

RESEARCH ARTICLE

# Exploring Canadian Echinoderm Diversity through DNA Barcodes

Kara K. S. Layton<sup>1\*</sup>, Erin A. Corstorphine<sup>2</sup>, Paul D. N. Hebert<sup>2</sup>

Centre for Biodiversity Genomics, Biodiversity Institute of Ontario, University of Guelph, Guelph, ON, N1G 2W1, Canada

☉ These authors contributed equally to this work.

✉ Current address: Centre for Evolutionary Biology, University of Western Australia, 35 Stirling Highway, Crawley, Western Australia, 6009, Australia

\* [kara.layton@research.uwa.edu.au](mailto:kara.layton@research.uwa.edu.au)



CrossMark  
click for updates

OPEN ACCESS

**Citation:** Layton KKS, Corstorphine EA, Hebert PDN (2016) Exploring Canadian Echinoderm Diversity through DNA Barcodes. PLoS ONE 11 (11): e0166118. doi:10.1371/journal.pone.0166118

**Editor:** Roberta Cimmaruta, Università degli Studi della Tuscia, ITALY

**Received:** February 5, 2016

**Accepted:** October 24, 2016

**Published:** November 21, 2016

**Copyright:** © 2016 Layton et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** DNA sequences are deposited in GenBank and Accessions are available in the Methods section of the paper. The dataset DS-COIECH is available on BOLD; [dx.doi.org/10.5883/DS-COIECH](https://dx.doi.org/10.5883/DS-COIECH).

**Funding:** This research was funded, in part, by the Natural Sciences and Engineering Research Council (NSERC) through a Discovery Grant to PDNH and a postgraduate scholarship to EAC. Sequence analysis was enabled by funding from the government of Canada through Genome Canada and the Ontario Genomics Institute in support of the International Barcode of Life Project

## Abstract

DNA barcoding has proven an effective tool for species identification in varied groups of marine invertebrates including crustaceans, molluscs, polychaetes and echinoderms. In this study, we further validate its utility by analyzing almost half of the 300 species of Echinodermata known from Canadian waters. COI sequences from 999 specimens were assigned to 145 BINs. In most cases, species discrimination was straightforward due to the large difference (25-fold) between mean intra- (0.48%) and inter- (12.0%) specific divergence. Six species were flagged for further taxonomic investigation because specimens assigned to them fell into two or three discrete sequence clusters. The potential influence of larval dispersal capacity and glacial events on patterns of genetic diversity is discussed for 19 trans-oceanic species. Although additional research is needed to clarify biogeographic patterns and resolve taxonomic questions, this study represents an important step in the assembly of a DNA barcode library for all Canadian echinoderms, a valuable resource for future biosurveillance programs.

## Introduction

DNA barcoding employs sequence variation in a 648 bp region of the mitochondrial cytochrome *c* oxidase subunit 1 (COI) gene as a tool for specimen identification and species discovery [1]. Its effectiveness has been shown in birds [2,3], bats [4], fishes [5–7], insects [8,9] and several marine invertebrate groups, including stomatopods [10], molluscs [11,12], crustaceans [13], and echinoderms [14]. The routine occurrence of a barcode gap, where interspecific divergence is much greater than intraspecific divergence, underpins its strong performance in taxonomic assignments versus the complexities which often complicate morphological approaches. For example, a survey of 15 marine phyla reported that one third of specimens could not be assigned to a species through morphological analysis [15] because of factors such as damage to diagnostic characters during collection. Immature stages are often particularly challenging to identify beyond a family via morphology, making it difficult to estimate the timing and duration of larval dispersal [16]. The use of DNA barcoding for species

(OGI-036). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

identification solves many of the problems associated with the application of morphology-based approaches. It has, for example, proven highly effective in identifying immature stages of fishes [17] and varied lineages of marine invertebrates [10,16,18]. It is particularly useful when a comprehensive barcode reference library is available. Moreover, divergence thresholds are a powerful tool for revealing newly encountered species, albeit they can overlook recently diverged species [10,19].

Although their phylogenetic relationships, developmental patterns, and reproductive biology have been well investigated, Canadian echinoderms have seen little genetic analysis. Three studies have examined introgression along secondary contact zones [20–22], while another investigated a cryptic species complex [23]. A further investigation documented patterns of mtDNA divergence in four trans-Arctic sea urchin species [24] while another used COI sequences to assess phylogeographic patterns in Arctic marine invertebrates, including several echinoderms from this region of North America [25]. While little information is available on patterns of COI sequence divergence in Canadian echinoderms, studies on this phylum in other regions have reported high interspecific divergences (2.5–24.2%) and low intraspecific distances (<1%) at COI [14,26,27]. Although these values are similar to those in other marine invertebrates [11,13,28] and fishes [6], low divergences were detected between some sibling species of echinoderms. For example, divergences of 1.1–1.2% for COI were observed between sister taxa of sea stars [29], while interspecific divergences of 2–3% were noted in closely-related echinoid species [30].

The variable climatic and hydrographic conditions of Canada's oceans during the late Cenozoic [31] means that its marine species have complex histories of range expansion and fragmentation. For example, the opening of the Bering Strait during the mid-Pliocene (3.5 Mya), coupled with ice-free Arctic waters [32], allowed extensive migration between the north Pacific and Arctic-Atlantic oceans [33,34]. Subsequent Pleistocene glaciations eradicated much of the fauna in the Arctic and northwest Atlantic where conditions were more severe than in the Pacific. Differing histories of re-colonization from Pacific and Atlantic refugia following deglaciation approximately 14,000 years ago [32] is reflected in varied levels of population structure, ranging from closely-related species complexes (*Macoma*, [35]; *Mallotus*, [36]) to widespread species with little genetic divergence (*Strongylocentrotus*, [24,37]).

This study begins the construction of a DNA barcode reference library for Canadian echinoderms, but since several of these taxa also occur in polar and temperate regions across the northern hemisphere, the library has broader utility. Approximately 300 species of echinoderms are known from Canada's oceans with two-thirds (217) restricted to the NE Pacific [38–40]. The present study tests the efficacy of DNA barcoding and examines phylogeographic patterns resulting from Canada's complex glacial history. By focusing on Canadian echinoderms, the size of the fauna (300 species) makes it feasible to develop a comprehensive reference library. Moreover, the circumboreal distributions of several species provide an opportunity to investigate patterns of intraspecific divergence at a large geographic scale.

## Methods

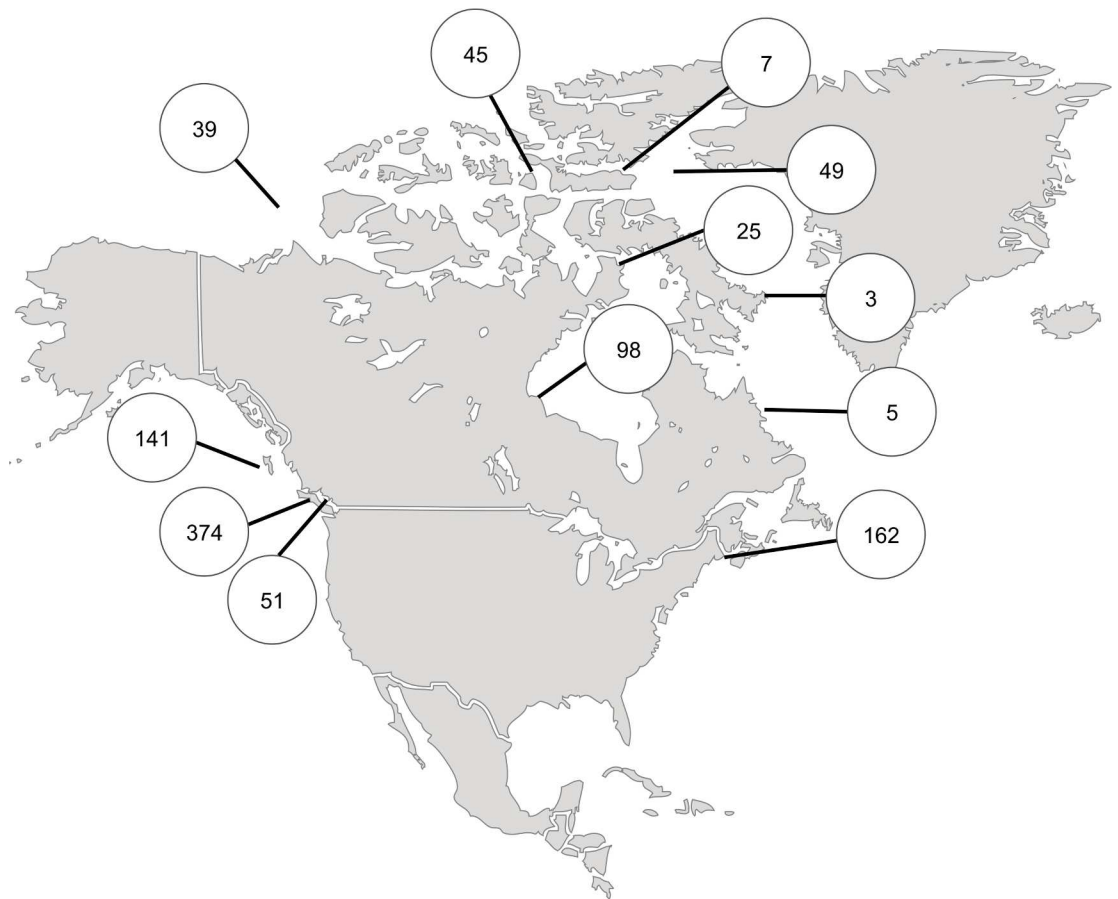
### Ethics Statement

Fieldwork in Churchill, Manitoba was conducted under permits issued by Manitoba Conservation Wildlife and Ecosystem Protection to the Churchill Northern Studies Centre (CNSC) for research in the Churchill Wildlife Management Area. Collections in British Columbia, Labrador, and New Brunswick were conducted under licenses from Fisheries and Oceans Canada. No specific permits were required for other collection activities as they were not conducted on privately owned or protected land. No field studies involved the collection of endangered or

protected species. A list of BOLD sample IDs, process IDs, GenBank accession numbers, and the institution storing each specimen in this study is available in [S1 Table](#).

## Specimen collection and taxonomy

Tissue samples were obtained from 1285 specimens in the collections of the Centre for Biodiversity Genomics (CBG) and the Royal British Columbia Museum (RBCM). These samples included 316 specimens representing 78 species from trawl collections made by the RBCM between 2000 and 2006 at depths > 200m at various localities along the coast of British Columbia ([Fig 1](#)). Another 252 specimens were obtained by SCUBA and dredging between 1995 and 2002 from Queen Charlotte Sound and Nunavut ([Fig 1](#)) while 589 specimens were collected by SCUBA, by dredging and by hand from sites near Baffin Island, British Columbia, Labrador, Manitoba, and New Brunswick between 2007 and 2014 ([Fig 1](#)). A final 128 specimens were gathered by trawls at depths < 700m from the Beaufort Sea, Baffin Bay and the Labrador Sea in 2010 as part of the Canadian Healthy Oceans Network ([Fig 1](#)). Scientific names follow the World Register of Marine Species (WoRMS, <http://www.marinespecies.org>). All RBCM specimens were identified by Philip Lambert, an echinoderm specialist. All CBG specimens were identified using regional taxonomic keys (British Columbia: [38–41]; Atlantic and Arctic Canada: [42]). When a species-level identification could not be made, an interim name was



**Fig 1. Sampling map.** Collection localities and sample sizes for 999 specimens that generated a COI sequence in this study.

doi:10.1371/journal.pone.0166118.g001

assigned that coupled a genus-level identification with a Barcode Index Number (BIN) designation [43]. All specimens are stored as vouchers in 95% ethanol or frozen at -20°C in collections at the RBCM or the CBG (S1 Table). Specimen and sequence data are available in the dataset DS-COIECH [dx.doi.org/10.5883/DS-COIECH](https://dx.doi.org/10.5883/DS-COIECH) on BOLD, the Barcode of Life Data Systems [44]. All sequences have also been deposited in GenBank under the following accession numbers: GU670162-65; GU670167-81; GU670187-94; HM400305-69; HM400539; HM405487; HM405870-914; HM473811-74; HM473876-89; HM473902-57; HM542062-420; HM542908-3073; JF891304-16; JN295388; JN314244; KU495734-918.

### DNA extraction, COI amplification, and sequencing

A DNA extract was prepared from the tube feet or gonadal tissue of each specimen. Tissue was placed into 96-well plates containing modified CTAB lysis buffer [45] and proteinase K (20mg/mL). Samples were incubated at 56°C for 18–24 hours before extraction using a 3um filter plate following a standard manual protocol [45]. Extracts were re-suspended in 40 µL of molecular grade water. COI was amplified using several primer combinations with products being 652bp, 658bp or 841bp in length depending on the primer set (Table 1). Polymerase chain reactions (PCRs) were carried out in 12.5 µL reaction volumes containing: 6.25 µL 10% trehalose, 2 µL molecular grade water, 1.25 µL 10X PCR Platinum Taq buffer, 0.625 µL MgCl<sub>2</sub> (50 mM), 0.125 µL of each primer (10 mM), 0.0625 µL dNTPs (10 mM), 0.06 µL of Platinum Taq polymerase and 2 µL of DNA template (20–60 ng) using a thermocycling profile of one cycle at 94°C for 1 min, five cycles at 94°C for 40 s, 45°C for 40 s, and 72°C for 1 min, 35 cycles of 94°C for 40 s, 51°C for 40 s and 72°C for 1 min, and a further extension period of 72°C for 5 min. PCR products were visualized on pre-cast 2% agarose gels (E-gel 96, Invitrogen) and PCR products with single, bright bands were selected for bidirectional sequencing using BigDye version 3.1 on a 3730XL DNA Analyser (Applied Biosystems). Cycle-sequencing reactions incorporated the same primers as those used to generate the selected PCR product. Sequences from specimens lacking a species designation were analyzed with both the BOLD identification engine and BLAST in GenBank [46] to determine if sequence matches existed. Congruence between identifications based on sequence data and morphological characters was confirmed prior to assigning names to unidentified specimens.

Four primer sets used in this study with their amplicon length and references for both forward and reverse primers, respectively.

### Data analysis

All sequences were edited manually using Sequencher 4.8 and aligned by eye in MEGA 6.0 [48]. As most sequences were barcode compliant (N = 987), they received a Barcode Index

**Table 1. Primers used in this study.**

Primer name	Primer sequence (5'-3')	Length of amplicon (bp)	References
EchinoF1	TTTCAACTAATCATAAGGACATTGG	841	[14]
COIeR1	GTCGTGTRTCTACRTCCAT		[26]
EchinoF1	TTTCAACTAATCATAAGGACATTGG	658	[14]
HCO2198	TAAACTTCAGGGTGACCAAAAATCA		[47]
LCOech1aF1	TTTTTCTACTAAACACAAGGATATTGG	658	This study
HCO2198	TAAACTTCAGGGTGACCAAAAATCA		[47]
COIeF1	ATAATGATAGGAGGRTTTGG	652	[26]
COIeR1	GTCGTGTRTCTACRTCCAT		[26]

doi:10.1371/journal.pone.0166118.t001

Number (BIN), aiding in species delimitation [43]. Maximum-likelihood trees were constructed with 1000 bootstrap replicates and either an HKY+I model in MEGA 6.0 [48] or with a GTR+G+I model in RAxML v8 [49]. Models with the highest AIC values generated by jModelTest v2 [50,51] were selected as most appropriate. Pairwise estimates of intra- and interspecific sequence divergence were calculated using the K2P distance model [52] and the Distance Summary tools on BOLD [44]. The presence of a barcode gap was analysed by plotting maximum intraspecific divergence against Nearest-Neighbor distance. All species with an intraspecific divergence >2% were flagged for further investigation. Maximum and mean intraspecific divergences were plotted against the number of individuals sampled within a species and regression analysis was carried out in RStudio using the Picante and VEGAN packages to determine the significance of this relationship [53,54].

## Results

### Sequence recovery

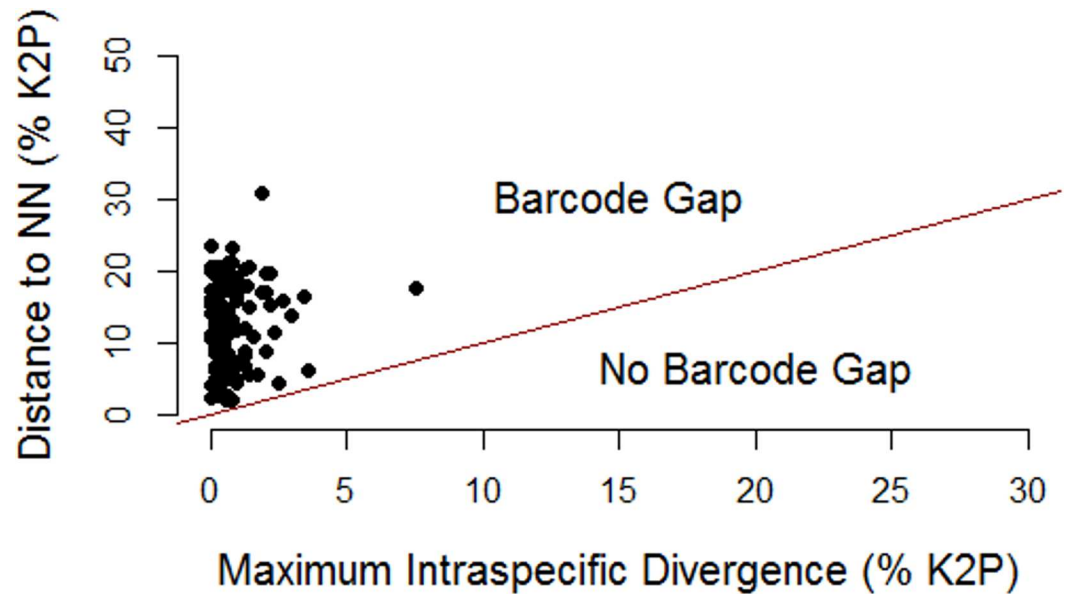
COI sequences were recovered from 999 specimens representing 141 species, 77 genera and 43 families for a success rate of 78% (999/1285). Sequences ranged in length from 347bp to 658bp but 95% were over 600bp. COI sequences could not be recovered from 18 species of deep-sea echinoderms from the RBCM collection despite their young age (2–10 years). Nuclear pseudogenes of mitochondrial origin (NUMTs, [55]) were occasionally encountered along with authentic COI, but most were <180bp so full sequences could be recovered following deletion of the NUMT region in the initial segment of each forward and reverse read.

### Taxonomic issues

The 141 presumptive species included 118 taxa with species-level identifications based on morphological study. The other 23 species could only be placed to a genus, but were assigned interim species identifications based on their BIN membership (e.g. *Ceramaster* sp. AAI7443). Many of these 23 taxa may represent undescribed species or taxa previously undocumented from Canada. For instance, two specimens of sea urchin (*Strongylocentrotus* sp. AAA9523) collected in arctic Canada did not match any species known from this region. Several characters commonly used to identify urchins, such as number of pore pairs and spine wedges, overlap in the other two Canadian members of this genus (*S. droebachiensis*, *S. pallidus*) and were similarly undiagnostic in this new taxon. However, it was distinguished morphologically by the deep purple colouration of its test, tube feet and spines. In addition, the density and size of spines on its test were reduced in comparison with other *Strongylocentrotus* specimens from Canada. Certain cases where there was initial discordance between barcode results and taxonomic assignments were resolved when taxonomic reanalysis led to their placement in the species corresponding with the barcode data. Many of these cases involved juveniles or genera where two species have subtle morphological differences.

### COI variation

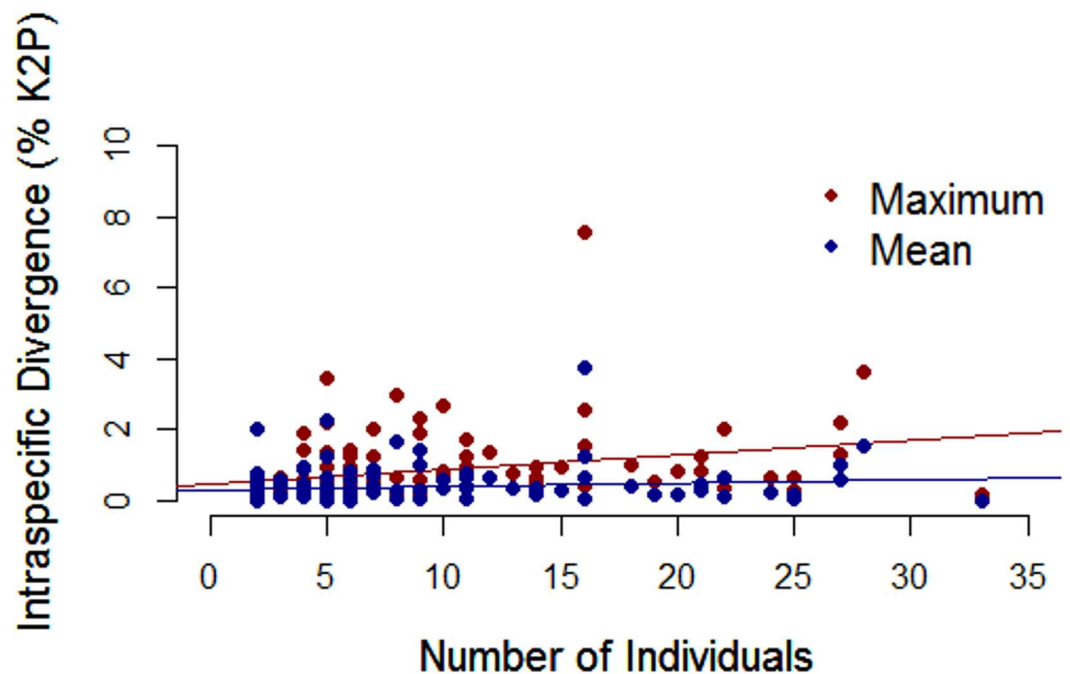
An average of seven individuals were sequenced per species (range of 1 to 33) with 113 taxa represented by multiple specimens and 28 taxa by singletons. Mean intraspecific distance was 0.48% (range 0.0–7.6%) while mean interspecific divergence was 25-fold greater at 12.0% (range 2.0–26.2%). S2 Table presents mean and maximum intraspecific distances for each of the 113 morphospecies represented by two or more specimens. A barcode gap was present for all taxa (Fig 2) although 12 species had maximum intraspecific divergences greater than 2%. Values of maximum intraspecific divergence showed a significant increase with the number of



**Fig 2. Barcode gap.** Nearest-Neighbor distances (% K2P) plotted against maximum intraspecific divergences (% K2P) for 113 taxa with two or more individuals. All taxa show a barcode gap.

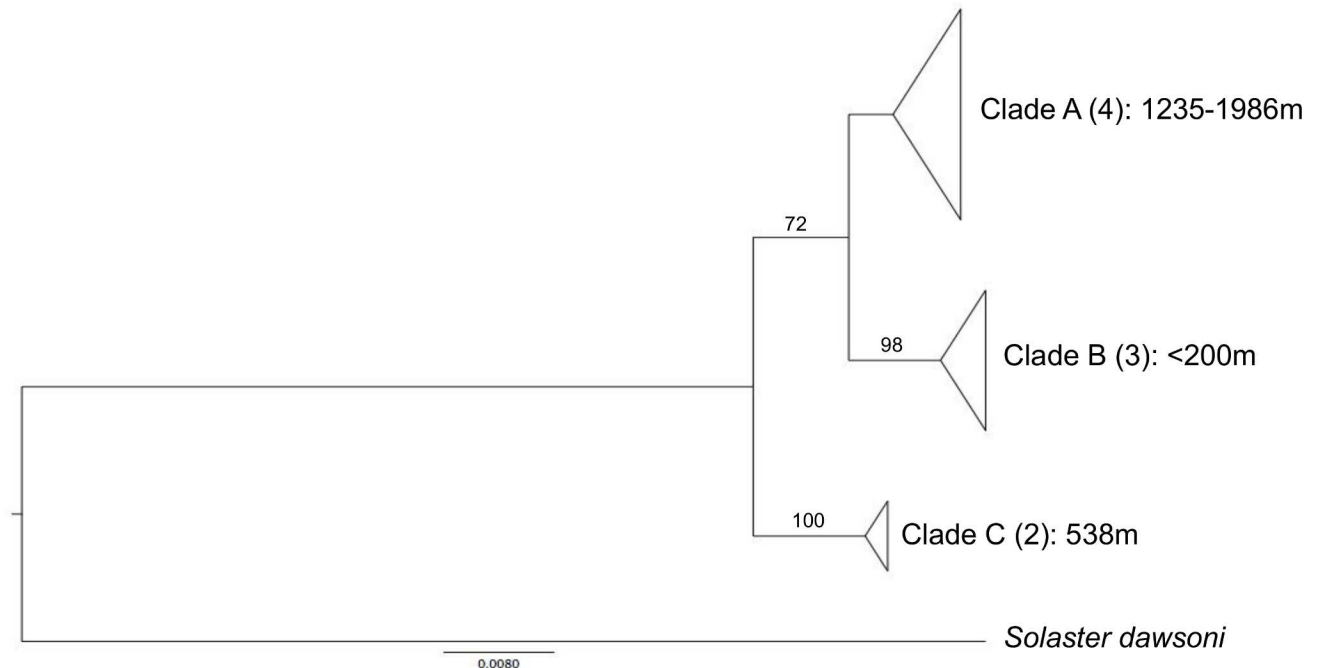
doi:10.1371/journal.pone.0166118.g002

individuals sampled, but the coefficient of determination was low (Fig 3). Four morphospecies were not assigned a BIN because their sequence length was less than 500bp, but the other 137 were assigned to 145 BINs reflecting the fact that six species (*Gorgonocephalus arcticus*, *Leptasterias hexactis*, *Leptosynapta clarki*, *Lophaster furcilliger*, *Ophiura sarsii*, *Pteraster militaris*)



**Fig 3. Relationship between COI distance and sample size.** Maximum and mean intraspecific divergences (% K2P) plotted against the number of individuals sampled for 113 species. The regression between maximum intraspecific divergence and the number of individuals sampled is significant ( $R^2 = 0.089$ ,  $p < 0.01$ ).

doi:10.1371/journal.pone.0166118.g003



**Fig 4. Depth partitioning in *Lophaster furcilliger*.** Maximum-likelihood tree (HKY+I) for specimens of *Lophaster furcilliger* from British Columbia. Groups A, B and C are partitioned by depth and sequence number is presented in brackets for each clade. Scale bar represents percent sequence divergence.

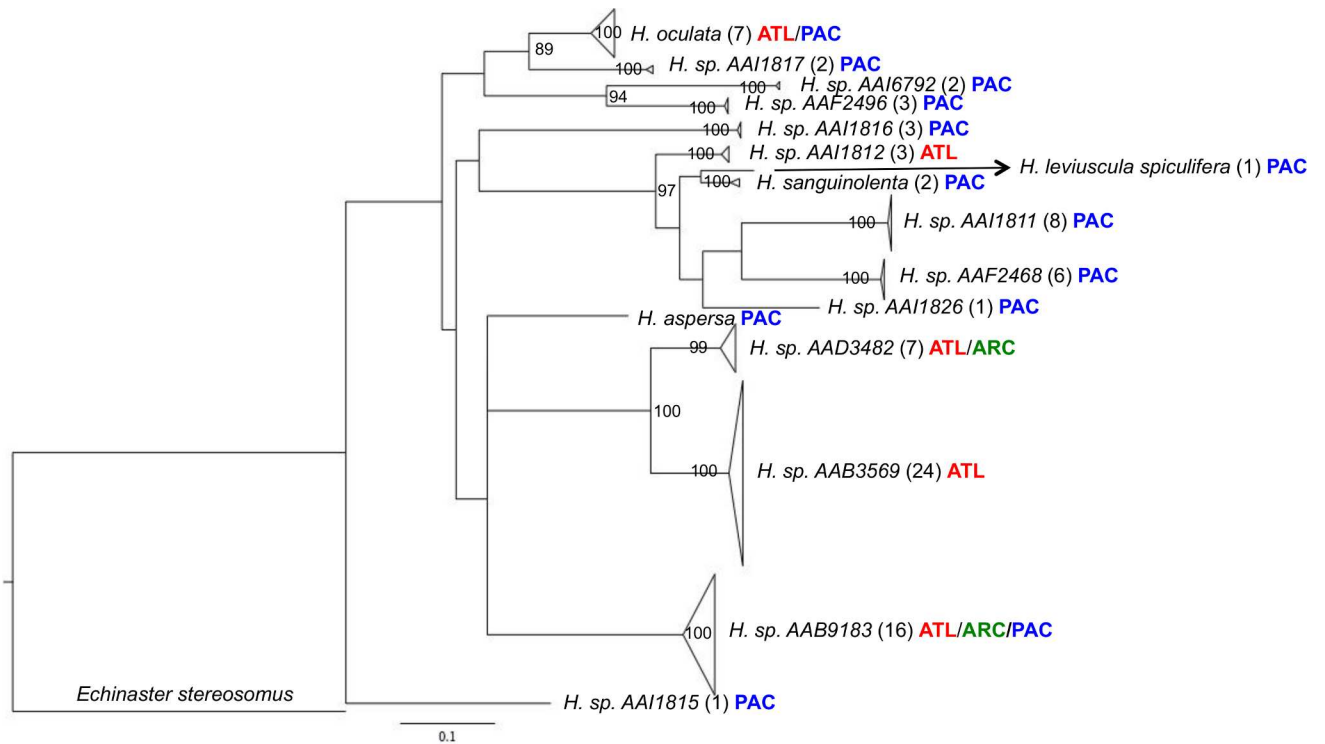
doi:10.1371/journal.pone.0166118.g004

included two or three clusters, each assigned to a different BIN. [S3 Table](#) reports the number of sequences and locality information for each BIN. A ML tree for one of these species, *L. furcilliger*, showed genetic partitioning associated with water depth with specimens in cluster A collected from 1200–2000m while those in cluster B were found at < 200m and cluster C at 500m ([Fig 4](#)). When these six taxa with multiple BINs (which may represent cryptic species) were removed from the analysis, the mean intraspecific divergence dropped to 0.32%. Mean Nearest-Neighbor distances were lowest between *S. droebachiensis* and *S. sp. AAA9523* (2.0%).

Some specimens of *Henricia* could not be assigned to a species because morphological differences among species in this genus are so conserved. However, the 87 sequences for this genus formed 16 clusters with relatively high interspecific distances (mean 13.7%; range 2.6–18.3%) and low intra-group divergences (mean 0.33%; range 0–1.6%) ([Fig 5](#)). Among the 12 clusters represented by multiple specimens, seven were only found in the Pacific, two in the Atlantic, one in both the Pacific and Atlantic, one in both the Arctic and Atlantic, and another in all three oceans.

### Divergence across oceans

Twelve of 19 species collected from two or more oceans ([Table 2](#)) had low intraspecific divergence with little to no geographic structure. Two others showed clear geographic structure although their maximum divergences were below 2%, while two species had maximum divergences above 2%, but without obvious geographic partitioning. Finally, three species showed clear genetic structure and maximum divergences above 2%. NJ trees are presented for the three species with the most conspicuous geographic structure; one with a maximum divergence below 2% (*Crossaster papposus*), and two with maximum divergences above 2% (*P*



**Fig 5. Deep divergences at COI within *Henricia*.** Maximum-likelihood tree (GTR+G+I) for 16 putative species of *Henricia*. Locality information is provided for each lineage (Atl = Atlantic, Arc = Arctic, Pac = Pacific) and sequence number is presented in brackets for each clade. Scale bar represents percent sequence divergence.

doi:10.1371/journal.pone.0166118.g005

*militaris*, *Solaster endeca*) (Fig 6). Sequences of *P. militaris* and *S. endeca* from the White Sea were included in the phylogenetic analysis and clustered with populations from the Atlantic Ocean. Divergences between lineages were relatively shallow (<3%) in the five species with geographic partitioning. Different lineages were normally allopatric, but *G. arcticus* was exceptional as two specimens from Nunavut clustered with Atlantic conspecifics rather than with other samples from Nunavut.

Four patterns of regional divergence in COI sequences were observed among the 19 echinoderm species analysed from two or more of Canada's oceans. Group A: species with maximum intraspecific divergences > 2% and geographic partitioning; Group B: species with maximum intraspecific divergences > 2% and no obvious geographic partitioning; Group C: species with maximum intraspecific divergences < 2% and geographic partitioning and Group D: species with maximum intraspecific divergences < 2% and no obvious geographic partitioning. Sample sizes are provided for each locality (Pac = Pacific, Arc = Arctic, Atl = Atlantic).

## Discussion

### Barcode Recovery

The present study recovered barcode sequences from 78% of the specimens that were analyzed. The failure to recover barcodes from all specimens is likely a consequence of the fact that current primer sets are not effective for all echinoderm lineages. Although the present study expanded the primer sets [14] available, more work is needed to develop a set that recovers barcodes from all echinoderms.



**Table 2. Divergences at COI for echinoderm species from multiple oceans.**

Group	Species	# BINs	Sample size	Maximum intraspecific divergence (%K2P)
			Pac/Arc/Atl	
A	<i>Gorgonocephalus arcticus</i>	2	0/11/5	2.53
	<i>Pteraster militaris</i>	3	3/0/5	2.98
	<i>Solaster endeca</i>	1	2/1/4	2.02
B	<i>Ophiopholis aculeata</i>	1	0/8/19	2.18
	<i>Ophiura sarsii</i>	2	3/13/0	7.58
C	<i>Crossaster papposus</i>	1	13/4/4	1.24
	<i>Leptasterias littoralis</i>	1	0/16/11	1.29
D	<i>Chirodota laevis</i>	1	0/1/5	0
	<i>Ctenodiscus crispatus</i>	1	10/15/0	0.62
	<i>Cucumaria frondosa</i>	1	0/9/8	0.98
	<i>Florometra serratissima</i>	1	8/13/0	0.80
	<i>Henricia oculata</i>	1	3/0/4	1.26
	<i>Henricia</i> sp. AAD3482	1	0/1/6	0.77
	<i>Henricia</i> sp. AAB9183	1	2/1/13	1.55
	<i>Ophiura robusta</i>	1	0/24/1	0.31
	<i>Psolus fabricii</i>	1	0/11/5	0.43
	<i>Psolus phantapus</i>	1	0/5/2	0.51
	<i>Strongylocentrotus droebachiensis</i>	1	11/0/8	0.53
	<i>Strongylocentrotus pallidus</i>	1	1/24/0	0.31

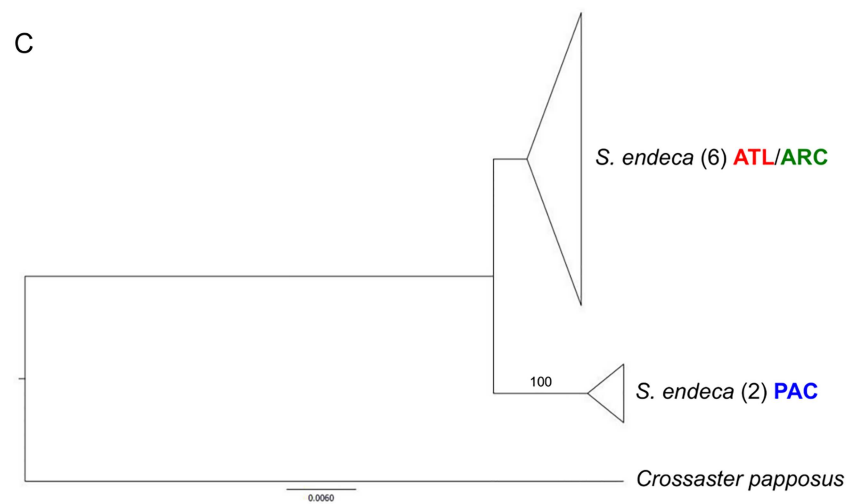
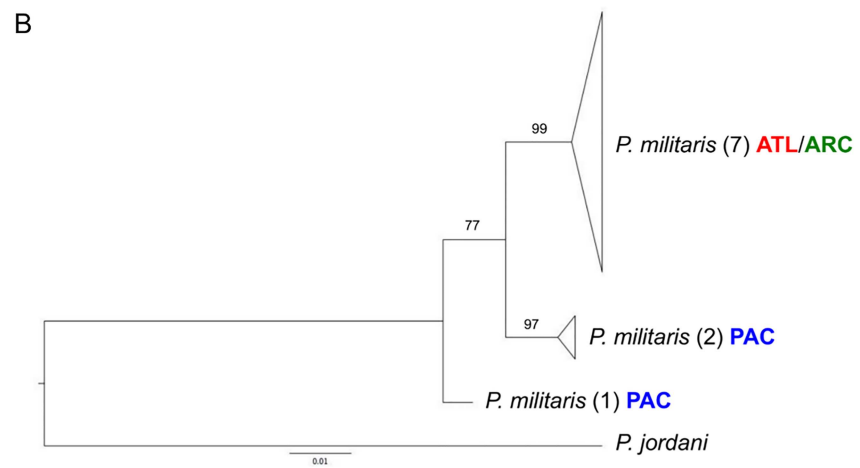
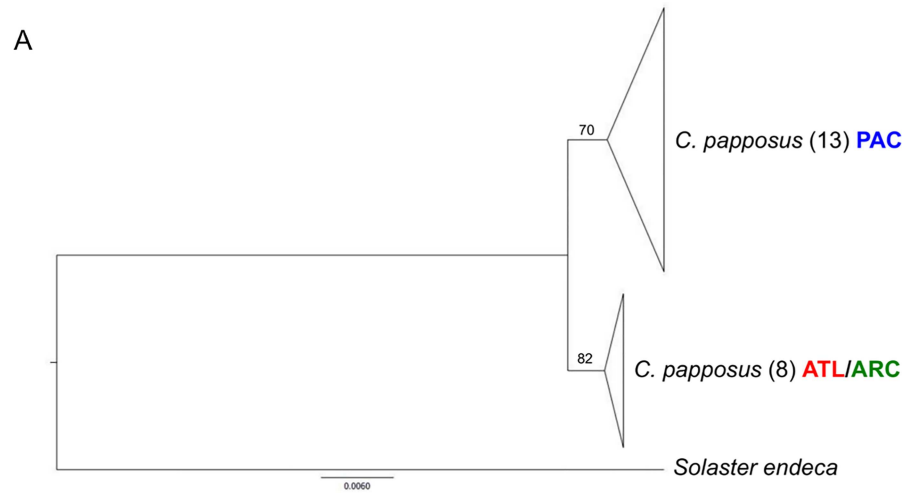
doi:10.1371/journal.pone.0166118.t002

### Species delineation

This study further validates the effectiveness of DNA barcoding as a tool for species discrimination in echinoderms. All 141 species examined in this study possessed diagnostic sequence variation at COI, and discrimination of species was generally straightforward because of the large difference (25-fold) between mean intra- (0.48%) and inter- (12%) specific divergence. This pattern is consistent with other studies which employed COI to examine relationships within and between echinoderm species [14,26,56–58]. Barcode data separated all currently recognized species examined in this study, and also revealed that specimens of six species were assigned to two or three BINs, suggesting possible overlooked taxa. Moreover, COI revealed evidence of phylogeographic patterns in five of the 19 species analyzed from more than one ocean. Because sample sizes for the 12 species with sequence divergences above 2% were small, further investigations are needed to clarify the taxonomic status of their component lineages. However, the present data do suggest that the current taxonomic system has not overlooked many species.

### Patterns of intraspecific divergence between allopatric populations

All nine species (*Ctenodiscus crispatus*, *Cucumaria frondosa*, *Florometra serratissima*, *O. sarsii*, *Psolus fabricii*, *Psolus phantapus*, *S. droebachiensis*, *S. pallidus*) which lack population structure across two or three of Canada’s oceans possess planktonic larval stages capable of long-range dispersal. However, three other species (*G. arcticus*, *P. militaris*, *S. endeca*) with planktonic larvae showed high intraspecific variation (max. 2–3%) at COI and clear geographic population structure. *G. arcticus* produces planktotrophic larvae while those of *S. endeca* are lecithotrophic [39]. *P. militaris* has a more complex reproductive strategy because females brood some



**Fig 6. Divergences at COI for three echinoderm species.** Maximum-likelihood (HKY+I) trees illustrating population structure at COI in three species of Canadian echinoderms: A) *Crossaster papposus*, B) *Pteraster militaris* and C) *Solaster endeca*. Sequence number is presented in brackets for each clade. Scale bars represent percent sequence divergence.

doi:10.1371/journal.pone.0166118.g006

young, but broadcast spawn others [59]. Because these three species produce planktonic larvae, little regional differentiation would be expected among their populations, suggesting that the deep divergences might indicate overlooked taxa. Since barcode records were obtained from both species of *Gorgonocephalus* (*G. arcticus*, *G. eucnemis*) known from Canada [60], the oversight of a known species cannot explain the deep divergence detected in *G. arcticus*. The situation for the other two genera (*Pteraster*, *Solaster*) is less certain because both genera include other Canadian species that lack barcode records [61]. As a consequence, the high intraspecific divergences for *P. militaris* and *S. endeca* might reflect misidentifications. When these divergent lineages are considered as different species, mean intraspecific divergences for these two taxa drop to 0.56% and the resulting interspecific distances (2–3%), although low, are similar to those observed for COI in other closely related echinoderm species (*Echinometra*: 2–3% [30]; *Leptasterias*: 0.4–2.2%, [23]; *Patiriella* 1.1–4.3%, [29]).

The genetic divergences in *S. endeca* and *P. militaris* involve differences between Pacific and Arctic-Atlantic populations, but they are too low to reflect isolation since the 3.5mya trans-Arctic interchange based on standard rates of mtDNA evolution [24]. Instead the 2–3% divergence in these species suggests that gene flow has occurred as recently as 1–1.5 million years ago. The divergence (2.0%) between Pacific and Arctic-Atlantic populations of *S. endeca* is considerably higher than that (1.2%) in *C. papposus*, a related species with a similar mode of dispersal [39]. Moreover, because high levels of population structure have been observed over just a few tens of kilometres in marine species with low dispersal [62], the divergence in Pacific populations of *P. militaris* (Fig 6B) may represent cryptic speciation. Additional sampling is required to better understand divergences between populations in different ocean basins. In any case, Atlantic and Arctic populations of *C. papposus*, *P. militaris*, and *S. endeca* are more similar to each other than to their Pacific counterparts, a pattern noted in several other North American echinoderms [25]. Three species with maximum intraspecific divergence above 2% have population structure associated with distance, but need further investigation to clarify their taxonomic status. It is also possible that further geographic sampling particularly across the Arctic will reveal situations in which lineages occur sympatrically, allowing a test of their reproductive isolation.

### Intraspecific divergence in sympatry

Seven of the 12 species with maximum intraspecific divergences >2% have relatively narrow geographic sampling, while two (*Ophiopholis aculeata*, *O. sarsii*) were collected from two oceans. *L. clarki*, a small brooding sea cucumber collected in the NE Pacific, had a maximum intraspecific divergence of 3.07%. The single divergent sequence of *L. clarki* might represent *L. transgressor*, the other species of this genus known from this region. These closely related species are difficult to distinguish morphologically, and have occasionally been treated as synonyms [38,63]. Alternatively, this sequence could represent a new cryptic species. Although further study is required, the sequence clusters in *L. furcilliger* may represent cryptic species with differing depth preferences. In contrast, *O. aculeata* was relatively well sampled across the Arctic and Atlantic (n = 28), but no population structure was observed. Larger sample sizes with broader geographic sampling and examination of additional characters are needed to explain the higher levels of intraspecific variation observed in these species. Lastly, *L. hexactis*

was represented by two BINs in British Columbia, a result that supports prior work which has highlighted cryptic complexes in several *Leptasterias* species using both allozyme analysis [64] and other molecular techniques [65,66].

## Patterns of divergence in a species-rich genus

Preliminary results suggest that there are over 100 species in the genus *Henricia* with particularly high levels of diversity in the NE Pacific (M. Strathmann and D. Eernisse pers. comm.). Closely related species in this genus often show subtle morphological differences and a tendency to hybridize, making identifications exceedingly difficult [40,67,68]. Although *Henricia* exhibits variable reproductive strategies ranging from brood-protection to pelagic lecithotrophy [39,69,70], challenges in species identification have made it difficult to ascertain if this variation occurs between or within species. The amphi-boreal distributions of *H. cf. oculata* and *H. sp. AAB9183* suggest that these species possess pelagic larvae, but do not rule out brooding because rafting by adults may also facilitate gene flow [71]. Although the present results suggest that DNA barcoding can advance understanding of species boundaries in this genus, future work should also incorporate nuclear markers to probe for evidence of hybridization.

## Conclusions

This study represents an important first step in the development of a DNA barcode library for the Canadian echinoderm fauna. However, additional sampling is needed, particularly in the deep-sea, as almost half of the fauna occur at depths greater than 200m with approximately 30% inhabiting depths greater than 1500m [37]. Continued expansion of the barcode reference library for echinoderms will greatly enhance the utility of DNA barcoding as an identification tool in support of a range of ecological, systematic and biodiversity studies.

## Supporting Information

**S1 Table. Specimens used in this study.** List of BOLD sample IDs, process IDs, GenBank accession numbers, and institution storing for each specimen in the dataset DS-COIECH [dx.doi.org/10.5883/DS-COIECH](https://doi.org/10.5883/DS-COIECH) on BOLD. \*sequences lacking a BIN ID (PDF)

**S2 Table. Intraspecific distances.** List of 113 species with two or more individuals and their corresponding mean and maximum intraspecific distances (%K2P). (PDF)

**S3 Table. Locality information for each BIN.** List of BINs with corresponding species ID, number of individuals, and locality information. Localities from a particular province/territory are separated by a comma, while semi-colon separates those from another province/territory. Species with multiple BIN assignments are in **BOLD**. Locality codes are as follows: Alutasivik Island (AI), Amundsen Gulf (AG), Baffin Bay (BB), Bamfield (BA), Barkley Canyon (BC), Barkley Sound (BS), Beaufort Sea (BE), Cape Scott (CS), Cape St. James (CJ), Churchill (CH), Cornwallis Island (CI), Devon Island (DI), Durban Harbour (DH), Haida Gwaii (HG), Howe Sound (HS), Hudson Bay (HB), Igloolik (IG), Kyoquot Sound (KS), Labrador Sea (LS), Lancaster Sound (LA), Nanaimo (NA), Nootka Sound (NS), Prince of Wales Strait (PW), Qikiqtarjuac (QK), Quatsino Sound (QS), Resolute (RE), Sechart (SE), Somerset Island (SI), St. Andrews (SA), Vancouver Island (VI). \*sequences lacking a BIN ID. (PDF)

## Acknowledgments

We thank Bridgette Clarkston, Christina Carr, Claudia Hand, Jeremy deWaard, Jim Boutillier, Katy Hind, Robert Frank, Sandra McCubbin, Tanya Brown, Tom Sheldon, and Victoria Frank for contributing specimens and aiding in collections. Kelly Sendall kindly provided access to the echinoderm collections at the Royal British Columbia Museum while we deeply appreciate the help that Chris Mah, Melissa Frey, and Phil Lambert provided with identifications. We also thank staff at the Canadian Centre for DNA Barcoding in the CBG for their aid in sequence acquisition. Lastly, we thank O.S. Klanten and two anonymous reviewers for their helpful suggestions on earlier versions of this manuscript.

## Author Contributions

**Conceptualization:** KKSL EAC PDNH.

**Data curation:** KKSL EAC PDNH.

**Formal analysis:** KKSL EAC PDNH.

**Funding acquisition:** EAC PDNH.

**Investigation:** KKSL EAC PDNH.

**Methodology:** KKSL EAC PDNH.

**Resources:** PDNH.

**Supervision:** PDNH.

**Validation:** KKSL EAC PDNH.

**Visualization:** KKSL EAC PDNH.

**Writing – original draft:** KKSL EAC PDNH.

**Writing – review & editing:** KKSL EAC PDNH.

## References

1. Hebert PDN, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. *Proc Royal Soc B Biol Sci.* 2003; 270: 313–321.
2. Kerr KCR, Stoeckle MY, Dove CJ, Weigt LA, Francis CM, Hebert PDN. Comprehensive DNA barcode coverage of North American birds. *Mol Ecol Notes.* 2007; 7: 535–543. doi: [10.1111/j.1471-8286.2007.01670.x](https://doi.org/10.1111/j.1471-8286.2007.01670.x) PMID: [18784793](https://pubmed.ncbi.nlm.nih.gov/18784793/)
3. Kerr KCR, Lijtmaer DA, Barreira AS, Hebert PDN, Tubaro PL. Probing evolutionary patterns in Neotropical birds through DNA barcodes. *PLoS ONE.* 2009; 4: e4379. doi: [10.1371/journal.pone.0004379](https://doi.org/10.1371/journal.pone.0004379) PMID: [19194495](https://pubmed.ncbi.nlm.nih.gov/19194495/)
4. Clare EL, Lim BK, Engstrom MK, Eger JL, Hebert PDN. DNA barcoding of Neotropical bats: species identification and discovery in Guyana. *Mol Ecol Notes.* 2007; 7: 184–190.
5. Steinke D, Zemlak TS, Hebert PDN. Barcoding Nemo: DNA-based identifications for the ornamental fish trade. *PLoS ONE.* 2009; 4: e6300. doi: [10.1371/journal.pone.0006300](https://doi.org/10.1371/journal.pone.0006300) PMID: [19621079](https://pubmed.ncbi.nlm.nih.gov/19621079/)
6. Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN. DNA barcoding Australia's fish species. *Philo Trans Royal Soc B Biol Sci.* 2005; 360: 1847–1857.
7. Zemlak TS, Ward RD, Connell AD, Holmes BH, Hebert PDN. DNA barcoding reveals overlooked marine fishes. *Mol Ecol Res.* 2009; 9: 237–242.
8. Robinson EA, Blagoev GA, Hebert PDN, Adamowicz SJ. Prospects for using DNA barcoding to identify species in species-rich genera. *ZooKeys.* 2009; 16: 27–46.
9. Smith MA, Fisher BL, Hebert PDN. DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. *Philo Trans Royal Soc B.* 2005; 360: 1825–1834.

10. Barber P, Boyce SL. Estimating diversity of Indo-Pacific coral reef stomatopods through DNA barcoding of stomatopod larvae. *Proc Royal Soc B Biol Sci.* 2006; 273: 2053–2061.
11. Layton KKS, Martel AL, Hebert PDN. Patterns of DNA barcode variation in Canadian marine molluscs. *PLoS ONE.* 2014; 9: e95003. doi: [10.1371/journal.pone.0095003](https://doi.org/10.1371/journal.pone.0095003) PMID: [24743320](https://pubmed.ncbi.nlm.nih.gov/24743320/)
12. Mikkelsen NT, Schander C, Willassen E. Local scale DNA barcoding of bivalves (Mollusca): a case study. *Zool Scripta.* 2007; 36: 455–463.
13. Radulovici AE, Sainte-Marie B, Dufresne F. DNA barcoding of marine crustaceans from the Estuary and Gulf of St Lawrence: a regional-scale approach. *Mol Ecol Res.* 2009; 9: 181–187.
14. Ward RD, Holmes BH, O'Hara TD. DNA barcoding discriminates echinoderm species. *Mol Ecol Res.* 2008; 8: 1202–1211.
15. Schander C, Willassen E. What can biological barcoding do for marine biology? *Mar Biol Res.* 2005; 1: 79–83.
16. Webb KE, Barnes DKA, Clark MS, Bowden DA. DNA barcoding: a molecular tool to identify Antarctic marine larvae. *Deep-Sea Res Part II Trop Stud Ocean.* 2006; 53: 1053–1060.
17. Pegg GG, Sinclair B, Briskey L, Aspden WJ. MtDNA barcode identification of fish larvae in the southern Great Barrier Reef, Australia. *Sci Mar.* 2006; 70(S2): 7–12.
18. Puillandre N, Strong EE, Bouchet P, Boisselier MC, Couloux A, Samadi S. Identifying gastropod spawn from DNA barcodes: possible but not yet practicable. *Mol Ecol Res.* 2009; 9: 1311–1321.
19. Meyer CP, Paulay G. DNA barcoding: error rates based on comprehensive sampling. *PLoS Biol.* 2005; 3: e422. doi: [10.1371/journal.pbio.0030422](https://doi.org/10.1371/journal.pbio.0030422) PMID: [16336051](https://pubmed.ncbi.nlm.nih.gov/16336051/)
20. Harper FM, Hart MW. Morphological and phylogenetic evidence for hybridization and introgression in a sea star secondary contact zone. *Invert Biol.* 2007; 126: 373–384.
21. Harper FM, Addison JA, Hart MW. Introgression versus immigration in hybridizing high-dispersal echinoderms. *Evolution.* 2007; 61: 4210–2418.
22. Wares JP. Biogeography of *Asterias*: North Atlantic climate change and speciation. *Biol Bull.* 2001; 201: 95–103. doi: [10.2307/1543530](https://doi.org/10.2307/1543530) PMID: [11526068](https://pubmed.ncbi.nlm.nih.gov/11526068/)
23. Hrnicevich AW, Rocha-Olivares A, Foltz DW. Phylogenetic analysis of molecular lineages in a species-rich subgenus of sea stars (*Leptasterias* subgenus *Hexasterias*). *Amer Zool.* 2000; 40: 365–374.
24. Palumbi SR, Kessing BD. Population biology of the Trans-Arctic exchange: mtDNA sequence similarity between Pacific and Atlantic sea urchins. *Evolution.* 1991; 45: 1790–1805.
25. Hardy SM, Carr CM, Hardman M, Steinke D, Corstorphine EA, Mah C. Biodiversity and phylogeography of Arctic marine fauna: insights from molecular tools. *Mar Biodiv.* 2011; 41: 195–210.
26. Arndt A, Marquez C, Lambert P, Smith MJ. Molecular phylogeny of eastern Pacific sea cucumbers (Echinodermata: Holothuroidea) based on mitochondrial DNA sequence. *Mol Phylo Evol.* 1996; 6: 425–437.
27. Hart MW, Byrne M, Smith MJ. Molecular phylogenetic analysis of life-history evolution in asterinid starfish. *Evolution.* 1997; 51: 1848–1861.
28. Carr CM, Hardy SM, Brown TM, Macdonald TA, Hebert PDN. A tri-oceanic perspective: DNA barcoding reveals geographic structure and cryptic diversity in Canadian polychaetes. *PLoS ONE.* 2010; 6(7): e22232.
29. Hart MW, Byrne M, Johnson SL. *Patiriella pseudoexigua* (Asteroidea: Asterinidae): a cryptic species complex revealed by molecular and embryological analyses. *J Mar Biol Assoc UK.* 2003; 83: 1109–1116.
30. Palumbi SR, Grabowsky G, Duda T, Geyer L, Tachino N. Speciation and population genetic structure in tropical Pacific sea urchins. *Evolution.* 1997; 51: 1506–1517.
31. Svitoch AA, Taldenkova EE. Recent history of the Bering Strait. *Oceanology.* 1994; 34: 400–404.
32. Dunton K. Arctic biogeography: The paradox of the marine benthic fauna and flora. *Trends Ecol Evol.* 1992; 7: 183–189. doi: [10.1016/0169-5347\(92\)90070-R](https://doi.org/10.1016/0169-5347(92)90070-R) PMID: [21236004](https://pubmed.ncbi.nlm.nih.gov/21236004/)
33. Durham JW, MacNeil FS. Cenozoic migrations of marine invertebrates through the Bering Strait region. In: *The Bering Land Bridge* (ed. Hopkins DM), pp. 326–349. Stanford: Stanford University Press; 1967.
34. Vermeij GJ. Anatomy of an invasion: the trans-Arctic interchange. *Paleobiology.* 1991; 17: 281–307.
35. Nikula R, Strelkov P, Vainola R. Diversity and Trans-Arctic invasion history of mitochondrial lineages in the North Atlantic *Macoma balthica* complex (Bivalvia: Tellinidae). *Evolution.* 2007; 61: 928–941. doi: [10.1111/j.1558-5646.2007.00066.x](https://doi.org/10.1111/j.1558-5646.2007.00066.x) PMID: [17439622](https://pubmed.ncbi.nlm.nih.gov/17439622/)
36. Dodson JJ, Tremblay S, Colombani F, Carscadden JE, Lecomte F. Trans-Arctic dispersals and the evolution of a circumpolar marine fish species complex, the capelin (*Mallotus villosus*). *Mol Ecol.* 2007; 16: 5030–5043. doi: [10.1111/j.1365-294X.2007.03559.x](https://doi.org/10.1111/j.1365-294X.2007.03559.x) PMID: [17944848](https://pubmed.ncbi.nlm.nih.gov/17944848/)

37. Palumbi SR, Wilson AC. Mitochondrial DNA diversity in the sea urchins *Strongylocentrotus purpuratus* and *S. droebachiensis*. *Evolution*. 1990; 44: 403–415.
38. Lambert P. Sea cucumbers of British Columbia, Southeast Alaska and Puget Sound. Vancouver: UBC Press; 1997.
39. Lambert P. Sea stars of British Columbia, Southeast Alaska and Puget Sound. Vancouver: UBC Press; 2000.
40. Lambert P, Austin WC. Brittle stars, sea urchins and feather stars of British Columbia, Southeast Alaska and Puget Sound. Victoria: UBC Press; 2007.
41. Kozloff EN. Marine invertebrates of the Pacific Northwest. Seattle: University of Washington; 1987.
42. Pollock LW. A practical guide to the marine animals of northeastern North America. New Brunswick: Rutgers University Press; 1998.
43. Ratnasingham S, Hebert PDN. A DNA-based registry for all animal species: The Barcode Index Number (BIN) System. *PLOS ONE*. 2013; 8: e66213. doi: [10.1371/journal.pone.0066213](https://doi.org/10.1371/journal.pone.0066213) PMID: [23861743](https://pubmed.ncbi.nlm.nih.gov/23861743/)
44. Ratnasingham S, Hebert PDN. BOLD: The Barcode of Life Data Systems ([www.barcodinglife.org](http://www.barcodinglife.org)). *Mol Ecol Notes*. 2007; 7: 355–364. doi: [10.1111/j.1471-8286.2007.01678.x](https://doi.org/10.1111/j.1471-8286.2007.01678.x) PMID: [18784790](https://pubmed.ncbi.nlm.nih.gov/18784790/)
45. Ivanova NV, Fazekas AJ, Hebert PDN. Semi-automated, membrane-based protocol for DNA isolation from plants. *Plant Mol Biol Rep*. 2008; 26: 186–198.
46. Benson DA, Boguski MS, Lipman DJ, Ostell J, Ouellette BFF, Rapp BA, Wheeler DL. GenBank. *Nuc Acids Res*. 1999; 27: 38–43.
47. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotech*. 1994; 3: 294–299.
48. Tamura K, Stecher G, Peterson D, Filipiński A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol Biol Evol*. 2013; 30: 2725–2729. doi: [10.1093/molbev/mst197](https://doi.org/10.1093/molbev/mst197) PMID: [24132122](https://pubmed.ncbi.nlm.nih.gov/24132122/)
49. Stamatakis A. RAxML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014; 30: 1312–1313. doi: [10.1093/bioinformatics/btu033](https://doi.org/10.1093/bioinformatics/btu033) PMID: [24451623](https://pubmed.ncbi.nlm.nih.gov/24451623/)
50. Guindon S, Gascuel O. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Syst Biol*. 2003; 52: 696–704. PMID: [14530136](https://pubmed.ncbi.nlm.nih.gov/14530136/)
51. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. *Nature Meth*. 2012; 9: 772.
52. Kimura M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol*. 1980; 16: 111–120. PMID: [7463489](https://pubmed.ncbi.nlm.nih.gov/7463489/)
53. Dixon P. VEGAN, a package of R functions for community ecology. *J Veg Sci*. 2009; 14: 927–930.
54. Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*. 2010; 26: 1463–1464. doi: [10.1093/bioinformatics/btq166](https://doi.org/10.1093/bioinformatics/btq166) PMID: [20395285](https://pubmed.ncbi.nlm.nih.gov/20395285/)
55. Bensasson D, Zhang D, Hartl DL, Hewitt GM. Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends Ecol Evol*. 2001; 16: 314–321. PMID: [11369110](https://pubmed.ncbi.nlm.nih.gov/11369110/)
56. Vogler C, Benzie J, Lessios H, Barber P, Worheide G. A threat to coral reefs multiplied? Four species of crown-of-thorns starfish. *Biol Lett*. 2008; 4: 696–699. doi: [10.1098/rsbl.2008.0454](https://doi.org/10.1098/rsbl.2008.0454) PMID: [18832058](https://pubmed.ncbi.nlm.nih.gov/18832058/)
57. Waters JM, O'Loughlin PM, Roy MS. Cladogenesis in a starfish species complex from southern Australia: evidence for vicariant speciation? *Mol Phylo Evol*. 2004; 32: 236–245.
58. Zigler KS, Lessios HA. Speciation on the coasts of the New World: phylogeography and the evolution of bindin in the sea urchin genus *Lytechinus*. *Evolution*. 2004; 58: 1225–1241. PMID: [15266972](https://pubmed.ncbi.nlm.nih.gov/15266972/)
59. McClary DJ, Mladenov PV. Reproductive pattern in the brooding and broadcasting sea star *Pteraster militaris*. *Mar Biol*. 1989; 103: 531–540.
60. Stöhr S. Gorgonocephalus Leach, 1815. In: Stöhr S.; O'Hara T. & Thuy B. (Eds) (2015) World Ophiuroidea database. Accessed through: Kennedy MK, L Van Guelpen, G Pohle, L Bajona (Eds.) (2015) Canadian Register of Marine Species at <http://www.marinespecies.org/Carms/aphia.php?p=taxdetails&id=123586> on 2015-11-2
61. Lambert P, Boutillier J. Deep-sea Echinodermata of British Columbia, Canada. Canadian Technical Report of Fisheries and Aquatic Sciences. Fisheries and Oceans Canada, Science Branch, Pacific Region, Pacific Biological Station, Nanaimo, British Columbia. 2011; 2929: 154p.
62. Palumbi SR. Marine speciation on a small planet. *Trends Ecol Evol*. 1992; 7: 114–118. doi: [10.1016/0169-5347\(92\)90144-Z](https://doi.org/10.1016/0169-5347(92)90144-Z) PMID: [21235975](https://pubmed.ncbi.nlm.nih.gov/21235975/)

63. Sewell MA, Thandar AS, Chia FS. A redescription of *Leptosynapta clarki* Hetling (Echinodermata: Holothuroidea) from the northeast Pacific, with notes on changes in spicule form, and size with age. *Can J Zool.* 1995; 73: 469–485.
64. Stickle WB, Foltz DW, Katoh M, Nguyen HL. Genetic structure and mode of reproduction in five species of sea stars (Echinodermata: Asteroidea) from the Alaskan coast. *Can J Zool.* 1992; 70: 1723–1728.
65. Flowers J, Foltz DW. Reconciling molecular systematics and traditional taxonomy in a species-rich clade of sea stars (*Leptasterias* subgenus *Hexasterias*). *Mar Biol.* 2001; 139: 475–483.
66. Foltz DW, Flowers J. Small-scale morphological differentiation in three brooding sea star species with limited dispersal. In: Harris L (ed) *Echinoderms 2006: Durham: Proceedings of the 12th International Echinoderm Conference.* Taylor & Francis, London, pp 167–172.
67. Clark AM, Downey ME. *Starfishes of the Atlantic.* New York: Chapman and Hall; 1992.
68. Mortensen T. *Handbook of the echinoderms of the British Isles.* Rotterdam: Clarendon Press Oxford; 1977.
69. Mercier A, Hamel JF. Depth-related shift in life history strategies of a brooding and broadcasting deep-sea asteroid. *Mar Biol.* 2008; 156: 205–223.
70. Strathmann MF. *Reproduction and development of marine invertebrates of the northern Pacific coast: data and methods for the study of eggs, embryos, and larvae.* Seattle: University of Washington Press; 1987.
71. Sponer R, Roy MS. Phylogeographic analysis of the brooding brittle star *Amphipholis squamata* (Echinodermata) along the coast of New Zealand reveals high cryptic genetic variation and cryptic dispersal potential. *Evolution.* 2002; 56: 1954–1967. PMID: [12449482](https://pubmed.ncbi.nlm.nih.gov/12449482/)