Ecological and Evolutionary Diversification within the Genus Carex (Cyperaceae): Consequences for Community Assembly in Subarctic Fens

Marcia J. Waterway,1,4 Kyle T. Martins,2 Anna Dabros,3 Alberto Prado,1,2 and Martin J. Lechowicz2

1Department of Plant Science, Macdonald Campus, McGill University, 21,111 Lakeshore Road, Ste-Anne-de-Bellevue, Quebec, Canada H9X 3V9.
2Department of Biology, McGill University, 1205 Avenue Dr. Penfield, Montreal, Quebec, Canada H3A 1B1.
3Department of Natural Resources Canada, Canadian Forest Service, 580 Booth Street, Ottawa, Ontario, Canada K1A 0E4.
4Author for correspondence (marcia.waterway@mcgill.ca)

Abstract—The concept of limiting similarity predicts that closely related taxa are less likely to co-occur than expected by chance. The degree to which the phylogenetic relatedness in plant communities is in accord with limiting similarity has been little tested at the scale where the consequences of adaptive differentiation during speciation should be most evident: the scale of neighboring, congeneric plants within a community. To quantify species co-occurrence patterns in relation to environment, we sampled sedge species, their rooting level relative to the water table, and the water pH in 2,124 0.25 m2 quadrats distributed across 29 subarctic fens in the central Labrador Peninsula. We estimated phylogenetic relationships using four DNA regions (ETS, ITS, matK, trnL-trnF) for all species of Carex (42), Eriophorum (6), and Trichophorum (2) in the region, of which 21, four, and two, respectively, occurred in the sampled fens. We demonstrate that closely related species of Carex are less likely to co-occur than expected by chance using 1) a probabilistic method to test the significance of pairwise co-occurrence patterns of species, and 2) linear mixed modeling to relate these patterns to phylogenetic relationships and ecological tolerances along gradients of substrate pH and rooting level in relation to the water table. The results also indicate that suites of species with significant mutual pairwise co-occurrence belong to distant lineages within the Cariceae-Dulichieae-Scirpeae clade of Cyperaceae and have stabilizing niche differences. We suggest that niche differentiation during the evolution and diversification of a clade of wetland Carex species over the past few million years, especially during the dynamic glacial cycles of the Pleistocene, has resulted in diverse sedge communities that share space and resources in harsh northern peatland habitats.

Keywords—Coexistence theory, community phylogenetics, ecological speciation, ecophylogenetics, limiting similarity, niche differentiation.

Contemporary perspectives on the processes of plant community assembly (Mittelbach and Schemske 2015) are firmly rooted in Gleason’s (1926, 1939) individualistic concept, which emphasizes that the presence and relative abundance of species in a plant community are predicated on individuals of a species having first dispersed to a site and then having successfully colonized and persisted at the site. The role of dispersal in this Gleasonian view of community assembly is currently expressed in discussions of the species pool (Ricklefs and Jenkins 2011), propagule pressure (Simberloff 2009), and neutral theory (Hubbell 2001, 2005), and the role of successful establishment in a community in the concept of environmental filtering (Gotzberger et al. 2012; Kraft et al. 2015), niche theory (Chesson 2011), coexistence theory (Chesson 2000; HilleRisLambers et al. 2012), and trait-based approaches to community assembly (McGill et al. 2006; Shipley 2010). A decidedly bottom-up, Gleasonian perspective organizes all these lines of contemporary research on plant community assembly, which typically is concerned with the functional diversity within and among the species in a region that allows a subset of those species to co-occur in a particular plant community associated with certain characteristic environmental conditions.

An allied field of research, variously characterized as community phylogenetics or ecophylogenetics (Mouquet et al. 2012; Ndiribe et al. 2013; Pearse et al. 2014), has become prominent since the publication of an influential paper by Webb et al. (2002) and the increasing availability of DNA sequence data. This research respects the elements of Gleason’s individualistic concept, but adopts a more top-down approach to assessing the nature of community assembly. The approach combines data on community composition and the traits of species in the community with a phylogenetic hypothesis for evolutionary relationships among those species to infer the relative importance of phylogenetic and functional constraints on community assembly (Cadotte et al. 2013). The line of inference that underpins this approach depends on the assumption that phylogenetic similarity is a good proxy for functional similarity, an assumption that has been widely applied at taxonomic levels above the genus and at large spatial scales (Vamosi et al. 2009) but that recently has come into question (reviewed in Gerhold et al. 2015). For plant communities, the community phylogenetic approach has scarcely been applied at all in what Vamosi et al. (2009) refer to as the Darwin-Hutchinson zone, the realm of direct interactions among closely related species. There are only a few studies of ecophylogenetic constraints on community assembly among species within a genus of vascular plants. The most notable are in herbaceous species of Tetraria (Slingsby and Verboom 2006) and among woody species in the genera Quercus (Cavender-Bares et al. 2004a; Cavender-Bares and et al. 2004b), Pinus, Ilex (Cavender-Bares et al. 2006), and Salix (Savage and Cavender-Bares 2012). However, the spatial scale at which these studies have been done is larger than the scale at which biotic interactions most often occur, that is, at the neighborhood spatial scale (Stoll and Weiner 2000), where the consequences of adaptive differentiation during past speciation events should be most evident during community assembly.

Herbaceous species in Carex and allied genera such as Eriophorum and Trichophorum in the family Cyperaceae provide an opportunity to extend tests of evolutionary and ecological theory related to assembly of plant communities down to the scale of interactions among neighboring individuals (Elliott et al. 2016) using plots on the order of 1 m2 or less. In this paper we employ sampling and analytic protocols congruent with the phylogenetic, spatial, and environmental scales at which the processes of community assembly actually occur in plant communities. First, we work at a spatial scale in which individual plants interact, considering
the patterns of co-occurrence among immediate neighbors within 0.25 m² quadrats in subarctic fens, a well-defined and well-studied habitat that has a high level of fine-scale environmental heterogeneity within sites (Bubier 1995; Payette and Rochefort 2001; Gignac et al. 2004; Dabros and Waterway 2008; Rochefort et al. 2012). Second, we work at the phylogenetic scale of a single clade within the large and diverse family Cyperaceae, comparing a set of species from the Cariceae-Dulichieae-Scirpeae clade (Muasya et al. 2009; Léveillé-Bourret et al. 2014; Global Carex Group 2015) within a single geographic region, the central Labrador Peninsula in eastern Canada. Third, we assess phylogenetic relationships among the sedges using a molecular data matrix with only a single missing cell across four gene regions that include the nuclear ribosomal spacers ETS and ITS, plus both a coding gene (matK) and non-coding regions (trnL intron, trnl-trnF intergenic spacer) from the plastid genome, to derive a nearly completely resolved phylogenetic hypothesis for the species of interest. Fourth, we quantify water pH in the root zone of the sampled individuals and their rooting position in relation to the water table, two environmental variables known to affect the distribution of the studied sedges (Gignac et al. 2004; Dabros and Waterway 2008). Combining data on phylogeny, co-occurrence, and environmental factors affecting their distribution allows a direct test of the contribution of phylogenetic constraints during community assembly, including the means to segregate environmental and phylogenetic factors influencing species co-occurrence. We focus in particular on testing the longstanding expectation, based on the concept of limiting similarity, that the more closely related species are, the less likely they are to occur together (Violle et al. 2011). We also consider the nature of stabilizing niche differences (Chesson 2000) that allow sets of more distantly related sedges to co-occur at the neighborhood scale.

Materials and Methods

Study Area and Sampling Strategy—We conducted this study in the central Labrador Peninsula near the iron-mining town of Schefferville, Quebec. This is a geologically and topographically complex region known as the Labrador Trough (Fig. 1; Conliffe 2015), from which the Wisconsinan ice retreated only about 6500 yr ago (Janson 2003; Dyke 2004). The regional vegetation is a mosaic of plant communities ranging from dry tundra on low, windswept ridges to lichen woodland and peatlands in the intervening valleys (Waterway et al. 1984). Glacial deposits and landforms combined with heterogeneous bedrock geology create a landscape especially rich in fens, minerotrophic peatlands that grade continuously from nutrient-poor fens in more acidic areas associated with runoff from shales to relatively rich fens associated with runoff from dolomite rocks (Payette and Rochefort 2001; Vitt 2002; Hajek et al. 2005; Rochefort et al. 2012). These numerous, topographically discrete fens with their high diversity of sedges (Cyperaceae) provide an excellent study system to explore the impact of evolutionary history on the assembly of plant communities.

The ecological data set is an expansion of the one analyzed by Dabros and Waterway (2008) to demonstrate the environmental affinities of sedges in the fens of the Schefferville region. The data matrix analyzed here includes the 21 fens sampled for that study plus an additional eight fens sampled in the following year to bring the total number of quadrats sampled to 2,124. Fen names, locations, estimated sizes, sedge species richness, the total number of quadrats sampled in each fen, along with environmental data, are given in Table 1, and the distribution of the fens on the landscape is shown in Fig. 1. The sampled fens varied in size from less than 1 hectare (10 fens) to more than 10 hectares (seven fens), with the largest being 19.6 hectares. Since environmental heterogeneity is high within these fens, we stratified fine-scale (0.25 m²) quadrat sampling across and within the 29 fens. We sampled a series of five transects placed across the short axis of each fen, with each transect randomly located in one of five equal-sized segments spaced along the long axis. For the three smallest fens, only three transects were sampled. Deep, open pools within the fens were not considered eligible for sampling because no vascular plants were growing in them. Percent cover of each species was estimated within the 0.25 m² quadrats that were randomly placed within each 5 m segment along each transect using the ‘‘ignorant man’’ technique (Ward 1974). The number of quadrats per transect was thus proportional to the width of the fen, resulting in more quadrats in large fens than in smaller ones. Abundance was estimated as percent cover using a modified Braun-Blanquet scale (Mueller-Dombois and Ellenberg 1974) 0 (not present), 1 (<5 %), 2 (5–25 %) and 4 (>75 %). Voucher specimens to verify and document identifications were deposited in the McGill University Herbarium (MTM). Position of the water table relative to the shoot bases of the plants (hereafter referred to as rooting level) and the pH of the water in the root zone (hereafter referred to as water pH or simply as pH) were measured in each quadrat using a calibrated pole and a portable Sutron 10489 pH meter (Welling Inc., Van der Waalspark, Netherlands), respectively. All field data were collected during midsummer (July to early August) when the water levels in the fens were stable except during and immediately after heavy rain events, when we did not do field sampling. Data on species occurrence, water pH, and rooting level are given for each of the 2,124 sampled quadrats in Supplemental Table S1. DNA from fresh samples was isolated using a modified CTAB protocol (Doyle and Doyle 1987) or from silica-dried samples of each species using a semi-automated Autogen 850 DNA extractor (Autogen, Holliston, Massachusetts) following the manufacturer’s recommended protocol for plants as described in Smith and Waterway (2008). ETS, ITS and trnl-trnF were amplified and sequenced as described in Waterway and Starr (2007). The matK region was amplified and sequenced using matk-21 forward and matK-5 reverse primers and PCR conditions as listed in Hollingsworth (2009, appendices 2 and 3). For this study, we used 146 sequences we had previously deposited in GenBank, six additional sequences from GenBank, and we added 71 new sequences that we also deposited in GenBank (Appendix 1). For 75% of the species, all four gene regions were represented by sequences from the same individual.

Phylogenetic Analysis—The DNA matrix for the regional species pool was augmented with sequences from Dulichium arundinaceum to serve as an outgroup. Tri-echium, representing the mono-typic genus Dulichium, is sister to the rest of the Cariceae-Dulichieae-Scirpeae (CDS) clade, a major lineage of Cyperaceae that includes all three genera found in the studied fens (Muasya et al. 2009; Léveillé-Bourret et al. 2014). Three species in the Siderostictae clade and C. gibba were also added to allow us to estimate clade divergence times. The Siderostictae clade has been shown to be sister to all other species of Carex (Waterway et al. 2009; Starr et al. 2015) and the divergence date between Siderostictae and non-Siderostictae species of Carex was estimated by Escudero et al. (2012) based on fossil evidence. Carex gibba is sister to all other species of Carex subg. Vignea in all published analyses, so it was included to help establish the crown age of the Vignea clade. DNA sequences were edited in ChromasPro v.1.7.6 (Technelysium Pty. Ltd, Brisbane, Australia), aligned with ClustalW v2 (Larkin et al. 2007) and the alignment manually refined using Mesquite v.3.02 (Maddison and Maddison 2008). The concatenated alignment file is deposited in the Dryad Digital Repository (http://dx.doi.org/10.5061/dryad.q810f).

All phylogenetic analyses were performed using the CIPRES Science Gateway v3.3 platform (Miller et al. 2010) and the cluster located in Andalusian Scientific Information Technology Center (CICA, Spain). The final phylogenetic analysis was implemented in MrBayes 3.2.6 (Ronquist and Huelsenbeck 2003) consisting of two independent runs of 10 million generations of a data set partitioned by the four genes with appropriate models applied
independently to each partition (Table 2). A second phylogenetic analysis to estimate divergence times was performed using BEAST 2.3.0 (Drummond et al. 2012; Bouckaert et al. 2014) with an uncorrelated log-normal model of molecular evolution and a calibrated Yule model as the tree prior for $1.5 \times 10^9$ generations (Heled and Drummond 2015).

Two independent calibration points with normal distribution were used to estimate the divergence times of the Carex phylogeny following Escudero et al. (2012): 42.2 million years ($\sigma = 6.5$) for the crown age of the genus Carex and 30.8 million years ($\sigma = 5$) for the crown age of the non-Siderostictae Carex. We assessed convergence and stationarity under Bayesian inference using the graphical analyses implemented in Tracer v1.4 (Rambaut and Drummond 2007); a total of 9,000 and 8,501 post-burn-in trees were pooled and used to generate a 50% majority-rule tree and a maximum clade credibility tree for MrBayes and BEAST results, respectively. We also tested two alternatives for calibrating the tree in BEAST, one using secondary calibration points based on the analysis of Escudero and Hipp (2013) for the root of the tree (43.6 million years, $\sigma = 4$) and for the crown age of Carex (32.6 million years, $\sigma = 2.9$), and another using the same secondary calibration point for the root, but using a late Eocene fossilized perigynium from C. colwellensis Chandler (33.9–41.3 million years), considered to be the oldest unequivocal fossil perigynium as reviewed in Smith et al. (2010) and Jiménez-Mejías and Martinetto (2013), to calibrate the crown age of Carex using a log-normal prior distribution with the mean of 32.7 million years placed at the origin of the crown node and the probability distribution accounting for the uncertainty.

Phylogenetic Distance Between Species Pairs — All statistical analyses described here and subsequently were conducted using R version 3.2.0 (R Core Team 2015). We computed the cophenetic distances (Paradis et al. 2004) for species pairs assessed in the co-occurrence analysis in two ways: one using branch lengths calibrated in number of changes/site/time on the 50% majority-rule tree described above from the MrBayes analysis as a measure of sequence divergence (hereafter referred to as dBayes), and the other using branch lengths calibrated in years on the ultrametric tree from the BEAST analyses (hereafter referred to as dBEAST) as a measure of phylogenetic distance in years. The cophenetic metric uses branch lengths of a given phylogenetic tree to measure the distance between species pairs. Following the suggestion of Letten and Cornwell (2015), we used a square-root transformation for the dBEAST distances to avoid overweighting deep time relative to recent time. The

Fig. 1. Locations of fens sampled along the western edge of the Labrador Trough near the town of Schefferville, QC, in the central Labrador Peninsula. Note the many elongated and interconnected water bodies organized by the strong NW-SE orientation of the ridges and valleys. See Table 1 for full names and other detailed information about the individual fens.
their different levels of species diversity and hence, divergence rates of divergence within as monophyletic and are easily distinguished by morphology. Levels ages found in the fens. These generic clades are each strongly supported level A to reflect the highest level of divergence among major sedge line-

\[ Eriophorum \]

\[ - \]

\[ SA \] Satellite 18 52 2.52 54.803260°, 
\[ RU \] Runway 20 167 11.58 54.804554°, 
\[ PH \] Philemon 7 32 0.85 54.728692°, 
\[ PA \] Paupercula 11 49 0.80 54.835136°, 
\[ OR \] Orchid 17 98 13.10 54.730925°, 
\[ NC \] Nasa 4 10 83 2.39 54.866498°, 
\[ MI \] MiniSatellite 11 17 0.25 54.805203°, 
\[ LO \] Leo 20 93 12.16 54.669671°, 
\[ KL \] Klucha 15 80 1.90 54.857691°, 
\[ KU \] Kluska 7 58 1.67 54.988979°, 
\[ GR \] Greenbush 24 106 14.91 54.874294°, 
\[ OR \] Orchid 17 98 13.10 54.730925°, 
\[ ND \] Nasa 3 10 83 2.39 54.866498°, 
\[ MI \] MiniSatellite 11 17 0.25 54.805203°, 
\[ MY \] Myszka 15 51 1.16 54.857691°, 
\[ NA \] Nasa 1 14 104 10.84 54.851714°, 
\[ NB \] Nasa 2 12 93 6.93 54.851714°, 
\[ NC \] Nasa 3 10 83 2.39 54.851714°, 
\[ ND \] Nasa 4 10 95 5.16 54.851714°, 
\[ OR \] Orchid 17 98 13.10 54.851714°, 
\[ PA \] Paupercula 11 49 0.80 54.835136°, 
\[ PE \] Peleliu 17 77 3.84 54.835136°, 
\[ PH \] Philemon 7 32 0.85 54.835136°, 
\[ RS \] Rusty 6 17 0.33 54.835136°, 
\[ RU \] Runway 20 167 11.58 54.835136°, 
\[ SA \] Satellite 18 52 2.52 54.835136°, 
\[ TA \] Tadpole 1 11 20 0.65 54.835136°, 
\[ TD \] Tadpole 2 11 20 0.65 54.835136°, 
\[ TK \] Teeka 8 30 0.67 54.835136°,
Fig. 2. Phylogram based on Bayesian analysis using MrBayes for sedge species in the Schefferville region plus selected Asian species added for rooting and dating the tree. Branch lengths are proportional to the number of nucleotide changes/site/time. Posterior probabilities > 0.7 are shown above the branches. Sectional placements are given for Carex species in the region, and habitat categories are based on Cayouette (2008) and Ball et al. (2002) and coded by color. Species sampled for co-occurrence are followed by an asterisk. Species not in the Schefferville regional species pool are shown in gray.
The most recent divergence level, the one between species within subclades, was defined as level D and included comparisons at this level from within both the Core Carex and the Vignea clades. For the fen species, comparisons at this level were those within clades representing sections Limosae, Vesiariae, and Paniicate in the Core Carex clade, and sections Glareosae and Stellulatae in the Vignea clade. All of these clades were strongly supported by the molecular and morphological synapomorphies. Level D1 mostly represented divergences within the last 5 million years with some being much younger (e.g. C. rostrata and C. utriculata estimated to have diverged only 200,000 yr ago, Fig. 3). We also tested an alternate definition of level D (D2) that differed only in how we treated the clades representing sections Limosae, Vesiariae, and Paniicate. These three subclades form a larger clade within Core Carex (Figs. 2, 3; Waterway et al. 2009) which also has a posterior probability of 1.0 in our analyses and is comparable in age to the Paniicate subclade (~11 MYA). In that alternative analysis we transferred all pairs in which the two species were from different subclades within the Wetland clade from the C level to the D level. These two different definitions for level D are the only divergence between what we term Clade Divergence Level 1 (CDL1) and Clade Divergence Level 2 (CDL2) in the analyses described below in which we test which factors best predict significant pairwise species co-occurrence.

Predicting Co-occurrence—Using the R package co-occur (Veech 2013; Veech 2014; Griffith et al. 2016) we analyzed the pairwise association between sedge species in fens of the Schefferville region at the boreal-subarctic transition in northern Quebec. This co-occurrence algorithm classifies species pairs as positively, negatively, or randomly associated by calculating observed and expected frequencies of co-occurrence between each pair of species and returning probabilities that a more extreme (either high or low) value of co-occurrence could have been obtained by chance (Veech 2013; Griffith et al. 2016). The effect size of co-occurrence, that is the strength of the association between a given species pair, is expressed as the difference between observed and expected frequencies (Griffith et al. 2016). The effect size quantifies the strength of species attraction (more positive values) or aversion (more negative values). We use ‘attraction’ and ‘aversion’ here and in the following discussion only as terms to indicate that the probability of pairwise co-occurrence is greater or less than expected by chance, with no implication of specific mechanisms to attract or avoid. This probabilistic model of species co-occurrence does not rely on data randomization and thus has the desirable properties of a low Type I error rate as well as the power of a low Type II error rate (Veech 2013). Following Veech (cf. Griffith et al. 2016), we dropped from subsequent analyses all species pairs that occurred in so few quadrats that their expected frequencies were less than one.

We conducted the initial co-occurrence analyses in two modes: 1) in a regional analysis pooling all the data without regard to the fen in which a quadrat occurred; and 2) in separate analyses for each of the 29 sampled fens. In the first mode the quadrates were considered as stratified random samples of the fen habitat in the Schefferville region. Note that in this mode of analysis fewer pairwise comparisons are discarded due to expected values falling below one, but on the other hand, any filtering effects due to environmental or size variation among fens will not be detected. The analyses in the second mode respect the possible effects of environmental differences along the gradient from poor to rich fens known to occur in the region (Bubier 1995; Dabros and Waterway 2008) but the dimensionality of the data matrices varied with the number of quadrats sampled in each fen and the number of sedge species observed. We standardized effect sizes (SES) calculated for each species pair by dividing them by the total number of quadrats sampled in a given fen to enable cross-fen comparisons. Although the fens retain their individual identity, sample size effects led to discarding more pairings with expected values below one than in the pooled data set. We report and compare the patterns of species association detected in both modes of analysis. In subsequent linear mixed modeling analyses relating species association to the degree of relatedness between species and differences in their environmental affinities, we used only the second, more conservative co-occurrence result, having combined SES values from all fens into a single table consisting of 1,278 fen by species-pair records (Supplemental Table 2, available from the Dryad Data Repository at http://dx.doi.org/10.5061/dryad.q8b10).

Predicting Co-occurrence—We used linear mixed modeling (R package lme4; Bates et al. 2015) to predict these fen by species-pair SES results as a function of the phylogenetic distance between species pairs (dBayes or dBEAST), the mean difference in the environmental affinities of each species pair (water pH and rooting level), and the two previously described factors coding for different definitions of clade divergence.
levels (CDL1 and CDL2). We first developed four model sets in order to (1) distinguish the relative importance of clade versus distance-based assessments of phylogeny, (2) assess the significance of either of two valid ways of defining clade levels (CDL1 vs CDL2) and phylogenetic distances (dBayes vs. dBEAST), and (3) test the importance of these six variables while avoiding collinearity between fixed effects since the two phylogenetic distance metrics were very similar to each other (r ˃ 0.6) as were the CDL variables. The four models all included differences in the environmental affinities of each species pair with regard to water pH and rooting level, but differed in whether they used dBayes or dBEAST distances and whether they used the CDL1 or CDL2 clade divergence levels. The four models thus represent all possible combinations of the two phylogenetic distance metrics with the two CDL schemes; however, no model included both types of distance metric and no model included both sets of CDL variables. We used standard diagnostic tests (Zuur et al. 2009) separately for each model set to select the best model in each case. Tests included verifying the significance of fixed effects using likelihood ratio tests, examining model residuals for heterogeneity and violations of normality and, finally, comparing model fit using the AIC. We then compared the best models from each model set to identify the single best model across all model sets. In all these models, we used two separate factors as crossed random effects after having verified their significance using likelihood ratio tests: (1) the fen in which the observations were made; and (2) the species pair being assessed. The overall intercept of our models thus reflects a weighted average over fens and species pairs. These analyses were repeated using the dBEAST values from the alternative dating calibrations described above to test their effect on the conclusions.

Mean Pairwise Differences in Environmental Affinities—We calculated mean differences in environmental affinities for water pH and rooting level for each species pair that was assessed in each fen for the co-occurrence analysis (see below). We first calculated the weighted mean differences of both pH and rooting level for a given species across all quadrats in which it occurred in a given fen. Values were weighted by the midpoints of the abundance classes of that species where it occurred in the fen. In this way we estimated multiple fen- and species-specific weighted mean values as opposed to a single value for each species across all fens. We then calculated the absolute difference in the weighted mean water pH (mean delta pH) and rooting level (mean delta rooting level) for each species pair at each fen.

For species pairs in the level D clade divergence category (both variants) that were significantly less likely to occur than expected by chance in the fen-by-fen co-occurrence analysis, we used Tukey’s honestly significant difference test (HSD) for significant differences in water pH and rooting level between the two species in each pair based only on quadrats in which one or the other of the two species in the pair occurred, but not those where they co-occurred. We compared environmental variables for these species pairs both at the regional level using appropriate quadrats from all fens and on a fen-by-fen basis.

Ordination of Quadrat Data on Sedge Community Composition—We used a detrended correspondence analysis (DCA), calculated using the R package vegan (Oksanen et al. 2015), to illustrate patterns of sedge species covariance across all fens and to see how the ordination reflected underlying pH and rooting level gradients. Specifically, we analyzed the entire 2,124 quadrat by 27 sedge species matrix using DCA with abundance estimates for each species on a 0–4 scale, as described above. We then fit vectors with measures of pH and rooting level at each quadrat onto the resulting ordination. To draw comparisons with results from the linear mixed model analysis, we color-coded the labels for species vectors according to their subclade membership (D1 divergence level), which also corresponded to their sectional affiliations.

Results

Species Distribution and Abundance—Sedges are so widespread and frequent in fens of the Schefferville region that only one of the 2,124 randomly sampled quadrats across the 29 fens lacked sedge species. The species varied in frequency with the most frequent being found in about a third (Carex aquatilis, T. cespitosum) to nearly half (C. limosa) of the 2,124 sampled quadrats, and the least frequent, C. saxatilis and E. vaginatum, found in only 10 and four quadrats, respectively (Table 3). Carex saxatilis and E. vaginatum were each found in only two fens, whereas C. magellanica and E. chamissonis were the two most widespread species, found in 27 and 26 of the 29 fens, respectively. In addition to the 27 sedge species found in the quadrats, four Carex species were infrequently observed in the fens but not sampled in any of the random quadrats: C. brunnescens, C. capitata, C. diandra, and C. vesicaria. The number of sedge species per 0.25 m² quadrat varied from one to eight, with a mean of 2.8 and a median of three. A total of 328 (15%) of the quadrats had only one sedge species and thus provided information on environmental tolerances but could not be used in the analysis of pairwise co-occurrence.

Phylogenetic Distance—Based on the ultrametric BEAST tree (Fig. 3), phylogenetic distances (dBEAST) measured in millions of years between species pairs of Carex in the ecological data set from the fens varied from 0.5–62 while phylogenetic distances between species pairs representing different genera in the CDS clade varied from 102–111 million years. The square roots of these distances are strongly correlated with mean phylogenetic distances based directly on levels of sequence divergence that were calculated from a set of 9000 trees in the clade credibility set from Mr. Bayes (dBayes; r = 0.96). The DNA data matrix is complete for these species, so no bias has been introduced by incomplete gene sampling in some parts of the tree compared to others, and as a result, the high correlation is not surprising. This also suggests that although the estimated divergence times have a high variance and depend on limited fossil evidence, the relative divergence times across the phylogenetic tree are likely to be robust.

We also used two alternative calibrations based on dates from Escudero and Hipp’s (2013) paper on Cyperaceae phylogeny. That paper used a younger (Eocene) fossil (C. colvensis) to calibrate the crown node of Carex rather than the older (Paleocene) putative fossil (C. tsagajanica) to calibrate the stem node of Carex. These two alternative calibrations resulted in smaller phylogenetic distances (as low as 0.5–47 million years for species pairs of Carex and as low as 77–84 million years for species pairs representing different genera in the CDS clade) (Supplemental Figs. 1, 2, available from the Dryad Data Repository at http://dx.doi.org/10.5061/dryad.q810f). Crown ages for the three largest Carex clades (Core Carex, Caricoid, and Vigno) in these two alternative analyses ranged from 19–44 million years, similar to the 15–16 million year crown ages found by Spalink et al. (2016). Despite these differences, the crown ages for the three largest Carex clades in our original analysis ranged from 23–19 million years, similar to the crown ages of approximately 20 million years found for these clades in a recent analysis of Poales using 9 fossils for calibration (Bouchenak-Khelladi et al. 2014) and in line with the fact that Carex fossil diversity was already established in the Oligocene, that is, prior to 23.7 million years ago. We therefore present our original BEAST analysis here (Fig. 3), with the caveat that estimating divergence times in Carex will remain an inexact science at least until additional fossils can be reliably determined. Although the absolute distances differed among analyses, the relative branch lengths were maintained across all three calibration methods, and all three gave the same results in our subsequent analyses.

Patterns of Pairwise Co-occurrence—When pooling data from all fens into a single regional analysis, we were unable to test 44 of the 351 possible species pairs because the expected probability of their co-occurrence was less than one
due to the rarity of a few species, particularly *E. vaginatum* and *C. saxatilis*. Of the 307 testable species pairs, 120 (39%) demonstrated significant levels of aversion and 72 (23%) showed significant levels of attraction (Fig. 4A; Table 4). Comparisons within subclades (D-level) differed markedly between the Core *Carex* clade and the *Vignea* clade. Six (67%) of the D1-level pairwise comparisons showed aversion while the rest were neutral.

In contrast, one species pair (17%) in the Cashสะพด clade showed significant attraction while the remainder were neutral.

The infrageneric pattern was different for the smaller genera: the two species of *Eriophorum* became closer, with the exception of the *E. vaginatum* vs. *E. angustifolium* pair showed significant aversion in the regional analysis, but no significant pattern in the fen-by-fen analysis.

The general tendency for more species pairs to show aversion than attraction varied depending on the clade divergence level of the pairs and on the particular clad (Fig. 4; Table 4). In the regional analysis, comparisons between species of different genera (A-level) and those between species from different major clades within *Carex* (B-level), as well as those between different subclades within the large Core *Carex* clade (C- and Cv-levels) all showed higher levels of aversion than attraction, substantially higher for the C levels (Fig. 4A; Table 4). Comparisons within subclades (D-level) differed markedly between the Core *Carex* clade and the *Vignea* clade. Six (67%) of the D1-level pairwise comparisons in Core *Carex* showed aversion while the rest were neutral. In contrast, one species pair (17%) in the *Vignea* clade showed significant attraction while the remainder were neutral.

The 2D-level comparisons, in which the D-level for the Core *Carex* clade was expanded to include all pairwise comparisons for the Wetland clade (see Fig. 3), still had 16 (55%) aversive pairs and only 10% attractive pairs, the others being neutral. The pattern was similar in the fen-by-fen analysis (Fig. 4B; Table 4). Overall, within *Carex* the number of pairwise comparisons demonstrating aversion increased as the phylogenetic relationship between the species pairs became closer, with the exception of the *Vignea* clade in which pairwise co-occurrence was neutral, or in one pair, significantly attractive. The infrageneric pattern was different for the smaller genera: the two species of *Eriophorum* showed strong attraction in both regional and fen-by-fen analyses, as did one pair of *Eriophorum* species. Another *Eriophorum* pair showed significant aversion in the regional analysis, but no significant pattern in the fen-by-fen analysis.
**Predicted Co-occurrence using Linear Mixed Modeling**—
A likelihood ratio test comparing models with and without the specified random effect structure showed that the model with the random effect structure was a better fit than the one without for all model sets (Table 5). Specifying the appropriate random effect structure, we then sequentially dropped non-significant fixed effects. The models not including the phylogenetic distance metrics (either dBayes or dBEAST) were better fit than the ones with these variables for all model sets (Table 5), i.e. cladistic structure was more informative than cophenetic distance per se. Once these distance metrics were removed, all remaining fixed effects were significant or had factor levels that were significant (Table 6). Both the mean delta pH and mean delta rooting level were Box-Cox transformed to reduce heterogeneity in the residuals. We compared the best models from all model sets, all of which included mean delta pH and mean delta rooting level as fixed effects but differed in whether cladistic relationships between species pairs were coded using CDL1 or CDL2 (i.e. the alternative cladestistics previously described). Using either of these two cladistics variables as fixed effects made no difference in terms of overall model fit ($\Delta AIC = 1.1, \chi^2 (df = 0) = 0, p = 1$). In all of the best models, SES decreased with increasing mean delta rooting level and mean delta pH (Table 6; Fig. 5), indicating, not surprisingly, that the greater the difference in environmental affinities, the less likely two species are to co-occur at the scale of 0.25 m$^2$ quadrats. In models with either CDL1 or CDL2, the level coding for the most closely related species pairs (D1 and D2, respectively) had, on average, significantly low values of SES, indicating they are unlikely to co-occur. Values of SES for D1 were slightly more negative and variable than for D2. The clade divergence level S had significantly high values for both CDL1 and CDL2 indicating that, in contrast to species pairs within Carex, species pairs within Eriophorum and Trichophorum are likely to co-occur. Mean values of SES for all other divergence levels (A, B, Cc and Cv) were not significantly different from each other (Table 6) in either CDL1 or CDL2 and were all negative.

Repeating the linear mixed modeling using dBEAST distances based on the alternative calibrations demonstrated that the results described above were robust with respect to the calibration methods used to estimate divergence times; values were nearly identical and all results remained the same.

**Fen Community Structure and Environmental Conditions**—
The number of sedge species in each fen varied from 3–24 with a positive correlation between species richness and fen size ($r^2 = 0.569, p < 0.001$). The fens also varied in the range of rooting levels and especially of water pH values that we recorded from the quadrats (Table 1; Fig. 6). In four fens,
Table 4. Tabular summary of co-occurrence results, giving the number of attractive, aversive and neutral pairs in each of the regional and fen-byn fen analyses for each of the two clade divergence level definitions (CDL1 and CDL2). Clade divergence levels are lettered as in Fig 3: A refers to divergence between species that belong to different genera of the Carex subgenus; B to species that belong to different major clades within the Carex clade; C to species that belong to different subclades within the Carex clade; and D to species that belong to the same subclade within either the Carex or Vignea clades. The total number of testable pairs pertains to those species pairs per CDL whose expected co-occurrence was greater than one for any given analysis.

<table>
<thead>
<tr>
<th>Clade Divergence Level</th>
<th>A</th>
<th>B</th>
<th>C1/C2</th>
<th>Cv</th>
<th>D1/D2</th>
<th>S</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDL1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regional</td>
<td>Attractive Pairs</td>
<td>26 (23.9%)</td>
<td>33 (28.0%)</td>
<td>7 (20.6%)</td>
<td>3 (11.1%)</td>
<td>1 (6.7%)</td>
<td>2 (50.0%)</td>
</tr>
<tr>
<td>Aversive Pairs</td>
<td>39 (35.8%)</td>
<td>47 (38.9%)</td>
<td>21 (61.8%)</td>
<td>6 (22.2%)</td>
<td>6 (40.0%)</td>
<td>1 (25.0%)</td>
<td>120 (39.1%)</td>
</tr>
<tr>
<td>Neutral Pairs</td>
<td>44 (40.4%)</td>
<td>38 (32.2%)</td>
<td>6 (17.6%)</td>
<td>18 (66.7%)</td>
<td>8 (53.3%)</td>
<td>1 (25.0%)</td>
<td>115 (37.5%)</td>
</tr>
<tr>
<td>Total Testable</td>
<td>109</td>
<td>118</td>
<td>34</td>
<td>27</td>
<td>15</td>
<td>4</td>
<td>307</td>
</tr>
<tr>
<td>Fen by Fen</td>
<td>Attractive Pairs</td>
<td>23 (24.5%)</td>
<td>29 (28.7%)</td>
<td>8 (22.9%)</td>
<td>0 (0.0%)</td>
<td>5 (16.7%)</td>
<td>2 (66.7%)</td>
</tr>
<tr>
<td>Aversive Pairs</td>
<td>31 (33.0%)</td>
<td>27 (26.7%)</td>
<td>16 (45.7%)</td>
<td>4 (22.2%)</td>
<td>5 (41.7%)</td>
<td>0 (0.0%)</td>
<td>83 (31.6%)</td>
</tr>
<tr>
<td>Neutral Pairs</td>
<td>40 (42.6%)</td>
<td>45 (44.6%)</td>
<td>11 (31.4%)</td>
<td>4 (77.8%)</td>
<td>6 (50.0%)</td>
<td>1 (33.3%)</td>
<td>117 (44.5%)</td>
</tr>
<tr>
<td>Total Testable</td>
<td>94</td>
<td>101</td>
<td>35</td>
<td>18</td>
<td>12</td>
<td>3</td>
<td>263</td>
</tr>
<tr>
<td>CDL2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regional</td>
<td>Attractive Pairs</td>
<td>26 (23.9%)</td>
<td>33 (28.0%)</td>
<td>4 (25.0%)</td>
<td>3 (11.1%)</td>
<td>1 (6.2%)</td>
<td>2 (50.0%)</td>
</tr>
<tr>
<td>Aversive Pairs</td>
<td>39 (35.8%)</td>
<td>47 (38.9%)</td>
<td>11 (68.8%)</td>
<td>6 (22.2%)</td>
<td>18 (66.7%)</td>
<td>1 (25.0%)</td>
<td>120 (39.1%)</td>
</tr>
<tr>
<td>Neutral Pairs</td>
<td>44 (40.4%)</td>
<td>38 (32.2%)</td>
<td>1 (6.2%)</td>
<td>18 (66.7%)</td>
<td>13 (39.4%)</td>
<td>1 (25.0%)</td>
<td>115 (37.5%)</td>
</tr>
<tr>
<td>Total Testable</td>
<td>109</td>
<td>118</td>
<td>16</td>
<td>27</td>
<td>33</td>
<td>4</td>
<td>307</td>
</tr>
<tr>
<td>Fen by Fen</td>
<td>Attractive Pairs</td>
<td>23 (24.5%)</td>
<td>29 (28.7%)</td>
<td>8 (22.9%)</td>
<td>0 (0.0%)</td>
<td>1 (8.3%)</td>
<td>2 (66.7%)</td>
</tr>
<tr>
<td>Aversive Pairs</td>
<td>31 (33.0%)</td>
<td>27 (26.7%)</td>
<td>16 (45.7%)</td>
<td>4 (22.2%)</td>
<td>5 (41.7%)</td>
<td>0 (0.0%)</td>
<td>83 (31.6%)</td>
</tr>
<tr>
<td>Neutral Pairs</td>
<td>40 (42.6%)</td>
<td>45 (44.6%)</td>
<td>6 (17.6%)</td>
<td>14 (77.8%)</td>
<td>8 (53.3%)</td>
<td>1 (33.3%)</td>
<td>117 (44.5%)</td>
</tr>
<tr>
<td>Total Testable</td>
<td>94</td>
<td>101</td>
<td>35</td>
<td>18</td>
<td>12</td>
<td>3</td>
<td>263</td>
</tr>
<tr>
<td>Total Testable Pairs</td>
<td>134</td>
<td>128</td>
<td>361</td>
<td>17</td>
<td>29</td>
<td>17</td>
<td>35</td>
</tr>
</tbody>
</table>

Table 5. \( \chi^2 \) values from likelihood ratio tests comparing models with and without variables listed in the “Variable Tested” column. Random effects pertain to crossed factors coding for the fen in which the observations were made and the species pair being assessed. Asterisks are printed beside \( \chi^2 \) values to signify significance levels, with ** for \( p < 0.001 \).

<table>
<thead>
<tr>
<th>Variable Tested</th>
<th>Test Statistic</th>
<th>Model Set 1</th>
<th>Model Set 2</th>
<th>Model Set 3</th>
<th>Model Set 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Effects</td>
<td>( \chi^2 ) (df=3)</td>
<td>143.51***</td>
<td>155.06***</td>
<td>155.07***</td>
<td>143.49***</td>
</tr>
<tr>
<td>dBays</td>
<td>( \chi^2 ) (df=1)</td>
<td>0.395</td>
<td>0.471</td>
<td>0.357</td>
<td></td>
</tr>
<tr>
<td>dBEAST</td>
<td>( \chi^2 ) (df=1)</td>
<td>-</td>
<td>0.382</td>
<td>-</td>
<td>0.537</td>
</tr>
</tbody>
</table>

mean rooting levels were below the water table and in seven fens, the means were above the water table, with the remainder having mean rooting levels more or less at the water table. Water pH values for the fens varied from 3.4–7. The range of values within each fen was also highly variable, with ranges spanning as little as 0.3–0.5 pH units, mostly at the acidic end of the gradient, to fens with pH ranges spanning up to 4 pH units, mostly in fens with mean pH values above 5 (Table 1; Fig. 6). The two environmental variables were very weakly correlated (\( r^2 = 0.0322, p < 0.001 \)).

Ordination of the 2,124 quadrats showed broad spread on both synthetic axes; not surprisingly given the large data matrix, the first DCA axis explained only 7.5% of the variance and the second 6.4% (Fig. 7). When fit onto the ordination space, the squared correlation coefficients for the first two DCA axes were 0.38 (\( p < 0.001 \)) and 0.37 (\( p < 0.001 \)), respectively. Species vectors on the ordination plot suggested that most of the species tended to segregate from their closest relatives on the pH gradient, the rooting level gradient, or both.

Ecological Tolerances of the Species—Abundance-weighted means, standard deviations, and ranges across all fens for rooting level in relation to the water table and water pH are given in Table 3. Carex saxatilis, C. rostrata, C. utriculata, C. livida, E. chamissonis, and E. angustifolium all tended to root below the water table and frequently occurred in quadrats alone or with only one or a few other sedges, often without other competing plant species. In contrast, C. trisperma, C. vagnata, C. gymnocrates, C. pauciflora, C. leptalea, C. rafirifor, and C. tenuiflora all had mean rooting levels more than 10 cm above the water table, and grew on low to high hummocks that were richer in species, including ericaceous shrubs and non-graminoid herbaceous plants. Many species tended to grow at or just above the water table so quadrats at this level frequently had 3–7 sedge species. Carex pauciflora, C. trisperma, C. oligosperma, C. magellanica, and E. vagnatum had mean water pH values below 4.8, while C. gymnocrates, C. leptalea, C. heleonastes, C. livida, E. viridicarinatum, and T. alpinum occupied the opposite end of the pH gradient with mean values of 5.5. Many of the species showed a remarkable range of tolerance along the pH gradient, in some cases with individuals occupying quadrats from pH 3 to pH 8 (e.g. C. limosa, C. aquatilis, T. cespitosum). Not only was the spread among individual points large, but the means within fens for these species varied nearly as much (Fig. 8). The range of tolerance for different water

Table 6. Values for the coefficients, standard errors and t-values for each fixed effect in the best linear mixed model (LMM). T-values represent the ratios between coefficient estimates and their standard errors. Cladistic associations are coded using CDL1. Asterisks are printed beside t-values to signify significance levels, with ** for \( p < 0.01 \) and *** for \( p < 0.001 \).

<table>
<thead>
<tr>
<th>Fixed Effect</th>
<th>Coefficient Estimate</th>
<th>Standard Error</th>
<th>t-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean delta rooting level</td>
<td>-0.006654</td>
<td>0.000623</td>
<td>-10.685**</td>
</tr>
<tr>
<td>Mean pH</td>
<td>-0.033702</td>
<td>0.001428</td>
<td>-2.592**</td>
</tr>
<tr>
<td>CDL B</td>
<td>0.004872</td>
<td>0.003160</td>
<td>1.542</td>
</tr>
<tr>
<td>CDL Cc</td>
<td>-0.005986</td>
<td>0.004189</td>
<td>-1.429</td>
</tr>
<tr>
<td>CDL Cv</td>
<td>-0.003715</td>
<td>0.006859</td>
<td>-0.542</td>
</tr>
<tr>
<td>CDL D1</td>
<td>-0.020870</td>
<td>0.006963</td>
<td>2.998***</td>
</tr>
<tr>
<td>CDL S</td>
<td>0.044620</td>
<td>0.013050</td>
<td>3.420***</td>
</tr>
</tbody>
</table>
levels was also quite broad but 78% of the species had mean and median rooting depths between 0 and 15 cm, and the three species mentioned above with broad water pH ranges all had rooting level means within a narrow range at or above the water table.

With one exception, pairs of closely related species (D-level) that showed significant aversion in the co-occurrence analyses were all significantly different along both water pH and rooting level gradients based on the Tukey HSD test using data pooled across all fens (Table 7). *Carex limosa* and *C. utriculata* did not differ significantly from each other along either gradient. Within individual fens, the differentiation among species along the environmental gradients was not as obvious because sample sizes were smaller and often unbalanced. The relative relationship of the means for rooting level within fens was the same as the regional
relationship for a high proportion of the species pairs. Among
the species with significant differences at the regional level,
the only exceptions to this trend were the C. oligosperma–
C. aquatilis and C. utriculata–C. aquatilis pairs, in both of
which the within-fen relationships between the means matched
the regional relationship only about half the time (Table 7). The
concordance between the regional relationship and the within-
fen relationships was much less consistent along the water
pH gradient. The relative difference between the species in a
pair was consistent with the regional relationship in less
than 60% of the fens for five pairs, more than 85% of the fens
for five pairs, and between these two values for the remaining
five pairs (Table 7).

The positions of species vectors on the ordination plot
(Fig. 7) are consistent with the pairwise tests of environmen-
tal tolerances between species. Most closely related species
segregate on the axis that is strongly correlated with rooting
level, or the axis that is strongly correlated with water pH,
or both. Although there is high variance along one or both
of these gradients for many species and 95% confidence ellipses
for species in the same D1 level subclade overlap (Fig. 8),
the central tendency, shown by the species names at the end points
of the species vectors, differs for all species pairs except the most
recently diverged pair, C. rostrata–C. utriculata.

Discussion

By focusing on sedges growing in subarctic fens on the
central Labrador Peninsula, we have circumscribed a study
system well suited to testing hypotheses about the evolu-
tionary ecology of closely related species able to coexist as
immediate neighbors within the same habitat. First, since
fens are extensively distributed on this recently deglaciated
landscape (Fig. 1), dispersal is unlikely to limit the coloniza-
tion of individual fens by species in the regional species
pool. There is good evidence for rapid and wide dispersal
within this region during deglaciation (Alsos et al. 2015;
Gajewski 2015). Second, cool climates, long harsh winters,
low nutrient levels, frequently anoxic substrates, and the
often extreme pH conditions that make nutrient uptake dif-
ficult all impose a strong environmental filter on species in
the fen habitat. The sedge species occurring in an individual
fen will have been subjected to habitat filtering in the sense
this term is used in ecophylogenetics (Mouquet et al. 2012;
Mittelbach and Schemske 2015), allowing us to focus on the
biotic interactions among species within the fen habitat
during the process of community assembly. Third, since
successful colonization of this inherently harsh habitat
requires an unusual suite of adaptations, the sedge commu-
nity is dominated by relatively few species in the Cariceae-
Dulichieae-Scirpeae (CDS) clade of Cyperaceae. The number
of sedge species involved is tractable, yet has sufficient func-
tional and phylogenetic diversity to allow testing of theoretic-
al expectations.

We observed 31 species from this CDS clade in the fens of
the Schefferville region, including 25 Carex, four Eriophorum
and two Trichophorum; all but four were common enough to
have been found in our random sampling of 2,124 quadrats
across 29 fens. Three predominantly northern circumboreal
lineages, Limosae (three species), Vesicariae (four species), and
Glaresae (four species), account for 11 of the 21 Carex
species in our quadrats. The remaining 10 species, with the excep-
tion of C. leptalea (North America and extending south to
Mexico and the West Indies) and C. exilis (temperate to
boreal Eastern North America) all belong to circumboreal
lineages (including some that are also bipolar) or to northern
clades within more broadly distributed groups (e.g. C. aquatilis
in section Phacocystis (Dragon and Barrington 2008) and sev-
eral species in section Vesicariae (Gebauer et al. 2014)). From
two to 24 of these species in the CDS clade co-occurred in
the same fen, and up to eight species (median 3) were found
in the same 0.25 m² quadrat. The frequent co-occurrence of
so many closely related species of Carex and allied genera as
immediate neighbors in the studied fens affords an opportu-
nity to test existing theory for community assembly and
co-existence. We focus primarily on testing the expectation
that the more closely species are related, the less likely they

Fig. 6. Box plots for the 29 fens showing variation among the
sampled quadrats in each fen for A: rooting level (cm) and B: water pH. Fens
are ordered in terms of their median values on each panel; the acronyms
refer to fen names in Table 1.
We discuss both the tendency for species in the same subclade to be neighbors less often than expected by chance and the tendency for suites of species that commonly occur together in particular microhabitats to be only distantly related within the CDS clade of Cyperaceae. We also comment on the importance of considering spatial and phylogenetic scale in interpreting the observed patterns.

**Testing Phylogenetic Limiting Similarity**—This line of inquiry arises from a longstanding expectation expressed by Darwin (1859) that the more closely related two species are, the stronger will be their struggle for existence, and by implication, the lower will be their likelihood of co-occurring as immediate neighbors. This expectation, rooted ecologically in the concepts of limiting similarity (Abrams 1983) and evolutionarily in its obverse, ecological character displacement (Stuart and Losos 2013), is deeply embedded in niche theory (Chase 2011). The relevant expectations from neutral theory (Hubbell 2001), which assumes no functional differences among species and focuses solely on dispersal limitation, are unlikely to apply in our study system except as a null model. A strong test of the phylogenetic limiting similarity hypothesis in our study system necessarily lies at the interface between the evolved characteristics of the sedge species in the regional pool and the ecological processes that

---

**Fig. 7.** Detrended Correspondence Analysis (DCA) ordination plot of the Schefferville fen data. Grey dots are used to represent the 2,124 quadrats, many of which overlap, and labels for the end points of species vectors are color-coded according to genus or sectional membership of the species. Vectors for rooting level and pH were added to the ordination diagram, depicted by black arrows, pointing in the direction of increasing values.
govern both community assembly and the co-occurrence of neighboring individuals within the fen habitat. The different adaptations of species in the regional pool that allow their successful colonization of the fen have arisen over time through adaptive evolution and are essentially fixed in terms of present day community assembly, although genotypic diversity within species may play some role (Violle et al. 2012; Siefert 2014). Hence, it is primarily Chesson’s coexistence theory (Chesson 2000; HilleRisLambers et al. 2012) that bears most directly on the question of co-occurrence of the various species at the neighborhood scale (Stoll and Weiner 2000) within a given fen.

Chesson’s analysis is framed in terms of the population dynamics of species in a community, which arise in the aggregate influence of interactions among neighboring plants. Coexistence theory predicts that co-occurrence can result from stabilizing niche differences, relative fitness differences, or a combination of these two effects. Stabilizing niche differences are attributable to traits that lead to spatiotemporal segregation among neighboring species, e.g. differences in rooting depth or timing of peak seasonal growth. Relative fitness differences in contrast originate in quantitative differences in traits that mediate competitive interactions, e.g. differences in the photosynthetic capacity or nutrient uptake rates. Using our data on the phylogenetic relatedness between neighboring plants, the distribution of sedge species along pH gradients, the levels at which the species root, and any ancillary information on the functional traits that may be available for the species, we first test the hypothesis that the likelihood of two sedge species co-occurring as immediate neighbors is inversely related to their phylogenetic relatedness against a null hypothesis of random co-occurrence.

**Fig. 8** Graphs of pH vs. rooting level (cm) for species within subclades showing symbols to represent quadrats where they occur (open circles), the abundance weighted means within fens (triangles) and 95% confidence ellipses for each species in different colors. A. section Limosae subclade. B. Paniceae–Bicolores subclade. C. section Glareosae subclade. D. Trichophorum subclade.
Second, we test whether neighboring sedge species differ significantly in ecological tolerances along two environmental gradients: water pH and rooting level in relation to the water table, both known to be important in the distribution of sedges (Gignac et al. 2004; Dabros and Waterway 2008). This uses character displacement (Stuart and Losos 2013) along environmental gradients within and across fens as an indirect test of expectations from the phylogenetic limiting similarity hypothesis. In discussing the results, we use selected examples of specific species pairs to illustrate the nature of the functional differences affecting the likelihood that sedge species in these subarctic fens do or do not occur as immediate neighbors. We then turn from the results of these general tests to discussion of selected examples of specific species pairs to illustrate the nature of the functional differences that characterize patterns of ecological specification affecting present processes of sedge community assembly in these subarctic fens.

### Species Aversion

The pairwise co-occurrence results show that nearly all of the most closely related species in the large Carex clade, especially those in the Wetland clade, show significant aversion both across fens and within them (Fig. 4). In the regional analysis, which considers the quadrats as random samples of fen habitat in the region (Fig. 4A), nine of the 10 pairwise comparisons within subclades representing sections Limosae, Viscaricae, and Panicae in Core Carex (clade divergence level D1, which represents sister pairs or sets of very closely related species) have sample sizes large enough to test significance. Six of these show significant aversion, that is, the two species in the pair co-occur in the 0.25 m² quadrats much less often than expected by chance, while the others are neutral. In the analysis considering pairings within fens (Fig. 4B), five of the six pairs that show aversion in the regional analysis also occur significantly less often than expected by chance within the fens, the remaining pair being neutral. These results are consistent with expectations under phylogenetic limiting similarity.

There is, however, one apparent exception: C. oligosperma and C. utriculata show no significant aversion or attraction within fens, but the regional analysis shows their significant aversion. This is a good example of niche segregation expressed more strongly among than within the regional fens, which range from more acidic poor fens to more alkaline-rich fens (Table 1; Fig. 6). These two species are known to segregate on this pH gradient, with C. oligosperma favoring more acidic fens than C. utriculata (Gignac et al. 2004; Dabros and Waterway 2008; Table 3). Several of the acidic fens have relatively narrow pH ranges, as do several of the fens with pH means above 6 (Fig. 6), making it likely that the strong pattern of aversion shown in 21 of the 25 possible species pairs that involve C. oligosperma result from filtering on the gradient in pH among regional fens that precludes C. oligosperma colonizing the full range of fen habitats. Carex oligosperma also shows a predominant pattern of significant aversion with other species within fens, but the number of pairings with C. oligosperma that show such aversion within fens is reduced from 21 to only nine, with most of its other pairings neutral within fens. Hence this species pair is also consistent with the expectations from limiting similarity, or more appropriately from the obverse perspective of ecological character displacement (Stuart and Losos 2013) expressed in their differing pH tolerances. The two species tend to occur at different extremes along the pH gradient among the fens in the region; in the infrequent cases where they both occur in an intermediate fen along the gradient, they occupy different pH microenvironments within that fen. Similarly, in a study of co-occurrence of two sister species in Carex section Racemosae that both occur in high elevation alpine tundra in the Rocky Mountains, Massatti and Knowles (2014) found microhabitat differentiation and thus pairwise aversion between these two closely related species based on different ecological tolerances along gradients in the frequency of water-saturated soils.

A similar situation prevails when using the alternative definition of D level clade divergence (D2), in which the entire Wetland clade is considered to be closely related. This increases the number of possible D2 pairings in Core Carex to 29, 26 of which have sample sizes large enough to potentially show significance. Sixteen and 15 of these 26 pairings show significant aversion in the regional and fen-by-fen
analyses, respectively, and all but three or four of the remaining pairings are neutral (Fig. 4A; Table 4). The _C. aquatilis_ – _C. magellanica_ pairing is significant in the regional analysis but not in the analysis within fens, probably for the same reason noted above; _C. magellanica_, like _C. oligosperma_, grows more commonly in fens with more acidic pH ranges that lack many of the other species. On the other hand, _C. limosa_ is significantly less likely to occur in the same quadrant as _C. oligosperma_ within fens, but not in the regional analysis pooling data from all the fens. Overall, it is apparent from the co-occurrence analyses that closely related species in the Core _Carex_ clade, following either the D1 or D2 definition of clade divergence level, are not very likely to co-occur as immediate neighbors in fens of the Schefferville region. At the same time, the level of fine-scale heterogeneity in water pH and rooting level within fens is high enough that species in these pairs can both grow within many of the same fens (cf. Figure 8), although not often as immediate neighbors.

In contrast to the pattern in Core _Carex_, D-level species pairs in the _Vignea_ clade (that is, within subclades representing sections _Glareosae_ and _Stellulatae_) show neither significant attraction nor significant aversion in either the regional analysis or the analysis within fens, with the exception of _C. canescens_ and _C. tenuiflora_, which tend to occur more often than expected by chance in both types of analyses. _Carex canescens_ and _C. tenuiflora_ are not sister species, but we still consider them closely related because they belong to sister clades of six and eleven species, respectively, within the larger _Glareosae_ clade, a group in which all of the species are quite similar in morphology and branch lengths are short (Maguilla et al. 2015). This pairwise attraction between two species in the group is contrary to expectations from the phylogenetic limiting similarity hypothesis but akin to the finding of different patterns of co-occurrence between the Core _Carex_ clade and the Core _Vignea_ clade in another community phylogenetic study of sedge species in the Schefferville region (Elliott et al. 2016). However, in that study, Core _Carex_ tended to exhibit clustering while the _Vignea_ clade tended toward overdispersion. That study sampled a broader spectrum of habitats, including ponds, lake and river shores, dry ridges, tundra, and wet to dry forests, so the outcome of the community phylogenetic analysis may have been driven by abiotic filtering effects. In any case, it is clear from both studies that the two major clades differ in the degree to which their species can co-occur, but further research is needed to understand the exact nature and functional basis for this phylogenetic contrast.

In summary, the results of the co-occurrence analyses reject the null hypothesis that closely related species co-occur at random and agree with the expectation that closely related species in the same habitat will have diverged enough in their ecological tolerances that they are unlikely to co-occur as immediate neighbors. Results from the linear modeling also support this expectation. The best model to predict pairwise co-occurrence included the level of clade divergence, as well as both pH and rooting level in relation to the water table (Fig. 5; Table 4). Significant aversion was predicted by the model for pairings of the most closely related species (D-level divergence) and _p_ values for divergence between subclades (C-level divergence) were 0.15 and 0.63 for the CLD1 and CLD2 analyses, respectively, suggesting that character displacement is limited to closely related species. The fact that both environmental factors were also significant further supports the idea that lineage diversification within the Wetland clade and the _Panicaceae_ clade has involved ecological speciation events. The BEAST analysis suggests that diversification in these flood-tolerant lineages started as early as the late Miocene and continued during the cooler and drier Pliocene and the glacial cycles of the Pleistocene; be that as it may, the species that colonized the Schefferville region in the 6,500 yr since the retreat of the Wisconsinan ice (Jansson 2003; Dyke 2004) have undoubtedly been subject to repeated cycles of glaciation throughout the Quaternary. The repeated alternation between periods of isolation in glacial refugia associated with different parent materials and recolonization of the sodden landscapes that accompanied the glacial retreats (Alsos et al. 2015; Gajewski 2015) would have provided ample opportunities for adaptive evolution and functional diversification relevant to present patterns of co-occurrence in cool wetland and peatland habitats.

**Species Attraction**—Expectations for significant attraction of species pairs in relation to their phylogenetic relationships are less clear. The phylogenetic limiting similarity hypothesis leaves open the question of a threshold at which relatedness no longer influences the likelihood of co-occurrence. Certainly there could be a threshold beyond which there is no longer aversion to co-occurrence, but conversely no necessity that more distantly related sedge species should have evolved characteristics that increase their chances of co-occurrence. Nevertheless, from coexistence theory (Chesson 2000; HilleRisLambers et al. 2012) we can expect that immediately neighboring individuals of different species should exploit resources in different ways, thus avoiding competition. Because of the high degree of horizontal and vertical heterogeneity in the fen environment (Rochefort et al. 2012; Macrae et al. 2013; Ulanowski and Branfireun 2013), there is ample opportunity for such functional segregation even between immediately neighboring plants. For example, the availability of N and P as the water table fluctuates seasonally is lower for plants rooted on a hummock as opposed to in an adjacent hollow (Rochefort et al. 2012; Macrae et al. 2013), but so is the risk of exposure to anaerobic conditions when water levels are high (Visser et al. 2000). Competitive interactions even for neighboring plants rooted at the same level can also be reduced by functional differences; e.g. _Eriophorum vaginatum_ and _Carex aquatilis_ take up nitrogen in different forms, a stabilizing niche difference (Chapin et al. 1993; Jonasson and Shaver 1999). Species from different lineages might be more likely to co-occur because they have evolved such differences or species may co-occur if one facilitates the growth of another in some way.

In this study, the most frequent set of species pairs that exhibit significant attraction root at or below the water table in areas with moderate to high pH. These species include _Carex limosa_, _C. livida_, _C. chordorrhiza_, _T. cespitosum_, _T. alpinum_, _E. chamissonis_, and _E. viridicarinatum_, all of which exhibit mutual pairwise attraction except the _C. chordorrhiza_ – _E. chamissonis_ pair (Fig. 4). Consistent with expectations for significant attraction, these seven species represent diverse lineages within the CDS clade: two Core _Carex_ species are from distantly related subclades, one species from the _Vignea_ clade, and two from each of the other genera. The two _Trichophorum_ species each represent one of the two major clades within that genus (Dhooge 2005), and the two _Eriophorum_ are from different subclades, both in our
phylogenetic tree and in a larger study of the CDS clade (Léveillé-Bourret et al. 2014).

In line with the general expectations about functional diversity described above, this set of species that frequently co-occur at fine scales shows clear differences in morphology that are likely to reflect different nutrient acquisition and space-filling strategies. For example, the three midsize species have a spreading habit but differ in how they spread as well as in their root structure and position relative to the water table. Both C. limosa and C. chordorrhiza spread by laying down their aerial stems at the end of each season and sending up new aerial stems from the nodes in the following year. The roots are relatively short and near the surface of the peat in both species, but C. limosa has abundant matted root hairs while C. chordorrhiza has smoother roots likely to differ in nutrient uptake capacity (Gerke 2015). The strategy of C. livida is the most different of the three; it has deep, long-spreading narrow rhizomes, and white, smooth, extremely long roots that extend several cm below the water surface. Carex livida is also capable of forming dauciform roots, a type of cluster root that facilitates phosphorus uptake and is induced by low phosphorus levels (Gerke 2015), while C. chordorrhiza is not (Bérubé, Waterway and Lechowicz, unpublished data).

The Trichophorum and Eriophorum species also exhibit disparate functional traits that may facilitate their co-existence in this network of co-occurring pairs. For example, Trichophorum species are also mid-sized, but in contrast to the Carex species in this network, which spread out radially, they grow in cespitose clumps: smaller looser clumps in T. alpinum, and very large dense clumps in T. cespitosum. Not much is known about either of their nutrient acquisition strategies, but there are examples of other cespitose sedge species that can recycle nutrients within the clump or release phosphorus to the zone of active roots (Jonasson and Chapin 1991). Carex limosa and C. chordorrhiza may even facilitate the establishment of the Trichophorum species; mosses colonize the horizontal stems and create floating mats on which Trichophorum species are frequently found. The two Eriophorum species in this network of co-occurring pairs are both larger in stature and spread with long, relatively narrow rhizomes that connect single culms or small clumps. They tend to spread quickly with widely spaced culms so individual genets can conceivably occupy a large area for nutrient capture. Carex rostrata, which is also large in stature and spreads by both short and rapidly growing, long, thick rhizomes that can grow quite deep due to their high proportion of aerenchyma for oxygen transport (Visser et al. 2000), is found with the other species in the network only when the two Eriophorum species, which are of similar size and growth form, are absent.

Similarly, hummock microhabitats are also characterized by the co-occurrence of species representing distantly related lineages within the CDS clade of Cyperaceae, although the patterns of mutual pairwise attraction are not as strong as those described above for the wetter microhabitats like flarks and pools. Hummocks vary in pH depending on the water pH where they are found and the particular assemblages of bryophyte species that form the hummock; those richer in species of Amblystegiaceae (brown mosses) tend to be found at higher pH than those dominated by Sphagnum species. Sedges that prefer rooting above the water level and are generalists with regard to water pH show positive co-occurrence relationships to specialist sedge species that occur at both ends of the pH gradient, but these specialists do not co-occur with each other. For example, the generalists C. aquatilis (Core Carex) and C. gynocrates (Vignea) show significant attraction to C. pauciflora (Caricoid) and C. echinata (Vignea) at the lower end of the pH gradient and with C. vaginata (Core Carex), C. leptalea (Caricoid), and T. alpinum at the higher end, but the two species from the Caricoid clade have different pH preferences and show significant aversion, and most of the other pairwise comparisons between hummock species with different pH preferences are neutral (Fig. 4). Many of the species that grow mostly or exclusively on hummocks are relatively short in stature, and the rhizomes of hummock species vary from threadlike in C. leptalea, C. gynocrates, and C. pauciflora, to much more robust in larger species like C. aquatilis, in which single individuals can spread across hummocks and hollows, or cross between the strings and flarks in patterned fens.

**Effects of Scale**—A key consideration in this discussion has been the effect of spatial and phylogenetic scale in studies of community assembly. We have argued that ecophylogenetic insights into the process of community assembly are strengthened by assessing interactions among congeners at the neighborhood scale (Stoll and Weiner 2000), and in particular through the consideration of pairwise comparisons between immediate neighbors that can help elucidate the functional basis for their coexistence. Defining the spatial scaling of a neighborhood relevant to pairwise tests of coexistence, however, is not entirely straightforward. Sessile organisms such as plants invite the use of either focal sampling centered on a specific individual (Ricotta et al. 2015; Elliott et al. 2016) or the use of suitably small, randomly placed quadrats as in this study. Such fine-scale sampling designed to discern interactions between individual plants is in contrast to traditional approaches in plant community ecology and community phylogenetics, which emphasize the use of quadrats sufficiently large to provide a “representative sample” of the community per se. We are aware of only one previous study (Slingsby and Verboom 2006) that has used a pairwise analysis characterized as ‘fine-scale’ to study community assembly in congeneric plant species.

Slingsby and Verboom (2006) analyzed the patterns of co-occurrence among species of Tetraria, another genus with high local diversity (Slingsby et al. 2014), and four co-distributed allied genera in tribe Schoeneae (Cyperaceae). They drew on vegetation data from 921 50 m² plots representing different plant community types at 11 sites from across the Cape Floristic Region (cf. McDonald 1993a, b); plots contained from two to 11 (median 5) Tetraria species and from three to 17 (median 10) schoenoid sedges in total (cf. Supplemental Table A2 in Slingsby and Verboom 2006). They considered these 50 m² plots to be a suitable scale at which to evaluate co-occurrence because at this scale there was a “visually assessed homogeneity of vegetation structure and habitat” (Wenger et al. 1972), although they acknowledge that this spatial scale does not directly capture interactions among neighboring individual plants. In an analysis of pairwise co-occurrence in these 50 m² plots, they detected evidence of significant pairwise aversion among the 15 species of the reticulate-sheathed Tetraria clade, but not at a broader phylogenetic scale that included other species of Tetraria and allied genera. Given the considerable diversity of plant community types sampled (McDonald 1993a, b)
as well as the likely environmental heterogeneity within the 50 m² plots, these results may well have arisen primarily through habitat filtering rather than biotic interactions. It would be informative to carry out a study of community assembly in the fynbos region using a focal sampling approach at a neighborhood spatial scale.

Conversely, we can easily collapse our neighborhood scale data to the level of individual fens as discrete samples of a habitat type on the regional landscape. We have reanalyzed our data to consider co-occurrence at the fen level rather than the quadrat level, that is, to consider what Silvertown et al. (2006) called the β-niche or habitat niche. Because subarctic fens are discrete habitats with easily defined boundaries and environmental conditions quite different from their surroundings as well as being inhospitable to most plant species, it is therefore not surprising that in this analysis, in which each of the 29 fens was considered as a single observation of sedge species co-occurrence, most species pairs were either attractive (56 pairs) or randomly associated (293 pairs) while only two pairs were aversive (unpublished results). Both aversive pairs were also aversive in our quadrat level analyses, but most other pairs that were aversive in the quadrat level analyses showed no significantly aversive or attractive tendencies at the fen level. Two species pairs (C. limosa – C. magellanica and C. limosa – C. oligosperma), showed significant attraction at the fen level, indicating that they were frequently found in the same fens, although they only rarely co-occurred at the quadrat level. These results underscore how critical it is to consider spatial scale in evaluating patterns of co-occurrence.

The common coexistence of many sedge species in the Schefferville region in the same fens (cf. Table 3) supports habitat filtering in the regional species pool, yet at the same time, most of the closely related species that occupy these fen habitats have diverged on one or more of the important environmental gradients that structure fen communities, and as a result, they do not co-occur at a fine scale within the fens. This suggests that niche differentiation has accompanied the process of lineage divergence, which is consistent with the expectation of Silvertown et al. (2006) that β niches are more conservative than α niches and that traits that allow plants to exploit particular habitats, such as subarctic fens, trace to earlier branching events in the phylogenetic tree. The Wetland clade in Core Carex appears to have diverged between 7.5 (plus or minus 2.5) and 12 (plus or minus 3) million years ago, depending on the calibration used (Fig. 3, Supplemental Figs. 1, 2, available from the Dryad Data Repository at http://dx.doi.org/10.5061/dryad.q810f). Nearly all species in that clade, worldwide, grow in wetland habitats that require some degree of flooding tolerance suggesting β-niche conservatism. Evolution of adaptations to acquire nutrients at different water levels and at different water pH levels appears to have occurred frequently in the Wetland clade, with diversification in actual rooting levels and pH tolerances varying at the tips of the phylogenetic tree and allowing coexistence of species with these different α-niches within a given fen.

In summary, working at fine spatial and phylogenetic scales, we have detected pairwise co-occurrence patterns suggesting that ecological speciation events at the tips of the phylogenetic tree allow coexistence of several species in the species-rich CDS clade of Cyperaceae within subarctic fens. We suggest that such coexistence within these harsh habitats, characterized by fine scale environmental heterogeneity, results from niche differentiation among closely related species along ecological gradients of rooting level in relation to the water table and water pH such that closely related species are significantly less likely to co-occur in the same microhabitats within the fens. We have also shown that mutually co-occurring species that show significant pairwise attraction also exhibit diverse strategies for acquiring nutrients and colonizing wetter microhabitats within the fen, thus allowing stable coexistence. Ecological tolerances within species along the pH gradient ranged from moderate to broad, but all species showed high levels of flexibility evidenced by wide variance among means across the 29 sampled fens. Ranges of rooting level were much narrower. Studies on mechanisms of nutrient capture in these species, and more detailed exploration of the evolution of traits that allow coexistence at the fen scale as well as co-occurrence at the fine scale are needed to further characterize these stabilizing niche differences. Finally, working with clade divergence levels rather than phylogenetic distances allowed us to evaluate co-occurrence patterns at different levels of the phylogenetic tree and to uncover contrasting patterns between the Vignea and Core Carex clades at the same level of divergence. It is notable that linear modeling consistently showed that clade divergence level was significant in predicting co-occurrence whereas phylogenetic distance was not, even when it was the only measure of phylogenetic relationship in the model. A likely explanation for this result is that phylogenetic distance metrics inappropriately treated all branches as if they have equal support whereas only clades with significant levels of support were used to determine clade divergence level. As phylogenetic hypotheses for Carex become more complete and well-resolved, there will be many opportunities to assess the generality of the results described here by comparing different lineages within and between habitats at broader geographic scales.

Acknowledgments. We thank Rachel Anderson Toews, Nadia Cavallin, and Chantal Duval for field assistance; Oksana Chouluk for logistic support at the McGill Subarcic Research Station in Schefferville; Christina Cho, Yaraslava Chtompel, Sofia Fuga, Anita Rogic, and Nadia Surdek, for assistance with DNA extraction and amplification, the McGill University & Genome Quebec Innovation Centre for DNA sequencing, Marcial Escudero for assistance with the BEAST analyses and advice about fossil calibration, Vahid Gazestani and Elijah Perez for help with data management, A.I. Luz for preparing Fig. 1, and Will Pearse for helpful discussions about the ecophylogenetic literature. We also thank Andrew Hipp for organizing the symposium, “Ecological diversification and niche evolution in the temperate zone’s largest genus: Carex,” in which the work in this paper was originally presented at Botany 2015, the American Society of Plant Taxonomists for funding the symposium, James Smith for managing publication of the symposium papers, and two anonymous reviewers for useful suggestions that improved the manuscript. The research was funded by an NSERC fellowship to AD, NSERC Discovery Grants to MJL and MJW, and Northern Scientific Training grants from the Department of Indian and Northern Development to AD and her field assistants.

Literature Cited


APPENDIX 1. Voucher information and GenBank accession numbers for sequence data used in this study. Taxa are ordered by genus, major clade within the genus, sectional classification, and then alphabetically by species. Each entry is in the following order: Taxon name, collection clade within the genus, sectional classification, and then alphabetically by species. Each entry is in the following order: Taxon name, collection locality, collector, collection number, voucher location with herbarium species. Each entry is in the following order: Taxon name, collection number, voucher location with herbarium species. Each entry is in the following order: Taxon name, collection number, voucher location with herbarium species. Each entry is in the following order: Taxon name, collection number, voucher location with herbarium species. Each entry is in the following order: Taxon name, collection number, voucher location with herbarium species. Each entry is in the following order: Taxon name, collection number, voucher location with herbarium species. Each entry is in the following order: Taxon name, collection number, voucher location with herbarium species. Each entry is in the following order: Taxon name, collection number, voucher location with herbarium species. Each entry is in the following order: Taxon name, collection number, voucher location with herbarium species. Each entry is in the following order: Taxon name, collection number, voucher location with herbarium species. Each entry is in the following order: Taxon name, collection number, voucher location with herbarium species. Each entry is in the following order: Taxon name, collection number, voucher location with herbarium species. Each entry is in the following order: Taxon name, collection number, voucher location with herbarium species.