CrossMark

ORIGINAL PAPER

DNA barcoding of aquatic insects reveals unforeseen diversity and recurrent population divergence patterns through broad-scale sampling in northern Canada

Ruben D. Cordero[1] · Santiago Sánchez-Ramírez[1,2] · Douglas C. Currie[1,3]

Abstract Arctic and subarctic environments are among the most inaccessible regions in the world, making biological surveys difficult to conduct. Thus, the insect fauna of these regions has remained inadequately surveyed. The aquatic insect orders Ephemeroptera, Plecoptera, and Trichoptera (EPTs) are particularly abundant and diverse at high latitudes, playing key roles in trophic chains where nutrients are scarce. However, particular aspects of their life cycle make them difficult to study. Specifically, species-level identification requires last-instar larvae or adults which, because they are short lived, are typically not available for all taxa during a particular collecting event. With the initial goal of surveying the biodiversity of these insect groups, we sampled ca.10,000 EPT specimens from 12 locations across northern Canada over two years. Approximately 800 of these were subsequently selected for COI DNA barcoding. Overall, we identified 155 EPT species (58 Ephemeroptera, 41 Plecoptera, 56 Trichoptera) based on a 2% divergence criterion. Compared to other similar studies on EPTs we found higher (particularly among the Plecoptera and Ephemeroptera) and more even diversity, potentially reflecting environmental differences in sampling localities. We further assessed phylogeographic divergence patterns among seven species, finding that eastern and western populations diverged during the Pleistocene Epoch (<2.5 Ma), with overlapping time frames. This finding highlights the role of potential glacial refugia and subsequent recolonization, as well as the dispersal potential of some EPT species. This study exemplifies how large-scale DNA-based surveys can be combined with phylogeographic inference to better understand the biodiversity and natural history of northern aquatic insect communities.

Keywords Aquatic insects · EPT diversity · DNA barcoding · Pleistocene · Molecular clock · Population divergence

Ruben D. Cordero and Santiago Sánchez-Ramírez have contributed equally to this work.

✉ Ruben D. Cordero
rubendario1939@gmail.com

1   Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks, Toronto, ON M5S 3B2, Canada

2   Environmental Genomics Group, Max Planck Institute for Evolutionary Biology, 24306 Plön, Germany

3   Department of Natural History, Royal Ontario Museum, 100 Queen's Park, Toronto, ON M5S 2C6, Canada

Introduction

Northern Canada is among of the most remote and isolated areas on Earth, making biological surveys costly and difficult. For insects, other complications involve short sampling windows (e.g., short summers) that reduce the chance of obtaining specimens in the appropriate life-history stage (especially adults) for species-level identification. Consequently, important components of northern biodiversity remain poorly studied, particularly in highly diverse organisms such as aquatic insects.

Mayflies (Ephemeroptera), stoneflies (Plecoptera), and caddisflies (Trichoptera)—the EPTs—along with Diptera, are the most prominent insect orders in northern aquatic ecosystems (Merritt et al. 2008). Collectively, these insects are a key component of energy flow in aquatic systems, as they constitute an important link between the algal/detrital

food base and higher trophic levels, including the fishes that feed upon them (Benke et al. 1984). EPTs are widely distributed in the northern Nearctic Region, with representatives occurring on most islands of the Canadian Arctic Archipelago, which makes them a suitable group for large-scale biodiversity assessments of that region.

DNA barcoding has received much attention for its potential for a rapid and relatively unbiased prospecting of biodiversity (Hebert and Gregory 2005). In addition, the approach has proven particularly useful for species-level identification in insects, regardless of gender (Ekrem et al. 2010) or developmental stage (Gullan and Cranston 2010; Zhou et al. 2009, 2010). The mitochondrial gene cytochrome oxidase subunit I (COI) is the genetic marker of choice for animals due to the benefits of working "universal" primers, length, alignment ease, and degree of intra- and intergenetic polymorphisms (Hebert et al. 2003). In addition to its utility for distinguishing known species, the COI region has also been found useful for revealing cryptic diversity (Hebert et al. 2004; Smith et al. 2008; Ståhls and Savolainen 2008; Monaghan et al. 2009; Velonà et al. 2015), as well as biogeographic and phylogeographic patterns (Pons et al. 2006; Hajibabaei et al. 2007; Craft et al. 2010). Thus, a barcoding approach promises to be a useful tool both for conducting large-scale biodiversity assessments, and exploring recent population divergence patterns (e.g., during Pleistocene glacial cycles).

Recent important works using the COI barcode approach have increased knowledge of biodiversity in various orders of aquatic insects in northern North America (Ball et al. 2005; Webb et al. 2012) and northern Europe (Boumans and Brittain 2012; Kjaerstad et al. 2012). In particular, notable are COI barcoding studies of EPTs from the vicinity of Churchill, Manitoba, conducted over a 6-year period, which established a baseline for biodiversity assessments in other parts of northern Canada (Zhou et al. 2009, 2010; Ruiter et al. 2013). While it is clear that the northern EPT fauna is much richer than previously supposed, these efforts were concentrated in a relatively small geographic area, leaving the vast majority of northern Canada poorly surveyed.

This study aims to redress this deficiency by undertaking the first broad-scale study of EPTs in northern Canada. The specific objectives of this study are to: (1) assess EPT diversity across multiple, widely separated locations using a COI barcoding approach, (2) compare species-specific cases of recent population-level divergence using a molecular clock approach, and (3) determine whether such genetic divergences are related to past climatic events. This is the first broad-scale study of EPTs in northern Canada that uses a DNA barcoding approach to assess species richness and compare patterns of intraspecific divergence.

Materials and methods

Fieldwork

Specimens of three orders of aquatic insects (EPT) were collected from 12 widely distributed sites in northern Canada during June and July of 2010 and 2011. Four sites were distributed among each of three northern ecoclimatic zones, as follows: Boreal—Goose Bay (NL), Moosonee (ON), Yellowknife (NT), Norman Wells (NWT); Subarctic—Schefferville (QC), Churchill (MB), Kugluktuk (NU), Ogilvie Mountains (YT); and Arctic—Iqaluit (NU), Lake Hazen (NU), Cambridge Bay (NU), and Banks Island (NT) (Fig. 1). Sampling took place for approximately two weeks at each site, with boreal sites sampled earliest in the season (followed, respectively, by subarctic and arctic sites) to ensure that insects were sampled at more-or-less the same stage of development in each ecoclimatic zone. Easternmost sites were sampled in 2010 and westernmost sites in 2011.

Standardized sampling protocols were implemented on an equal number of lentic (standing water) and lotic (running water) habitats at each site. Each habitat was sampled using a D-framed aquatic dip net for 20 min to collect the immature (i.e., aquatic) stages of EPTs. Net contents were emptied into shallow pans of water, with larvae of target taxa removed using fine forceps. A further 15 min was spent sweeping the riparian vegetation surrounding each habitat to collect adult EPTs. Aquatic and terrestrial specimens were fixed mainly in 95% ethanol to facilitate both morphological identification and DNA barcoding. However, in instances when ethanol supplies became exhausted toward the end of a sampling period at some locations, material was fixed in 80% isopropyl alcohol or propylene glycol, and later transferred to 95% ethanol. Vouchers are deposited in the Entomological collection at the Royal Ontario Museum, Toronto, Canada.

Laboratory procedures

Approximately 10,000 EPT specimens were collected and sorted under magnification with a Nikon stereomicroscope. Specimens were initially identified to the lowest possible taxonomic level following taxonomic keys in Merritt et al. (2008), Wiggins (1996) and Ruiter (1995). Selected adults and mature larvae were sent to taxonomic specialists for confirmation of species-level identifications: Ephemeroptera (S. Burian, Southern Connecticut State University), Plecoptera (R. Baumann, Brigham Young University and B. Kondratieff, Colorado State University), and Trichoptera (D. Ruiter, Grants Pass, Oregon). Finally, 800 specimens—700 larvae and 100 adults—representing all the morphospecies

Fig. 1 Sampled sites across the three main ecoclimatic region (arctic, subarctic and boreal) in northern Canada. *Dark circles* eastern sites: Goose Bay (GB), Iqaluit (IQ), Lake Hazen (LH), Moosonee (MO) and Schefferville (SV). *White circles* western sites: Banks Island (BI), Churchill (CH), Cambridge Bay (CB), Kugluktuk (KT), Norman Wells (NW), Ogilvie Mountains (OM) and Yellowknife (YK)
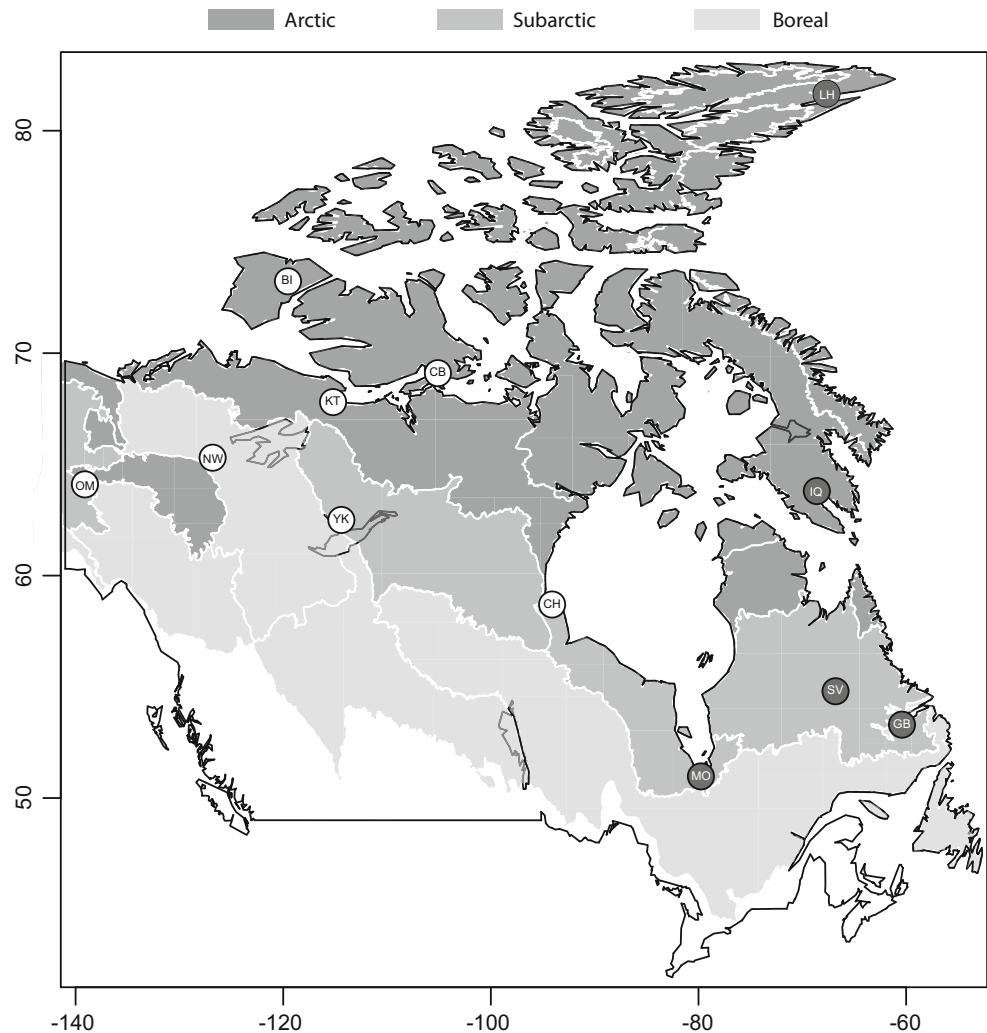
identified by us, were selected for DNA extraction and COI barcode sequencing. Specimens from Churchill, Manitoba, were not barcoded because all morphospecies collected by us corresponded with those already identified and barcoded from that site by Zhou et al. (2009, 2010). Accordingly, sequence data from those studies were used as a proxy for the EPTs collected in the present study.

Genomic DNA was extracted from the hind leg of selected specimens using the DNeasy Blood & Tissue Kit (Qiagen, Inc.). Extracted DNA was eluted in 150 μl of elution buffer AE (10 mM Tris–HCl, 0.5 mM EDTA, pH 9.0). We amplified the cytochrome oxidase I (COI) gene for all 800 specimens using the primer set LCO1490 (5′-ATTCAACCAATCATAAAGATATTGG-3′)—HCO2198 (5′-TAAACTTCAGGGTGACCAAAAAATCA-3′) following the methods described by de Waard et al. (2008) and Hajibabaei et al. (2005). PCR products were visualized and excised from a 1% agarose gel, followed by, spin column centrifugation for PCR product clean-up, and Sanger sequencing using the BigDye Terminator kit v3.1.

All sequencing reactions were performed using a Hitachi Applied Biosystems 3730 DNA Analyzer.

Genetic analyses

Obtained sequences were individually compared against the NCBI GenBank and BOLD systems v3 database using BLASTn to identify the species name using a 98% similarity cutoff. A sequence threshold of 2% was chosen to delimit haplotype clusters to compare with previous studies of northern EPT's by Zhou et al. (2009, 2010). For sequences with greater than 2% divergence, voucher specimens were sent to taxonomic specialists of each group to confirm identification. In order to improve geographic sampling, we included COI sequences for 107 species produced by Zhou et al. (2009, 2010) from Churchill (Fig. 1; Online Resource 1). Sequences produced in this study were deposited in GenBank (NCBI) under accession numbers KJ674822—KJ675402, and the information of each sequence is provided in Online Resource 2.

Sequences were grouped according to order and then aligned and edited using Geneious® 6.1.6 (Biomatters Inc.) and ClustalX (Higgins and Sharp 1988) with default parameters. The best substitution model for each alignment was obtained using jModelTest v2 (Darriba et al. 2012) and the Akaike Information Criterion (AIC). Phylogenetic reconstructions for each insect order were obtained using BEAST v.1.7.5 (Drummond et al. 2012), by selecting a constant coalescent tree prior and a strict clock model with the rate fixed to 1.0, in order to estimate relative divergence times. Twenty million MCMC generations were run, sampling every 2000 states, to produce log and tree files with 10.000 states each. Convergence and chain mixing were assessed by visually inspecting each parameter trace from the log files using Tracer v.1.5 (Drummond and Rambaut 2007) considering effective sample sizes above 200 as a good indicator. After specifying a burnin of 10%, a maximum clade credibility tree (MCC) with mean heights was produced using TreeAnnotator V.1.7.5. MCC trees were visualized in FigTree v.1.4 to identify widespread species (Suppl. Figs. S1–3). Once species were identified, individual alignments were re-analyzed using the same procedure as above, but with an MCMC chain of 5 million generations, sampling every 500 states. MCC trees were observed and edited with the package APE (Paradis et al. 2004) for R v3.0.1 (R Core Team 2015). The time to the most recent common ancestor (TMRCA) was rescaled post-analysis based on a 3.54% divergence rate per Myr (0.0177 substitutions per site per Myr) proposed for insect COI sequences (Papadopoulou et al. 2010).

Statistical analyses

To compare the TMRCA for eastern and western populations of selected widespread species, the *treeModel.-rootHeight* parameter was extracted from each species BEAST log file in R, discarding the initial 10% of the MCMC samples as burnin. We estimated the amount of overlapping time between the TMRCA of each species measuring the distance between means D with the formulae:

$$D = \frac{\mu_1 - \mu_2}{\sigma_1}$$

where μ_1 is the mean TMRCA of the "reference" species, μ_2 is the mean TMRCA of the "compared" species, and the σ_1 standard deviation of the "reference" species. With D, the proportion of overlap p can be calculated using a normal density function P (pnorm() function in R):

$$p = 2P\left[X \le \left(\frac{|D|}{2}\right)\right]$$

where X is a random variable from a normal distribution. Since the distribution of the MCMC TMRCA estimates was not symmetrical, we estimated p on each side of the distribution for each pair of species.

To account for the effect of randomness in the observed distribution of subpopulations, we performed a permutation test with 10,000 iterations. The test consisted of randomizing the distribution of all terminals into "East" and "West" and counting the number of times the observed distribution of subpopulations h_o was found for each iteration. To obtain a p value, we used the following formula,

$$p = \frac{\sum_{i=1}^{n} ri}{n} \quad \text{if } r_i = h_o = 1$$

where r_i represents the number of times h_o was found in iterations. Permutation analyses and the statistical figures were made in R Core Team (2015).

Results

COI barcode subunit 1 sequences ranging between 650 and 660 bp were recovered for 587 individuals (73%) from 10 sites of 12 initial sites. Barcode data for Churchill, Manitoba, EPTs were derived from Zhou et al. (2009, 2010). Only one EPT morphospecies—an immature trichopteran of the family Limnephilidae —was collected from Lake Hazen on Ellesmere Island; however, we were unable to obtain sequences for that species.

Sequences were obtained from 87% of Ephemeroptera, 83% of Plecoptera and 71% of Trichoptera specimens sampled. Of these, a total of 155 species were identified based on BLAST using a 2% divergence cutoff. The number of species by order was: Ephemeroptera (58 spp.), Plecoptera (41 spp.), and Trichoptera (56 spp.) (Figure 2).

The best substitution model obtained for each order-level alignment was the generalized time reversible (GTR, Tavaré 1986). Nineteen species from MCC trees (Online Resources 3–5) had populations distributed on both east and west of Hudson Bay. Of these, only seven exhibited both geographic and genetic structure (Fig. 3): Ephemeroptera (*Acentrella lapponica, Ameletus inopinatus, Baetis tricaudatus* and *Ephemerella aurivillii)*, Plecoptera (*Skwala compacta*), and Trichoptera (*Ceratopsyche alternans* and *Onocosmoecus unicolor*). Nine other species exhibited marked genetic divergence, but without clear geographic structure: Ephemeroptera (*Heptagenia pulla, Epeorus vitreus, Leptophlebia cupida, Baetis hudsonicus, Baetis brunneicolor*) (ESM 3), Plecoptera (*Nemoura arctica, Isoperla transmarina*) (ESM 4), and Trichoptera (*Limnephilus picturatus*) (ESM 5). Finally, three other widely distributed species—all trichopterans— exhibited no clear genetic or geographic structure (*Asynarchus lapponica, Anabolia bimaculata, Grensia praeterita*) (ESM 5).
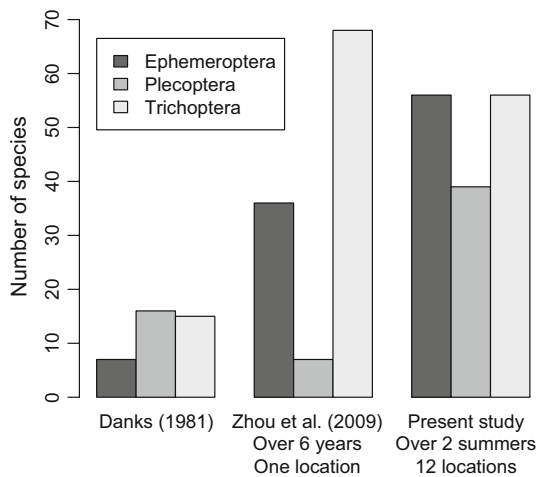
Fig. 2 Barplot comparing EPT diversity between three studies in nothern Canada; Danks (1981) and Zhou et al. (2009) and the present study

The best substitution models for each of the species-level alignments were: the GTR model for *A. inopinatus* (–log$L$ 1071.1 *SE* 0.102), *C. alternans* (–log$L$ 869.4 *SE* 0.106), *E. aurivillii* (–log$L$ 975.3 *SE* 0.084), and *O. unicolor* (–log$L$ 1141.1 *SE* 0.1639); and the HKY (Hasegawa et al. 1985) model for *B. tricaudatus* (–log$L$ 1119.5 *SE* 0.081), *A. lapponica* (–log$L$ 887.5 *SE* 0.091) and *S. compacta* (–log$L$ 1004.7 *SE* 0.077). Posterior probabilities (clade support) for each subpopulation (i.e., east versus west) ranged from 0.94 to 1.0 except for two species: *Ameletus inopinatus* and *Ephemerella aurivillii*. In these cases, western subpopulations had values of 0.62 and 0.64, respectively, for each species (Fig. 3). The interpopulation divergence percentage, measured from branch edges (with length in substitutions per site) connecting both subpopulations, was 0.4% for *E. aurivillii*, 0.6% for *Ceratopsyche*
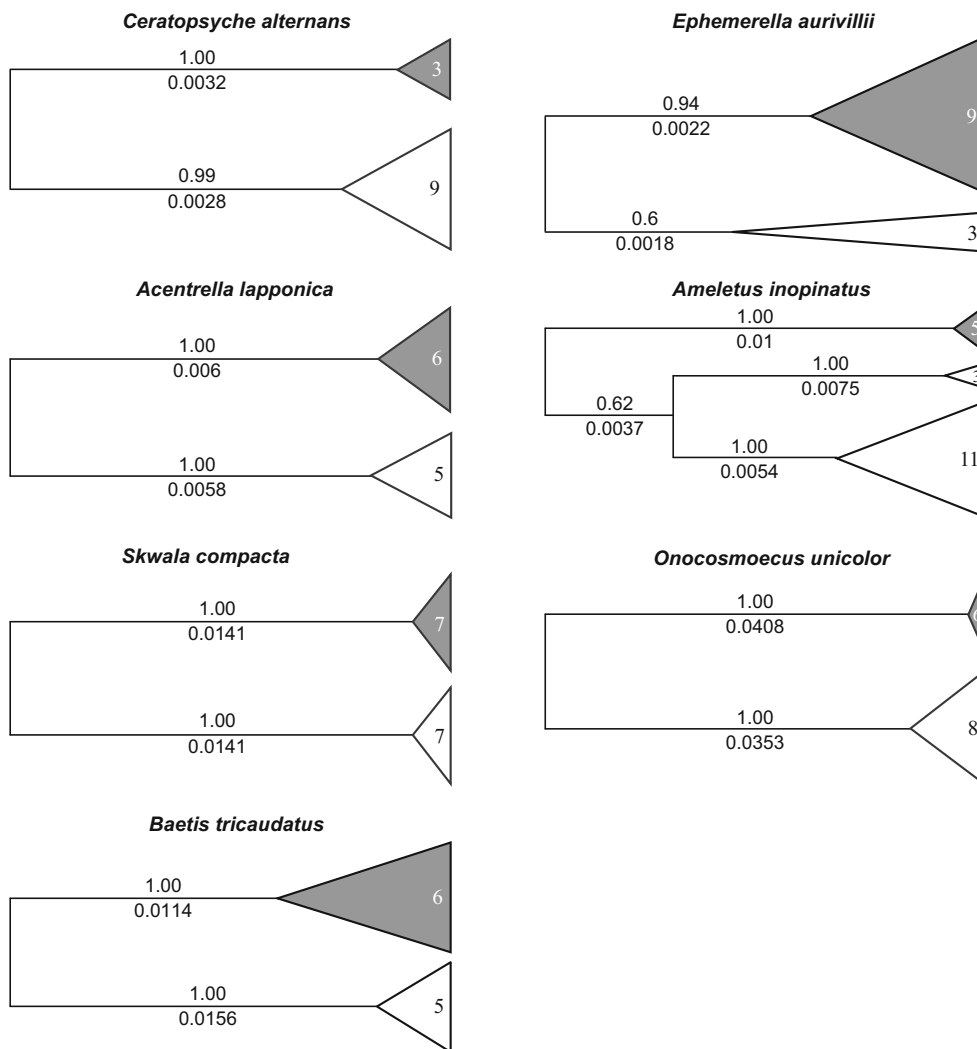

Fig. 3 Bayesian maximum clade credibility trees of seven EPT species present with divergence eastern and western populations in northern Canada. *Numbers above* branches indicate the posterior probability (clade support). *Numbers below* branches indicate the genetic distance (in substitutions per site) between populations. *Numbers within triangles* indicate individuals. *White triangles* represent collapsed lineages of eastern populations, while *gray triangles* those of western populations
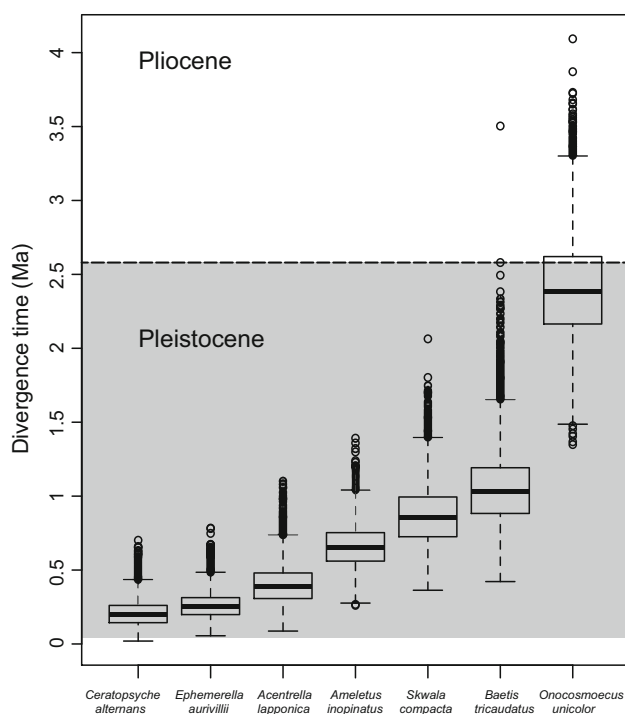
Fig. 4 Time to the most recent common ancestor (TMRCA) of eastern and western populations of seven EPT species widely distributed in northern Canada. Species populations show divergence times during the Pleistocene within a close range except for *Onocosmoecus unicolor*. A total of 10,000 iterations were performed to account for the effect of randomness in the observed distribution of subpopulations of all terminals into "East" and "West." *Box plot* and *black circles* represent 95 and 5% of the iteration results, respectively

alternans, 1.18% for *A. lapponica*, 1.38% for *A. inopinatus*, 2.65% for *B. tricaudatus*, 2.82% for *S. compacta,* and 8.66% for *O. unicolor*.

The posterior mean of the TMRCAs of eastern and western populations ranged between 0.23 and 2.41 Ma (Fig. 4). The maximum percentage of TMRCA overlap was between *C. alternans* and *E. aurivillii*, with almost 76% (Table 1). Most species pair combinations overlapped more than 10%, except for *C. alternans* and *B. tricaudatus*, which overlapped less than 9%. *Onocosmoecus unicolor* had the least TMRCA overlap with any other species; no more than 5% overlap (Table 1). Permutation tests for each order-level tree indicate that distributions are significantly different from random assignments (Ephemeroptera $P = 0.0369$, Plecoptera $P = 1e-04$, and Trichoptera $P = 0.0425$).

Discussion

The vastness of northern Canada, in combination with the difficulties of conducting fieldwork in remote locations, has presented a great challenge for revealing patterns of aquatic

insect biodiversity and community structure in that region. Moreover, the taxonomy of most aquatic insects is based on adults (especially males), which are typically short lived and difficult to collect at northern latitudes. In contrast, the immature stages of aquatic insects are relatively easy to collect, but are difficult or in most cases impossible to identify to species-level using available keys.

The biodiversity of northern aquatic insects was relatively poorly known before the advent of molecular taxonomic techniques. For example, Danks (1981) reported a total of just 38 species of EPTs distributed across the entirety of North American's Arctic zone from Alaska to Greenland. This represented the sum total of knowledge up to that time, although relatively few additional records were published through the end of the century. In contrast, Zhou et al. (2009, 2010), based on collections made during a 6-year period, reported 112 species from just a single site in the subarctic ecoclimatic zone of northern Manitoba. Although subarctic environments are expected to support higher diversity than arctic ones, it is nonetheless surprising to find a nearly threefold increase in species richness from just a single subarctic site—albeit, one that borders the northernmost margin of the Boreal ecoclimatic zone. In our broader-scale study conducted over a much shorter timeframe (i.e., only 4 months of sampling over two years), we recovered a total of 88 EPT species from four arctic sites and four subarctic sites (Fig. 2)—exceptional diversity given the limited sampling effort. Indeed, this figure represents well over twice the historical benchmark for arctic EPTs developed over the previous century, reflecting, in part, the higher ratio of cryptic diversity revealed by DNA barcodes relative to traditional taxonomy (Hebert et al. 2004; Witt et al. 2006). When comparing our data with the longer-term DNA barcoding studies of Zhou et al. (2009, 2010), we find about 35% more species of Ephemeroptera and about four times more species of Plecoptera in our study. The differences in plecopteran diversity might reflect order-specific habitat preferences, as most plecopteran species were found in more northerly locations (data not shown, Cordero et al. in prep.). In addition, our study shows higher diversity evenness across orders compared to Zhou et al. (2009, 2010). This may reflect ecoregional differences at a broader spatial scale compared to essentially a single location—Churchill, Manitoba.

Special comment is warranted about the relatively low success rate of obtaining sequences from particular specimens. While we obtained sequences for all the adult specimens sampled, we had a considerably lower success rate for larvae. This difference is perhaps related to the fact that immature EPTs were collected en masse, sometimes with hundreds of specimens preserved during a particular collecting event. Although an effort was made to frequently

change or refresh alcohol during the initial fixation period, we suspect these efforts failed when too many specimens were included in a single collecting jar. The problem was especially acute for the Trichoptera, where sequences were obtained from less than half the specimens sampled from Moosonee and Norman Wells. Large sample sizes from those locations no doubt contributed to the problem; however, the situation was compounded in case-dwelling larvae—perhaps because they were more resistant to preservation when they remained within their retreats during fixation. Our limited success of obtaining sequences from larvae highlights the importance of balancing the conflicting objectives of collecting the maximum number of specimens while conserving ethanol, which is difficult to obtain and transport in remote northern locations. Nonetheless, even with the relatively low success rate, we are confident that—with up to 10 specimens sampled for each morphospecies—the great majority of EPT species present at each site were successfully identified.

One of the promises of DNA barcoding is that it integrates multiple biological disciplines, such as taxonomy, biological databasing, ecology, systematics, phylogenetics, and phylogeography (Monaghan et al. 2009). Studies that have succeeded in such endeavors have not only characterized diversity in exceptional detail for some taxa (e.g., Smith et al. 2008) but have also complemented

important aspects about their natural and evolutionary histories. DNA barcoding often yields high levels of cryptic diversity (Hebert et al. 2004; Smith et al. 2008; Velonà et al. 2015). Cryptic diversity is often explained by imperceptible morphological, ecological, or behavioral differences among closely related lineages (Bickford et al. 2007). We find, based on the order-level phylogenetic trees (ESM 3–5), 22 species of EPTs with COI divergence >2.5% that are not represented in GenBank or BOLD databases. Hence, these results suggest that as much as 14% of the diversity we recovered represents hitherto unrecognized diversity, yet to be examined by taxonomy experts. A striking example is seen in the trichopteran *Onocosmoecus unicolor*, where a COI difference of >8% was found between eastern and western populations (Fig. 3) and a TMRCA > 2 Myr. This particular caddisfly has had a long and checkered taxonomic history, with at least six synonyms currently recognized (Wiggins 1996; Rasmussen and Morse 2014). Given the amount of morphological variation within *O. unicolor* as currently defined (Wiggins and Richardson 1986), it seems likely that at least two valid species are currently included under that name.

The Pleistocene glacial period is a well-known speciation driver, which can lead to cryptic diversity through the generation of incipient species after recurrent periods of

Table 1 Comparison of time to the most common recent ancestor (TMRCA) for species pairs

Species 1	Species 2	μ	σ^2	D_1	D_2	P_1	P_2
C. alternans	*E. aurivillii*	0.21	0.09	−0.58	0.39	0.7720	0.8450
S. compacta	*B. tricaudatus*	0.87	0.2	−0.92	0.75	0.6460	0.7080
A. inopinatus	*S. compacta*	0.66	0.15	−1.42	1.02	0.4780	0.6100
E. aurivillii	*A. lapponica*	0.26	0.13	−1.05	1.05	0.6000	0.6000
C. alternans	*A. lapponica*			−2.18	1.46	0.2760	0.4650
A. inopinatus	*B. tricaudatus*			−2.7	1.58	0.1770	0.4300
A. lapponica	*A. inopinatus*	0.4	0.13	−1.97	1.8	0.3250	0.3680
A. lapponica	*S. compacta*			−3.53	2.3	0.0776	0.2500
A. lapponica	*B. tricaudatus*			−4.93	2.63	0.0137	0.1890
E. aurivillii	*A. inopinatus*			−3.02	2.75	0.1310	0.1690
E. aurivillii	*S. compacta*			−4.58	2.99	0.0220	0.1350
C. alternans	*A. inopinatus*			−5.12	3.13	0.0105	0.1180
E. aurivillii	*B. tricaudatus*			−5.99	3.19	0.0027	0.1110
C. alternans	*S. compacta*			−7.46	3.26	0.0002	0.1030
C. alternans	*B. tricaudatus*			−9.56	3.4	0.0000	0.0891
B. tricaudatus	*O. unicolor*	1.06	0.25	−5.39	3.98	0.0070	0.0466
S. compacta	*O. unicolor*			−7.53	4.54	0.0002	0.0232
A. inopinatus	*O. unicolor*			−11.95	5.15	0.0000	0.0100
A. lapponica	*O. unicolor*			−15.07	5.92	0.0000	0.0031

Posterior TMRCA (million years ago) mean (μ) and standard deviation (σ^2) for species 1. The mean distance (*D*) and percentage of overlap (*p*) are for comparison between species 1 and species 2, and viceversa. *Af, Acentrella lapponica; Ai, Ameletus inopinatus; Bt, Baetis tricaudatus; E. aurivillii, Ephemerella aurivillii; S. compacta, Skwala compacta; C. alternans, Ceratopsyche alternans* and *Ou, Onocosmoecus unicolor*

isolation (Hewitt 2000, 2004; Knowles 2001; Hawlitschek et al. 2012; Sánchez-Ramírez et al. 2015). A possible example of such cryptic diversity can be seen in *Isoperla transmarina* (see ESM 4), wherein divergence data suggests the present of at least two separate species under that name. During the last glacial maximum (LGM), present-day arctic and subarctic regions in Canada were covered by massive continental ice sheets (Pielou 1991; Clark and Mix 2000). This implies that these communities likely persisted in refugia, later recolonizing northern territories as climatic conditions became favorable (Bernatchez and Wilson 1998; Keppel et al. 2012). Our coalescent-based population-level exploration revealed seven species with strong east–west genetic structure (Fig. 3). Furthermore, molecular clock analyses indicate that these geographic populations diverged during the Pleistocene Epoch (Fig. 4), probably as result of isolation in eastern (possibly the Appalachian Mountains and the Atlantic coastal plains) and western (most likely Beringia) refugia (Hopkins 1967). Besides molecular clock TMRCA estimates, we find a substantial degree of temporal overlap (Table 1) among the divergence times of these populations. The synchronicity in divergence times might be viewed as an additional indicator of the effect of an extrinsic factor on population divergence, such as Pleistocene glacial dynamics. Furthermore, we detected 12 other species displaying strong genetic divergence (e.g., high clade posterior probabilities, see ESM 3–5), but with no east–west geographic structure. This pattern could be the result of post-glacial dispersal.

In this study, we show that a rapidly conducted biological survey using a DNA barcoding approach has a great utility for assessing biodiversity in remote and difficult to access territories such as northern Canada. DNA barcoding studies executed over a broad-geographic area not only aids biodiversity discovery, but also helps reveal how EPT biodiversity originated and is maintained in northern Canada. Further, as many aquatic insects are difficult to identify, particularly in the aquatic stage, a combined barcoding-phylogeographic approach can serve as a powerful tool to gain preliminary evolutionary and ecological insights that can be tested using additional data (e.g., multiple genes, more localities) and emerging analytical tools.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Ball SL, Hebert PD, Burian SK, Webb JM (2005) Biological identifications of mayflies (Ephemeroptera) using DNA barcodes. J N Am Benthol Soc 24:508–524

Benke AC, Van Arsdall TC, Jr Gillespie DM (1984) Invertebrate productivity in a subtropical blackwater river: the importance of habitat and life history. Ecol Monogr 54:25–63

Bernatchez L, Wilson CC (1998) Comparative phylogeography of Nearctic and Palearctic freshwater fishes. Mol Ecol 7:431–452

Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I (2007) Cryptic species as a window on diversity and conservation. Trends Ecol Evol 22:148–155

Boumans L, Brittain JE (2012) Faunistics of stoneflies (Plecoptera) in Finnmark, northern Norway, including DNA barcoding of Nemouridae. Norw J Entomol 59:196–215

Clark PU, Mix AC (2000) Global change: ice sheets by volume. Nature 406:689–690

Craft KJ, Pauls SU, Darrow K, Miller SE, Hebert PD, Helgen LE, Novotny V, Weiblen GD (2010) Population genetics of ecological communities with DNA barcodes: an example from New Guinea Lepidoptera. Proc Natl Acad Sci USA 107:5041–5046

Danks HV (1981) Arctic Arthropods: a review of systematics and ecology with particular reference to the North American fauna. Entomol Soc Can, Ottawa

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 9:772

de Waard JR, Ivanova NV, Hajibabaei M, Hebert PDN (2008) Assembling DNA barcodes: analytical protocols. In: Martin C (ed) Methods in molecular biology: environmental genetics. Humana Press, Totowa, pp 275–293

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7:214

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 8:1969–1973

Ekrem T, Stur E, Hebert PD (2010) Females do count: documenting Chironomidae (Diptera) species diversity using DNA barcoding. Org Div Evol 10:397–408

Gullan PJ, Cranston PS (2010) The insects: an outline of entomology. Wiley, New York

Hajibabaei M, Ivanova NV, Ratnasingham S, Dooh RT, Kirk SL, Mackie PM, Hebert PD (2005) Critical factors for assembling a high volume of DNA barcodes. Philos Trans R Soc B 360:1959–1967

Hajibabaei M, Singer GA, Hebert PD, Hickey DA (2007) DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. Trends Genet 23:167–172

Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22:160–174

Hawlitschek O, Hendrich L, Espeland M, Toussaint EF, Genner MJ, Balke M (2012) Pleistocene climate change promoted rapid diversification of aquatic invertebrates in Southeast Australia. BMC Evol Biol 12:142

Hebert P, Gregory TR (2005) The promise of DNA barcoding for taxonomy. Syst Biol 54:852–859

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. Proc R Soc Lond B 270:313–321

Hebert PD, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proc Natl Acad Sci USA 101:14812–14817

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. Nature 405:907–913

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. Philos Trans R Soc B 359:183–195

Higgins DG, Sharp PM (1988) CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73:237–244

Hopkins DM (1967) The Cenozoic history of Beringia—a synthesis. In: Hopkins DM (ed) The Bering land bridge. Stanford University Press, Stanford, pp 451–484

Keppel G, Van Niel KP, Wardell-Johnson GW, Yates CJ, Byrne M, Mucina L, Schut AGT, Hopper SD, Franklin SE (2012) Refugia: identifying and understanding safe havens for biodiversity under climate change. Global Ecol Biogeogr 21:393–404

Kjaerstad GA, Webb JM, Ekrem TO (2012) A review of the Ephemeroptera of Finnmark-DNA barcodes identify Holarctic relations. Norw J Entomol 59:182–195

Knowles LL (2001) Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshopprers. Mol Ecol 10:691–701

Merritt RW, Cummins KW, Berg MB (2008) An introduction to the aquatic insects of North America. Kendall/Hunt Publishing Company, Dubuque

Monaghan MT, Wild R, Elliot M, Fujisawa T, Balke M, Inward DJ, Lees DC, Ranaivosolo R, Eggleton P, Barraclough TG, Vogler AP (2009) Accelerated species inventory on Madagascar using coalescent-based models of species delineation. Syst Biol 58:298–311

Papadopoulou A, Anastasiou I, Vogler AP (2010) Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. Mol Biol Evol 27:1659–1672

Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290

Pielou EC (1991) After the ice age: the return of life to glaciated North America. University of Chicago Press, Chicago

Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst Biol 55:595–609

R Core Team (2015) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Austria

Rasmussen AK, Morse JC (2014) Distributional checklist of Nearctic Trichoptera (Summer 2014 Revision). Unpublished, Florida A&M University, Tallahassee 487 pp. http://www.Trichoptera. org. Accessed July 28

Ruiter DE (1995) The Adult *Limnephilus* Leach (Trichoptera: Limnephilidae) of the New World. In: Cafazzo VM (ed) Bulletin new series, Ohio Biol Surv Bull 11:1

Ruiter DE, Boyle EE, Zhou X (2013) DNA barcoding facilitates associations and diagnoses for Trichoptera larvae of the Churchill (Manitoba, Canada) area. BMC Ecol 13:1

Sánchez-Ramírez S, Tulloss RE, Guzmán-Dávalos L, Cifuentes-Blanco J, Valenzuela R, Estrada-Torres A, Ruán-Soto F, Díaz-Moreno R, Hernández-Rico N, Torres-Gómez M, León H, Moncalvo JM (2015) In and out of refugia: historical patterns of diversity and demography in the North American Caesar's mushroom species complex. Mol Ecol 24:5938–5956

Smith MA, Rodriguez JJ, Whitfield JB, Deans AR, Janzen DH, Hallwachs W, Hebert PD (2008) Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. Proc Natl Acad Sci USA 105:12359–12364

Ståhls G, Savolainen E (2008) MtDNA COI barcodes reveal cryptic diversity in the *Baetis vernus* group (Ephmeroptera, Baetidae). Mol Phylogenet Evol 46:82–87

Tavaré S (1986) Some probabilistic and statistical problems in the analysis of DNA sequences. In: Miura RM (ed.) Some mathematical questions in biology–DNA sequence analysis. Lectures on mathematics in the life sciences, American Mathematical Society, 17:57–86

Velonà A, Brock PD, Hasenpusch J, Mantovani B (2015) Cryptic diversity in Australian stick insects (Insecta; Phasmida) uncovered by the DNA barcoding approach. Zootaxa 3957:455–466

Webb JM, Jacobus LM, Funk DH, Zhou X, Kondratieff B, Geraci CJ, DeWalt RE, Baird DJ, Richard B, Phillips I, Hebert PD (2012) A DNA barcode library for North American ephemeroptera: progress and prospects. Plos One 7:e38063

Wiggins GB (1996) Larvae of the North American caddisfly genera (Trichoptera). University of Toronto Press, Toronto

Wiggins GB, Richardson JS (1986) Revision of the *Onocosmoecus unicolor* group (Trichoptera: Limnephilidae, Dicosmoecinae). Psyche 93:187–216

Witt JDS, Threloff DL, Hebert PDN (2006) DNA barcoding reveals extraordinary cryptic diversity in an amphipod genus: implications for desert spring conservation. Mol Ecol 15:3073–3082

Zhou X, Adamowicz SJ, Jacobus LM, Dewalt RE, Hebert PDN (2009) Towards a comprehensive barcode library for arctic life—Ephemeroptera, Plecoptera, and Trichoptera of Churchill, Manitoba, Canada. Front Zool 6:30

Zhou X, Jacobus LM, Dewalt RE, Adamowicz SJ, Hebert PDN (2010) Ephemeroptera, Plecoptera, and Trichoptera fauna of Churchill (Manitoba, Canada): insights into biodiversity patterns from DNA barcoding. J N Am Benthol Soc 29:814–837