth, University of Massachusetts, 161 p.

SEIFERT K.A., 2009 — Progress towards DNA barcoding of fungi. *Molecular ecology resources* 9: 83-89.

SEIFERT K.A., SAMSON R.A., DEWAARD J.R., HOUBRAKEN J., LEVESQUE C.A., MONCALVO J.M., LOUIS-SEIZE G. & HEBERT P.D.N., 2007 — Prospects for fungus identification using C01 DNA barcodes, with *Penicillium* as a test case. *Proceedings of the National Academy of Sciences U.S.A.* 104: 3901-3906.

SHERWOOD A.R. & PRESTING G.G., 2007 — Universal primers amplify a 23S rDNA plastid marker in eukaryotic algae and cyanobacteria. *Journal of phycology* 43: 605-608.

SHERWOOD A.R., VIS M.L., ENTWISLE T.J., NECCHI O. & PRESTING G.G., 2008 — Contrasting intra versus interspecies DNA sequence variation for representatives of the Batrachospermales (Rhodophyta): Insights from a DNA barcoding approach. *Phycological research* 56: 269-279.

SUBBOTIN S.A., VOVLAS N., CROZZOLI R., STURHAN D., LAMBERTI F., MOENS M. & BALDWIN J.G., 2005 — Phylogeny of Criconematina Siddiqi, 1980 (Nematoda: Tylenchida) based on morphology and D2-D3 expansion segments of the 28S-rRNA gene sequences with application of a secondary structure model. *Nematology* 7: 927-944.

SWOFFORD D.L., 2002 — PAUP* v4.0b10. Sinauer Associates, Sunderland, Massachusetts.

VERBRUGGEN H., TYBERGHEIN L., PAULY K., VLAEMINCK C., VAN NIEUWENHUYZE K., KOOSITRA W. H. C. F., LELIAERT F. & DE CLERCK O., 2009 — Macroecology meets macroevolution: evolutionary niche dynamics in the seaweed *Halimeda*. *Global ecology and biogeography* 18: 393-405.

WALKER R.H., BRODIE J., RUSSELL S., IRVINE L.M. & ORFANIDIS S., 2009 — Biodiversity of coralline algae in the northeastern Atlantic including *Corallina caespitosa* sp. nov. (Corallinoideae, Rhodophyta). *Journal of phycology* 45: 287-297.

WATANABE K.I., EHARA M., INAGAKI Y. & OHAMA T., 1998 — Distinctive origins of group I introns found in the COXI genes of three green algae. *Gene* 213: 1-7.

WHITE T.J., BRUNS T., LEE S. & TAYLOR J., 1990 — Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In*: Innis M.A., Gelfan D.H., Sninsky J.J. & White T. J. (Eds) *PCR Protocols: A guide to methods and applications.* Academic Press, Inc., San Diego, California, pp. 315-322.

ZUCCARELLO G.., PRICE N., VERBRUGGEN H. & LELIAERT F., 2009 — Analysis of a plastid multigene data set and the phylogenetic position of the marind macroalga *Caulerpa filiformis* (Chlorophyta). *Journal of phycology* 45: 1206-1212.