Biodivers Conserv (2014) 23:1411–1424
DOI 10.1007/s10531-014-0672-z

ORIGINAL PAPER

Molecular taxonomic tools provide more accurate estimates of species richness at less cost than traditional morphology-based taxonomic practices in a vegetation survey

Ken A. Thompson · Steven G. Newmaster

Received: 17 January 2014 / Revised: 1 March 2014 / Accepted: 7 March 2014 /
Published online: 19 March 2014
© Springer Science+Business Media Dordrecht 2014

Abstract Vegetation surveys are conducted to obtain a catalogue of the plant species that occupy an area of interest, and are used to inform the decisions of policymakers about conservation, development, and remediation efforts. Currently, vegetation surveys rely on traditional morphology-based taxonomic practices to identify collected specimens. By implementing recent advances in molecular taxonomy, it may be possible to improve upon these methods and reduce the associated costs. In this study, we used both morphological and molecular taxonomic methods to sample 337 forest vegetation plots in northeastern Ontario, Canada. DNA barcoding—a molecular taxonomic tool—was used to identify specimens collected in the molecular taxonomic survey. The molecular taxonomic survey identified a mean of 12.4 species per plot and 202 species in total, whereas the morphological taxonomic survey identified a mean of 9.8 species per plot and 142 species in total; both surveys provided identical estimates of community similarity. The morphological taxonomic survey was 37 % more expensive than molecular taxonomic survey, owing largely to the increased time required in the field to collect specimens that flowered at different times. Our results indicate that molecular taxonomic tools are more cost-effective than traditional morphology-based taxonomic practices for species identification in vegetation surveys. Taxonomy underpins all conservation, and the implementation of molecular taxonomic tools for vegetation surveys has promise to lessen the consequences of the taxonomic impediment and increase the effectiveness of conservation efforts.

Keywords Community inventory · DNA barcoding · DNA taxonomy · Ecological survey · Environmental impact assessment · Plant identification

Communicated by Dirk Sven Schmeller.

K. A. Thompson (✉) · S. G. Newmaster
Department of Integrative Biology, University of Guelph, Guelph, ON N1G 2W1, Canada
e-mail: kthomp1063@gmail.com

Industry and academia have a need for survey tools to assess biological diversity that are underpinned by species identification (Schemske et al. 1994; Claridge et al. 1997; Mace 2004). Ecological surveys are conducted to monitor the status of natural resources in a region, and are a common requirement of environmental impact assessments (Lawrence 2003). The results of ecological surveys are used to inform the decisions of policymakers about conservation, development, and remediation efforts (Beanlands and Duinker 1983), and accurate surveys are important because they allow for better decisions to be made (Resh and Unzicker 1975). One type of ecological survey is the vegetation survey, which catalogues the plant species present in an area and provides information about species richness and community similarity (Curtis 1959; Austin and Heyligers 1989). Currently, vegetation surveys rely on traditional morphology-based taxonomic practices to identify collected specimens. Because of recent advances in DNA sequencing technology (Schuster 2008) and DNA-based plant identification (Newmaster et al. 2006; Fazekas et al. 2012), molecular taxonomic tools have promise to improve upon traditional methods of species identification in vegetation surveys (de Mattia et al. 2012).

Conducting a vegetation survey using traditional morphology-based taxonomic practices for species identification can be arduous (Margules et al. 1994). Technicians require extensive botanical training, and must collect hundreds of obscure specimens for identification at the home institution. Moreover, there is a growing shortage of skilled taxonomists—a trend known as the 'taxonomic impediment' (de Carvalho et al. 2005); perhaps as a result, misidentification and poor resolution of taxa to species-level abound in vegetation surveys (Scott and Hallam 2003). To determine the identity of unknown species, botanists employ diagnostic taxonomic keys that typically compare floral characters between taxa. Because plant species flower at different times throughout the growing season (Campbell 1959), surveyors must revisit a site several times before all plants can be identified (de Mattia et al. 2012). Juveniles or plants that do not flower annually may take years before they can be identified. Although traditional morphology-based taxonomy is critical for describing new species and resolving difficult taxonomic issues, it may not be the best tool for routine species identification in vegetation surveys.

DNA-based molecular taxonomic tools have promise to improve the accuracy and efficiency of routine vegetation surveys (de Mattia et al. 2012; Kuzmina et al. 2012). Firstly, molecular taxonomic tools only require a small section of tissue to identify plants (Gonzalez et al. 2009), and field technicians could save time and space by not having to collect large specimens. Secondly, the use of a referenced DNA database to identify specimens provides archived molecular evidence for the identification (Ratnasingham and Hebert 2007); many plant surveys include on-site identifications with no collections or vouchers as supportive evidence. Thirdly, botanical knowledge is not required to identify plants using molecular taxonomy if there is a comprehensive reference database. There is evidence that amateur technicians can sample diversity in the field as well as expert taxonomists (McCune et al. 1997), so a lack of trained botanists should not inhibit a thorough survey if molecular techniques are used. Lastly, since flowers are not required to identify plants using molecular taxonomic tools, sampling can be completed in a single trip, rather than at several times throughout the growing season. An effective sampling design is important regardless of the primary method of species identification (Austin and Heyligers 1989).

DNA barcoding is a taxonomic method that uses short, standardized, DNA gene marker regions that are conserved within species but vary among species to generate 'barcodes' which are used to determine the identity of organisms (Hebert et al. 2003). Published plant barcoding research supports a multi-region approach to generating DNA barcodes (Kress

et al. 2005; Newmaster et al. 2006; Kress and Erickson 2007; CBOL Plant Working Group 2009). Plant barcodes are constructed from two 'tiered' gene regions; an easily amplified and aligned region is used for the first tier, and acts as a scaffold on which data from a more variable second-tier region are placed (Newmaster et al. 2006). The chloroplast gene *rbcL* is used as the standard first-tier marker because of its universality and demonstrated success for differentiating congeneric plant species (Chase et al. 2006; Newmaster et al. 2006). In 2009 the CBOL Plant Working Group suggested that the standard first tier barcoding marker for plants should be *matK*. However, many published studies have concluded that *matK* is not suitable for this purpose because it is very difficult and expensive to sequence, has relatively low species resolution, and performs inconsistently across land plant taxa (Fazekas et al. 2008; Newmaster et al. 2008; Chen et al. 2010; Kress et al. 2010; von Cräutlein et al. 2011; Li et al. 2012a, 2012b; Sandionigi et al. 2012; Zhang et al. 2012). Today, the highly-variable nuclear *ITS2* region is widely used as the standard second tier marker because of its ability to distinguish between closely related species, and its ease of amplification and sequencing (Chen et al. 2010; Yao et al. 2010; Li et al. 2011). The high success rate of *ITS2* has been verified in several recent applications (Kool et al. 2012; Kuzmina et al. 2012; Liu et al. 2012; García-Robledo et al. 2013; Ghahramanzadeh et al. 2013; Newmaster et al. 2013a, b; Laiou et al. 2013; de Boer et al. 2014).

Before molecular taxonomic tools can be implemented as the primary method of species identification in routine vegetation surveys, two questions must be answered: (1) Do molecular taxonomic tools provide results that are comparable to—or better than—those of surveys conducted with traditional morphology-based taxonomic methods? (2) Is the use of a molecular taxonomic tool as the primary method of species identification more cost-effective than using traditional morphology-based taxonomic methods? Here, we sample 337 vegetation plots in northeastern Ontario, Canada, using both morphological and molecular taxonomic methods. We compare estimates of species richness and community similarity between surveys, as well as the associated costs to assess the utility of molecular taxonomic tools for species identification in vegetation surveys.

Methods

Sampling methods

Sampling took place from June to September 2010, in the boreal mixed-wood forests near Timmins, Ontario (48°N, 81°W). A total of 337 plots (50 m^2) were sampled with both molecular and traditional morphology-based taxonomic methods (hereafter the *molecular survey* and *morphological survey*, respectively) at the NEBIE Plot Network (Bell et al. 2008). All vascular plants, bryophytes, and lichens were sampled. For each plot, species richness was recorded as the number of unique species-level identifications. We recorded the genus or family for all specimens that could not be resolved to species. Environmental variables associated with each plot (slope, aspect, landform, effective texture, moisture regime, soil depth, organic depth, and canopy closure) were recorded for 100 m^2 plots aligned with the main plot's center.

A team of four trained botanists surveyed all plots and identified species in the morphological survey. Common species were identified in the field, while obscure species were pressed and identified at the University of Guelph Herbarium in the Center for Biodiversity Genomics (CBG) at the Biodiversity Institute of Ontario. Plots were sampled repeatedly over 4 months because some plants could not be identified without flowers.

A team of three undergraduate field technicians and one botanist surveyed all plots and collected specimens for identification with DNA barcoding in the molecular survey. Undergraduates received basic training in recognizing patterns in vegetation and in distinguishing species. Small samples (10 cm^2) of leaf tissue were collected from each perceived species and stored in vials containing silica gel.

Costs for both surveys were calculated based on the actual value of personnel, travel and lab expenses. This included the time and rate-of-pay for technicians, costs associated with traveling to the site and completing data collection on site, as well as food and campsite rental. The associated lab costs include herbarium materials and those associated with the botanical DNA barcode research lab at the CBG.

DNA barcoding

Leaf samples collected in the molecular survey were identified using DNA barcoding. Leaf samples (0.5 cm^2) were processed at the CBG, and sent to the Canadian Center for DNA Barcoding (CCDB) for automated high through-put barcoding. This facility has processed over 2.86 million barcodes for over 200,000 species (flora and fauna) and is an international hub for DNA barcoding research that has received over $100 million in support of DNA barcoding. The DNA barcode library used for this project was constructed as part of a campaign to barcode the Flora of Ontario (Newmaster and Ragupathy 2013). The library is deposited as a public database entitled "University of Guelph, OAC Herbarium" in the Barcode of Life Data System (http://www.boldsystems.org). The cost for barcoding a specimen ranges from $15 to $40 CAD, which is dependent on the taxonomic group being surveyed; use of the on-line database (BOLD) is included in this cost. All specimens were vouchered, and barcodes were uploaded to BOLD by the University of Guelph, OAC Herbarium.

DNA-barcodes were constructed using the *rbcL* and *ITS2* marker regions. DNA extraction followed the standard protocol for plant DNA barcoding (Ivanova et al. 2005, 2008, 2011; Kuzmina and Ivanova 2011a); final DNA concentration was 20–40 ng/μL. Markers were amplified with Platinum® *Taq* DNA polymerase (Invitrogen, Carlsbad, CA), and pre-made frozen plates (Ivanova and Grainger 2006; Fazekas et al. 2012). We obtained strong amplification of the gene regions with low concentrations of primers (0.1 μM), dNTPs (0.05 mM), and Taq polymerase (0.024 U/μL). Amplicons were diluted 5–10× and sequenced without PCR purification. The primers rbcLa-F (Levin et al. 2003) and rbcLa-R (Kress et al. 2009) were used for *rbcL* analysis, while ITS-S2F (Chen et al. 2010) was used for ITS2 analysis (primer sequences for *rbcL* and *ITS2* are available on the CCDB Protocols website (Kuzmina and Ivanova 2011b)). Products were analyzed using an ABI3730xl capillary sequencer (Applied Biosystems, Foster City, CA). Chromatographic traces were aligned and contigs were generated from codons with Codoncode Aligner v3.0 (CodonCode, Centerville, MA).

We used DNA barcodes to determine species membership. We used BLAST to compare marker region sequences against a standard reference barcode library with a minimum BLAST cut-off of 97 % identity for top matches (Newmaster et al. 2013a). Results were verified using neighbour-joining tree analysis where the branches of marker regions from leaf samples were compared to sequences of reference species. Species were considered to be taxonomically resolved if their members showed diagnostic differences from other taxa and formed a clade.

Data analysis

We compared estimates of plot species richness generated in both surveys. We conducted a paired *t* test in R version 3.0.2 (R Core Team 2013) to determine if estimates of mean plot species richness were significantly different between surveys. We also retrieved conservation status ranks (hereafter *S-ranks*) from the Flora of Ontario Integrated Botanical System (Newmaster and Ragupathy 2013) and compared them across surveys.

We used multivariate analysis to explore variation in community similarity among plots. We produced presence-absence matrices for all plots separately for both the molecular and morphological taxonomic surveys. We did not restrict analysis to only species-level identifications because of an a priori assumption that identifications to genus are important for community similarity estimates. We used CANOCO 5.0 (Ter Braak and Smilauer 2012) to explore variation among taxa within plots constrained by the eight measured environmental variables. To identify the length of the ordination axis (i.e. the extent of variation in the axis scores), we conducted a canonical correspondence analysis (CCA; Ter Braak 1986); the length of the gradient (4 SD) justified the use of a CCA to characterize variation among species. We considered species turnover between plots in the CCA using Hill's scaling (Hill and Gauch 1980); plots that are 2 SD apart on the ordination axis share less than one-third of their species, and plots that are 4 SD apart have no species in common (Jongman et al. 1995). We used plot scores from the CCA to position the plot on a complex gradient, which was standardized to a within-plot variance of 1 (Ter Braak 1986). To compare axis scores across surveys, we calculated Pearson product-moment correlation coefficients using the 'Hmisc' package for R (Harrell and Dupont 2012).

Results

Species richness

Across both surveys, we identified a total of 7,948 specimens (Table 1). Among these, there were 202 unique species from 75 families. 142 unique species were identified in the morphological survey, while 23 genera had members that could not be identified to species; twenty specimens could only be identified to family. Approximately 12.5 % of specimens collected in the morphological survey could not be identified to species—most (57 %) of these were from taxonomically difficult graminoid, pteridophyte, bryophyte, or lichen taxa. DNA barcodes were produced for all specimens collected in the molecular survey. All 142 species identified in the morphological survey were accounted for by the molecular survey, as well as 60 additional species unique to the molecular survey. Of these 60 species, 11 were from genera not detected by the morphological survey and 49 were from genera that could not be resolved using morphology-based taxonomy.

The two surveys differed in their estimates of species richness at the plot level (Table 1). A paired *t* test revealed that the molecular survey returned significantly higher estimates of plot species richness compared to the morphological survey ($t_{336} = 26.1$; $p < 0.001$; 95 % CI 2.40–2.79; Fig. 1). The difference in mean estimated plot richness between surveys was still significantly higher in the molecular survey when all unique taxa (i.e. taxa identified to genus or family only) were considered ($t_{336} = 18.8$; $p < 0.001$; 95 % CI 1.07–1.32).

Most (89 %) species identified in the molecular survey that were not in the morphological survey were of little or no conservation concern (S-rank S4 or S5). However, there

Table 1 Summary of the morphological and molecular vegetation surveys

	Morphological	Molecular
Specimens collected	3,772	4,176
Unique species detected	142	202
Identification rate		
To species	0.87	1
To genus	0.99	1
Plot richness (species)		
Mean	9.8	12.4
Median	9	12
Mode	8	11
Plot richness (taxa)		
Mean	11.2	12.4
Median	11	12
Mode	10	11
Associated costs		
Field survey	$86, 800	$19, 600
Specimen ID	$11, 200	$42, 100
Total	$98, 000	$61, 700

were two critically imperiled exotic rare invasive species (S-rank SE1), *Actaea pachypoda* (Ranunculaceae) and *Hieracium praealtum* (Asteraceae), that were unaccounted for by the morphological survey, as well as an imperiled (S-rank S2) moss, *Callicladium haldanianum* (Hypnaceae), and a vulnerable (S-rank S3) grass, *Poa saltuensis* (Poaceae).

Community similarity

Relationships between species, plots, and environmental variables were interpretable in the canonical correspondence analysis (Fig. 2; Table 2). The overall variance was 4.0 SD for both surveys, indicating that species composition differed considerably among plots. The seven environmental variables explained 81.9 % and 79.4 % of the variance in species composition at the plot level for the morphological and molecular surveys, respectively (Table 3). Canopy closure and organic depth had the strongest effect on plot inter-set correlation coefficients. Plots with closed canopies were in the lower left quadrat, and plots with deep organic layers, high soil moisture, and deeper soils were in the lower right quadrat. Plots with steeper slopes were in the top left quadrat. Axis scores were highly correlated across surveys (both axes Pearson's $r = 1.00$; $n = 337$; $p < 0.001$), indicating that estimates of community similarity did not differ between surveys.

Associated costs

The two surveys differed considerably in associated costs (Table 1). Collecting specimens in the field cost $86,800 and $19,600 CAD for the morphological and molecular surveys respectively. This difference in total cost was largely because the morphological survey took 4 months in the field to collect all specimens while they were in flower, while the molecular survey required only 1 month to sufficiently sample all plots. Wage differences

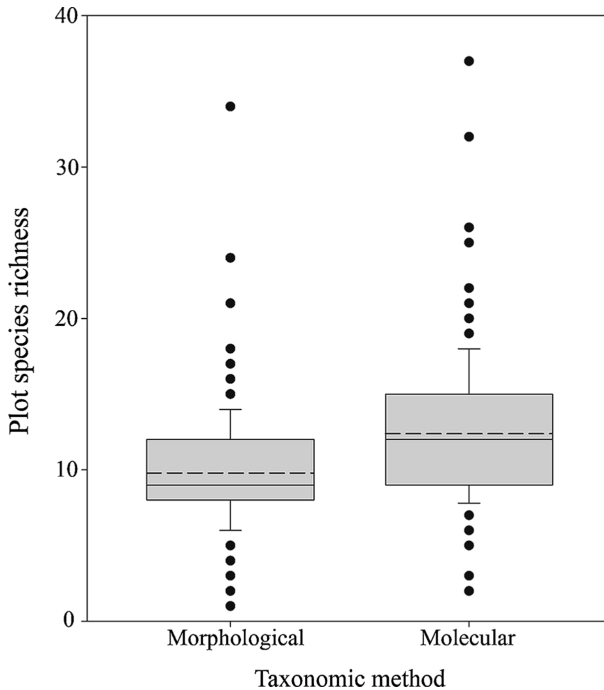

Fig. 1 Box plot comparing the number of taxa identified to species in each of 337 plots for both the morphological and molecular surveys. *Boxes* represent the interquartile range (IQR), and *bars* include the lowest/highest datum within 1.5 IQR of the upper and lower quartiles. The mean number of species per plot is indicated with a *dashed line*, and the median with a *solid line*

also added to the difference in cost—botanists received an hourly wage of $20 CAD while undergraduate technicians received an hourly wage of $15 CAD. Since the morphological survey team consisted of four botanists, it cost more per hour than the molecular survey, which had a team of three undergraduates and one botanist. The cost of identifying specimens collected in the field was higher for the molecular survey—DNA barcoding cost $42,100 whereas identifying specimens in the herbarium using traditional morphological methods cost $11,200. This difference in cost was because processing and barcoding fees at the CBG and CCDB were higher than the cost of hiring a botanist to identify collected specimens with traditional morphology-based taxonomy.

In sum, field collection and specimen identification cost $98,000 and $61,700 CAD for the morphological and molecular surveys, respectively. Costs per plot were $292 CAD and $184 for the morphological and molecular surveys, respectively. If only species-level identifications are considered, then it cost approximately $29.8 to identify each species in a plot for the morphological survey, and $14.9 for the molecular survey.

Discussion

The results of this study demonstrate that vegetation surveys conducted using molecular taxonomic tools as the primary method of species identification provide better species

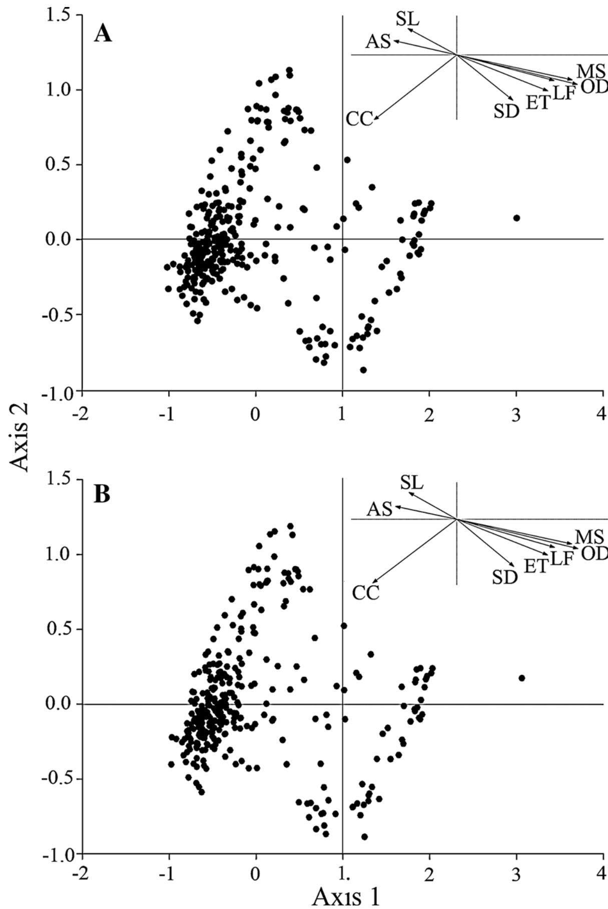

Fig. 2 Canonical correspondence analysis ordination of 337 sample plots for the (**A**) morphological and (**B**) molecular taxonomic surveys constrained by 8 environmental variables. Matrices were produced for all unique taxa, rather than only species-level identifications. Bi-plot *arrows* represent the relative canonical correlations for environmental variables, all of which explain a significant ($p < 0.05$) portion of the variation among plots on at least one axis. Plot scores on both axes are highly correlated across surveys ($r = 1.00$, $p < 0.001$)

richness estimates at less cost than surveys conducted using morphology-based taxonomic methods. Gonzalez et al. (2009) made the first attempt to use DNA barcoding to expedite vegetation surveys, but found that it had low (<70 %) rates of species identification; because of this, the authors concluded that the tool was not yet suitable for routine species identification in vegetation surveys. With recent advancements in plant DNA barcoding, de Mattia et al. (2012) demonstrated that vegetation surveys conducted using molecular taxonomic tools for species identification are more efficient than those that use morphology-based taxonomy. The present study corroborates the findings of de Mattia et al. (2012), and builds upon their result by demonstrating that vegetation surveys conducted using molecular taxonomic tools provide more accurate estimates of species richness at a much lower cost than surveys conducted using morphology-based taxonomy for species identification. Successful implementation of molecular taxonomy requires a comprehensive

Table 2 Statistics in canonical correspondence analysis of 337 plots sampled using both morphological and molecular taxonomic methods constrained by 8 environmental variables

Environmental variable	Inter-set correlation		*t* value		*p* value	
	Axis 1	Axis 2	Axis 1	Axis 2	Axis 1	Axis 2
Morphological						
Slope	−0.3310	0.1980	0.9853	3.2477	0.325	**<0.001**
Aspect	−0.4258	0.1066	−3.0066	−0.9889	**0.003**	0.323
Landform	0.6608	−0.1910	−0.7209	−2.0878	0.471	**0.038**
Effective texture	0.6183	−0.2694	−0.2117	−3.0765	0.832	**0.002**
Moisture regime	0.7844	−0.1862	6.9772	0.8326	**<0.001**	0.406
Soil depth	0.3804	−0.3390	−3.3885	−4.3674	**<0.001**	**<0.001**
Organic depth	0.8187	−0.2199	9.6367	−1.0749	**<0.001**	0.283
Canopy closure	−0.5615	−0.4872	−9.9674	−13.9369	**<0.001**	**<0.001**
Molecular						
Slope	−0.3251	0.2065	1.1818	3.6678	0.238	**<0.001**
Aspect	−0.4176	0.0983	−2.6420	−1.4726	**0.009**	0.142
Landform	0.6681	−0.2157	−0.9375	−3.0563	0.349	**0.002**
Effective texture	0.6204	−0.2747	−0.0954	−3.2815	0.924	**<0.001**
Moisture regime	0.7857	−0.1899	6.6480	1.1284	**<0.001**	0.260
Soil depth	0.3896	−0.3623	−2.9574	−5.2267	**0.003**	**<0.001**
Organic depth	0.8276	−0.2255	10.3220	−0.8019	**<0.001**	0.423
Canopy closure	−0.5732	−0.4891	−10.3740	−14.7678	**<0.001**	**<0.001**

Significant *p* values ($\alpha = 0.05$) are highlighted in bold

Table 3 Summary of canonical correspondence analysis for 337 plots sampled using both morphological and molecular taxonomic techniques constrained by 8 environmental variables

Summary variable	Axis			
	1	2	3	4
Morphological				
Eigenvalue	0.387	0.137	0.129	0.072
Environmental variable correlations	0.894	0.698	0.664	0.57
Cumulative % of variance explained	43.7	59.2	73.8	81.9
Molecular				
Eigenvalue	0.386	0.142	0.125	0.075
Environmental variable correlations	0.899	0.724	0.686	0.578
Cumulative % of variance explained	42.1	57.6	71.2	79.4

DNA reference database, and recent evidence indicates that these databases can be assembled relatively inexpensively if local herbaria are well provisioned (Kuzmina et al. 2012).

The molecular survey consistently reported higher plot species richness than the morphological survey. Some species were likely detected only in the molecular survey because

of misidentifications in the morphological survey. For example, *Waldsteinia fragarioides* (Rosaceae) closely resembles the common strawberry (*Fragaria* spp.; Hill 2003), and may have been identified only as such in the morphological survey. This may reflect a gap in the training of our field botanists, but is important nonetheless because human error is common in ecological surveys (Lepš and Hadincová 1992; Scott and Hallam 2003; Archaux 2009; Ensing et al. 2012). Some species were not identified in the morphological survey because of problems inherent to morphology-based taxonomy. For example, *Salix* spp. (Salicaceae) had finished flowering before the morphological survey began and could only be identified to genus.

The high correlation of CCA axes across surveys (Fig. 2) indicates that information about community similarity did not differ between taxonomic surveys. This suggests that there are diminishing returns in information about community similarity as species-level information increases. Because molecular taxonomic tools did not change estimates of community similarity, prior work making such estimates using morphology-based taxonomic methods (e.g. Curtis 1959) will be comparable to new estimates made using molecular taxonomic tools. This backward compatibility suggests that unanimous implementation of molecular taxonomic tools will not disrupt the continuity of long-term community-level data. This result is especially important if conservation decisions are based on community-similarity measures (Su et al. 2004).

As skilled taxonomists become scarcer, the need for accurate and inexpensive species identification in field surveys will increase in importance. Molecular taxonomic tools, paired with online publicly accessible DNA sequence databases will make it easier for scientists to access taxonomic information (Ratnasingham and Hebert 2007). These databases provide a way around the taxonomic impediment because skilled taxonomists—who are normally needed to identify difficult specimens—are not required if there is a match in a DNA reference library (Tautz et al. 2003). Recently, DNA barcoding has been used by non-taxonomists to efficiently identify plant roots (Kesanakurti et al. 2011) and freshwater benthic invertebrates (Hajibabaei et al. 2011) collected in the field, demonstrating its utility for the routine identification of difficult specimens. Taxonomists will always be needed to describe new species (alpha taxonomy) and to provide species-level identifications for building referenced DNA barcode databases. Thus, it is important to note that we are not advocating a ubiquitous replacement of morphology-based taxonomy. Rather, we suggest selectively implementing molecular taxonomic tools in areas where evidence indicates that morphology-based taxonomy is unnecessary.

An inability to generate and access taxonomic information will act as a bottleneck that slows the progress of conservation research (Claridge et al. 1997; Dubois 2003; Tautz et al. 2003). This is because the first step in the cycle of conservation is observing and monitoring species or populations to demonstrate that conservation efforts are required, and conservation efforts will be impeded if taxa that require conservation cannot be identified (Mace 2004). Species richness values are—perhaps too often (Su et al. 2004)—used for allocating conservation funding (Balmford et al. 2000; Gotelli and Colwell 2001). In our study, the morphological taxonomic survey provided estimates of plot species richness that were lower than those of the molecular taxonomic survey. If DNA-based taxonomy were implemented for routine vegetation surveys, then the more accurate estimates of species richness that result will ensure appropriate allocation of conservation funds. Since molecular taxonomic tools reduce costs associated with conducting vegetation surveys, their implementation will allow for a higher frequency of more accurate surveys to be conducted on a fixed budget. This has promise to better the monitoring of rare species and enable earlier detection of invasive species—both of which have important implications for

conservation (Prendergast et al. 1993; Schemske et al. 1994; Gurevitch and Padilla 2004). Other species of interest—for conservation purposes or otherwise—could be effectively located and monitored with GPS co-ordinates of accurately identified populations on a referenced DNA database. Today, identifications made in vegetation surveys are typically not anchored in online databases, and DNA databases ensure that this data remains freely available into the future. This could have applications for reducing sampling bias in the field and improving ecological niche modeling (Costa et al. 2010), and may also be useful for biomonitoring (Hajibabaei et al. 2011).

Our study had several limitations. Firstly, the study site was distant (~750 km drive) from the home institution. This added considerably to costs because the survey teams had to make several trips to the site, and had to pay for food and lodging. If costs associated with travel, lodging and food are removed, costs become $66,000 and $53,700 for the morphological and molecular taxonomic surveys, respectively; the difference in cost is still higher for the morphological method because of the wage differences and the time required for sampling the plots throughout the growing season. Secondly, this study was conducted out of a facility that possessed all equipment necessary to conduct a molecular taxonomic survey, although we were charged the standard rate for barcoding at CCDB. Because of this, costs that would normally be associated with purchasing DNA sequencing equipment or shipping samples for barcoding did not apply. Lastly, the study was conducted in the boreal forest of Ontario, Canada, where there is a publically available comprehensive and fully referenced DNA barcode database. Without such a database in place, future molecular taxonomic surveys will not share the degree of success that we had.

We conclude that greater effort should be made to implement molecular taxonomic tools for vegetation surveys because they allow for more accurate estimates of species richness to be made at less cost than when traditional morphology-based methods of species identification are used. Molecular markers can identify most major groups of organisms (Vences et al. 2005; Begerow et al. 2010; Hajibabaei et al. 2011), and may have promise for complementing or replacing traditional species-identification practices in non-botanical field surveys. In our rapidly changing world where at least 15 % of species face extinction (Thomas et al. 2004), cost-effective conservation efforts are of great importance (Murdoch et al. 2007). Molecular taxonomic tools will continue to increase in performance (Frézal and Leblois 2008) and decrease in price (Hebert and Gregory 2005; Schuster 2008), and their effective implementation has promise to improve conservation efforts globally.

Acknowledgments The authors thank W. Bell and the Ontario Forest Research Institute for assistance with fieldwork, N. Webster for assistance with morphology-based specimen identification, and S. Ragupathy for assistance with DNA barcoding. SGN was supported with Grants from NSERC CRD, OMNR, Forest Ecosystem Co-op, and the Forest Research Partnership.

References

Archaux F (2009) Could we obtain better estimates of plot species richness from multiple-observer plant censuses? J Veg Sci 20:603–611

Austin MP, Heyligers PC (1989) Vegetation survey design for conservation: gradsect sampling of forests in North-eastern New South Wales. Biol Conserv 50:13–32

Balmford A, Gaston KJ, Rodrigues ASL, James A (2000) Integrating costs of conservation into international priority setting. Conserv Biol 14:597–605

Beanlands GE, Duinker PN (1983) An ecological framework for environmental impact assessment in Canada. Institute for Resource and Environmental Studies, Dalhousie

Begerow D, Nilsson H, Unterseher M, Maier W (2010) Current state and perspectives of fungal DNA barcoding and rapid identification procedures. Appl Microbiol Biotechnol 87:99–108
Bell F, Parton J, Stocker N, Joyce D (2008) Developing a silvicultural framework and definitions for use in forest management planning and practice. Forest Chron 84:678–693
Campbell JB (1959) Flowering sequence of a local flora. J Range Manag 12:127–132
CBOL Plant Working Group (2009) A DNA barcode for land plants. Proc Natl Acad Sci USA 106:12794–12797
Chase M, Fay M, Devey D et al (2006) Multigene analyses of monocot relationships: a summary. Aliso 22:62–74
Chen S, Yao H, Han J et al (2010) Validation of the *ITS2* region as a novel DNA barcode for identifying medicinal plant species. PLoS ONE 5:e8613
Claridge MF, Dawah HA, Wilson MR (1997) Species: The units of biodiversity. Chapman & Hall, London
Core Team R (2013) R: a language and environment for statistical computing. Austria, Vienna
Costa GC, Nogueira C, Machado RB, Colli GR (2010) Sampling bias and the use of ecological niche modeling in conservation planning: a field evaluation in a biodiversity hotspot. Biodivers Conserv 19:883–899
Curtis JT (1959) The vegetation of Wisconsin: an ordination of plant communities. University of Wisconsin Press, Madison
de Boer HJ, Ouarghidi A, Martin G et al (2014) DNA barcoding reveals limited accuracy of identifications based on folk taxonomy. PLoS ONE 9:e84291
de Carvalho MR, Bockmann FA, Amorim DS et al (2005) Revisiting the taxonomic impediment. Science 307:353
de Mattia F, Gentili R, Bruni I et al (2012) A multi-marker DNA barcoding approach to save time and resources in vegetation surveys. Bot J Linn Soc 169:518–529
Dubois A (2003) The relationships between taxonomy and conservation biology in the century of extinctions. C R Biol 326:9–21
Ensing DJ, Moffat CE, Pither J (2012) Taxonomic identification errors generate misleading ecological niche model predictions of an invasive hawkweed. Botany 91:137–147
Fazekas AJ, Burgess KS, Kesanakurti PR et al (2008) Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. PLoS ONE 3:e2802
Fazekas AJ, Kuzmina ML, Newmaster SG, Hollingsworth PM (2012) DNA barcoding methods for land plants. In: Kress WJ, Erickson (ed) DNA barcodes: methods and protocols. Methods Mol Bol 853:223–252
Frézal L, Leblois R (2008) Four years of DNA barcoding: current advances and prospects. Infect Genet Evol 8:727–736
García-Robledo C, Erickson DL, Staines CL et al (2013) Tropical plant-herbivore networks: reconstructing species interactions using DNA barcodes. PLoS ONE 8:e52967
Ghahramanzadeh R, Esselink G, Kodde LP et al (2013) Efficient distinction of invasive aquatic plant species from non-invasive related species using DNA barcoding. Mol Ecol Resour 13:21–31
Gonzalez MA, Baraloto C, Engel J et al (2009) Identification of Amazonian trees with DNA barcodes. PLoS ONE 4:e7483
Gotelli NJ, Colwell RK (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecol Lett 4:379–391
Gurevitch J, Padilla DK (2004) Are invasive species a major cause of extinctions? Trends Ecol Evol 19:470–474
Hajibabaei M, Shokralla S, Zhou X et al (2011) Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. PLoS ONE 6:e17497
Harrell Jr FE, Dupont MC (2012) Hmisc: Harrell miscellaneous. R package version 3.10-1. [http://CRAN.R-project.org/package=Hmisc]
Hebert PDN, Gregory TR (2005) The promise of DNA barcoding for taxonomy. Syst Biol 54:852–859
Hebert PDN, Cywinska A, Ball SL (2003) Biological identifications through DNA barcodes. Proc R Soc London Series B Biol Sci 270:313–321
Hill SR (2003) Conservation assessment for Barren strawberry (*Waldsteinia fragarioides* ssp. *fragarioides*). United States Forest Service, Washington, DC
Hill MO, Gauch HG (1980) Detrended correspondence analysis: an improved ordination technique. Vegetatio 42:47–58
Ivanova NV, Grainger C (2006) Pre-made frozen PCR and sequencing plates. [www.dnabarcoding.ca/CCDB_DOCS/CCDB_Advances_Methods_Release_No4_Dec1st_2006.pdf]
Ivanova NV, DeWaard JR, Hajibabaei M, Hebert PDN (2005) Protocols for high volume DNA barcoding; [http://www.dnabarcoding.ca/]

Ivanova NV, Fazekas AJ, Hebert PDN (2008) Semi-automated, membrane-based protocol for DNA isolation from plants. Plant Mol Biol Rep 26:186–198

Ivanova NV, Kuzmina ML, Fazekas AJ (2011) CCDB Protocols. Manual protocol employing centrifugation: Glass fiber plate DNA extraction protocol for plants, fungi, echinoderms, and mollusks. [http://www.ccdb.ca/CCDB_DOCS/CCDB_DNA_Extraction-Plants.pdf]

Jongman RHG, ter Braak CJF, Van Tongeren OFR (1995) Data analysis in community and landscape ecology. Cambridge University Press, Cambridge

Kesanakurti PR, Fazekas AJ, Burgess KS et al (2011) Spatial patterns of plant diversity below-ground as revealed by DNA barcoding. Mol Ecol 20:1289–1302

Kool A, de Boer HJ, Krüger Å et al (2012) Molecular identification of commercialized medicinal plants in southern Morocco. PLoS ONE 7:e39459

Kress WJ, Erickson DL (2007) A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. PLoS ONE 2:e508

Kress WJ, Wurdack KJ, Zimmer EA et al (2005) Use of DNA barcodes to identify flowering plants. Proc Natl Acad Sci USA 102:8369–8374

Kress WJ, Erickson DL, Jones FA et al (2009) Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. Proc Natl Acad Sci USA 106:18621–18626

Kress WJ, Erickson DL, Swenson NG et al (2010) Advances in the use of DNA barcodes to build a community phylogeny for tropical trees in a Puerto Rican forest dynamics plot. PLoS ONE 5:e15409

Kuzmina ML, Ivanova NV (2011a) PCR amplification for plants and fungi. [http://www.ccdb.ca/CCDB_DOCS/CCDB_Amplification-Plants.pdf]

Kuzmina ML, Ivanova NV (2011b) CCDB Protocols. [http://www.ccdb.ca/CCDB_DOCS/CCDB_PrimerSets-Plants.pdf]

Kuzmina ML, Johnson KL, Barron HR, Hebert PDN (2012) Identification of the vascular plants of Churchill, Manitoba, using a DNA barcode library. BMC Ecol 12:25

Laiou A, Mandolini LA, Piredda R et al (2013) DNA barcoding as a complementary tool for conservation and valorisation of forest resources. Zookeys 213:197–213

Lawrence DP (2003) Environmental impact assessment. Wiley & Sons, Hoboken

Lepš J, Hadincová V (1992) How reliable are our vegetation analyses? J Veg Sci 3:119–124

Levin RA, Wagner WL, Hoch PC et al (2003) Family-level relationships of Onagraceae based on chloroplast *rbcL* and *ndhF* data. Am J Bot 90:107–115

Li D-Z, Gao L-M, Li H-T et al (2011) Comparative analysis of a large dataset indicates that internal transcribed spacer (*ITS*) should be incorporated into the core barcode for seed plants. Proc Natl Acad Sci 108:19641–19646

Li CP, Yu ZG, Han GS, Chu KH (2012a) Analyzing multi-locus plant barcoding datasets with a composition vector method based on adjustable weighted distance. PLoS ONE 7:e42154

Li H-Q, Chen J-Y, Wang S, Xiong S-Z (2012b) Evaluation of six candidate DNA barcoding loci in *Ficus* (Moraceae) of China. Mol Ecol Resour 12:783–790

Liu Z, Zeng X, Yang D et al (2012) Identification of medicinal vines by *ITS2* using complementary discrimination methods. J Ethnopharmacol 141:242–249

Mace GM (2004) The role of taxonomy in species conservation. Philos Trans R Soc Lond B Biol Sci 359:711–719

Margules CR, Austin MP, Mollison D, Smith F (1994) Biological models for monitoring species decline: the construction and use of data bases [and discussion]. Philos Trans R Soc London Series B Biol Sci 344:69–75

McCune B, Dey JP, Peck JE et al (1997) Repeatability of community data: species richness versus gradient scores in large-scale lichen studies. Bryologist 100:40–46

Murdoch W, Polasky S, Wilson KA et al (2007) Maximizing return on investment in conservation. Biol Conserv 139:375–388

Newmaster SG, Ragupathy S (2013) Flora Ontario Integrated Botanical Information System (FOIBIS). BIO/OAC Herbarium, Biodiversity Institute of Ontario, Guelph. [http://www.uoguelph.ca/foibis]

Newmaster SG, Fazekas AJ, Ragupathy S (2006) DNA barcoding in land plants: evaluation of *rbcL* in a multigene tiered approach. Botany 341:335–341

Newmaster SG, Fazekas AJ, Steeves RAD, Janovec J (2008) Testing candidate plant barcode regions in the Myristicaceae. Mol Ecol Resour 8:480–490

Newmaster SG, Ragupathy S, Dhivya S et al (2013a) Genomic valorization of the fine scale classification of small millet landraces in southern India. Genome 56:123–127

Newmaster SG, Grguric M, Shanmughanandhan D et al (2013b) DNA barcoding detects contamination and substitution in North American herbal products. BMC Med 11:222

Prendergast JR, Quinn RM, Lawton JH et al (1993) Rare species, the coincidence of diversity hotspots and conservation strategies. Nature 365:335–337

Ratnasingham S, Hebert PDN (2007) BOLD: The barcode of life data system. Mol Ecol Notes 7:355–364. http://www.boldsystems.org

Resh V, Unzicker J (1975) Water quality monitoring and aquatic organisms: the importance of species identification. Water Pollut Control Fed 47:9–19

Sandionigi A, Galimberti A, Labra M et al (2012) Analytical approaches for DNA barcoding data-how to find a way for plants? Plant Biosyst 146:803–813

Schemske DW, Husband BC, Ruckelshaus MH et al (1994) Evaluating approaches to the conservation of rare and endangered plants. Ecology 75:584–606

Schuster SC (2008) Next-generation sequencing transforms today's biology. Nat Methods 5:16–18

Scott WA, Hallam CJ (2003) Assessing species misidentification rates through quality assurance of vegetation monitoring. Plant Ecol 165:101–115

Su JC, Debinski DM, Jakubauskas ME, Kindscher K (2004) Beyond species richness: community similarity as a measure of cross-taxon congruence for coarse-filter conservation. Conserv Biol 18:167–173

Tautz D, Arctander P, Minelli A et al (2003) A plea for DNA taxonomy. Trends Ecol Evol 18:70–74

ter Braak CJF (1986) Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67:1167–1179

Ter Braak, CJF, Smilauer P (2012) CANOCO reference manual and CanoDraw for Windows user's guide: software for canonical community ordination (version 5.0). Microcomputer Power, Ithaca

Thomas CD, Cameron A, Green RE et al (2004) Extinction risk from climate change. Nature 427:145–148

Vences M, Thomas M, van der Meijden A et al (2005) Comparative performance of the 16S rRNA gene in DNA barcoding of amphibians. Front Zool 2:5

von Cräutlein M, Korpelainen H, Pietiläinen M, Rikkinen J (2011) DNA barcoding: a tool for improved taxon identification and detection of species diversity. Biodivers Conserv 20:373–389

Yao H, Song J, Liu C et al (2010) Use of *ITS2* region as the universal DNA barcode for plants and animals. PLoS ONE 5:e13102

Zhang C-Y, Wang F-Y, Yan H-F et al (2012) Testing DNA barcoding in closely related groups of *Lysimachia* L. (Myrsinaceae). Mol Ecol Resour 12:98–108