Mitochondrial genomics of gadine fishes: implications for taxonomy and biogeographic origins from whole-genome data sets

Mark W. Coulson, H. Dawn Marshall, Pierre Pepin, and Steven M. Carr

Abstract: Phylogenetic analysis of 13 substantially complete mitochondrial DNA genome sequences (14 036 bp) from 10 taxa of gadine codfishes and pollock provides highly corroborated resolution of outstanding questions on their biogeographic evolution. Of 6 resolvable nodes among species, 4 were supported by >95% of bootstrap replications in parsimony, distance, likelihood, and similarly high posterior probabilities in bayesian analyses, one by 85%–95% according to the method of analysis, and one by 99% by one method and a majority of the other two. The endemic Pacific species, walleye pollock (*Theragra chalcogramma*), is more closely related to the endemic Atlantic species, Atlantic cod (*Gadus macrocephalus*), than either is to a second Pacific endemic, Pacific cod (*Gadus macrocephalus*). The walleye pollock should thus be referred to the genus *Gadus* as originally described (*Gadus chalcogrammus* Pallas 1811). Arcto-Atlantic Greenland cod, previously regarded as a distinct species (*G. ogac*), are a genomically distinguishable subspecies within pan-Pacific *G. macrocephalus*. Of the 2 endemic Arctic Ocean genera, Polar cod (*Boreogadus*) as the outgroup to Arctic cod (*Arctogadus*) and *Gadus* sensu lato is more strongly supported than a pairing of *Boreogadus* and *Arctogadus* as sister taxa. Taking into consideration historical patterns of hydrogeography, we outline a hypothesis of the origin of the 2 endemic Pacific species as independent but simultaneous invasions through the Bering Strait from an Arcto-Atlantic ancestral lineage. In contrast to the genome data, the complete proteome sequence (3830 amino acids) resolved only 3 nodes with >95% confidence, and placed Alaska pollock outside the *Gadus* clade owing to reversal mutations in the ND5 locus that restore ancestral, non-*Gadus*, amino acid residues in that species.

Key words: gadines, *Gadus*, *Theragra*, mtDNA, mitogenomics, phylogeny, phylogeography, biogeography, systematics.

Résumé : Une analyse phylogénétique des séquences essentiellement complètes de 13 génomes mitochondriaux (14 036 pb) provenant de 10 taxons de morues et de goberges offre des réponses fortement supportées à des questions irrésolues quant à leur évolution biogéographique. Des six nœuds identifiés chez ces espèces, quatre étaient supportés dans plus de 95 % des rééchantillonnages lors d'analyses de parcimonie, de distance, et de vraisemblance tout en montrant des probabilités postérieures tout aussi élevées lors d'analyses bayesiennes. Un autre nœud était supporté à 85-95 % selon la méthode d'analyse et un dernier nœud était supporté à 99 % à l'aide d'une méthode et de façon majoritaire à l'aide de deux autres méthodes. Le goberge d'Alaska (*Theragra chalcogramma*), une espèce endémique du pacifique est plus apparenté à la morue de l'Atlantique (*Gadus morhua*), une espèce endémique de l'Atlantique, qu'à la morue du Pacifique (*Gadus macrocephalus*). Le goberge d'Alaska devrait donc plutôt appartenir au genre *Gadus* tel que cela était le cas initialement (*Gadus chalcogrammus* Pallas 1811). La morue du Groenland, une espèce arcto-atlantique, antérieurement considérée comme une espèce distincte (*G. ogac*), constitue un sous-clade distinct sur le plan génomique du *G. macrocephalus* pan-pacifique. Des deux genres endémiques de l'océan Arctique, le positionnement de la morue polaire (*Boreogadus*) en tant que groupe externe à la morue arctique (*Arctogadus*) et aux *Gadus* sensu lato est plus fortement supporté que le positionnement des *Boreogadus* et *Arctogadus* comme taxons frères. En considérant les données hydro-géographiques historiques, les auteurs présentent une hypothèse pour l'origine des deux espèces endémiques du Pacifique où celles-ci résulteraient d'invasions indépendantes mais simultanées via le détroit de Béring d'un lignage ancestral arcto-atlantique. Contrairement aux données génomiques, la séquence complète du protéome (3 830 acides aminés) n'a produit que trois nœuds à plus de 95 % de confiance et plaçait le goberge d'Alaska en dehors du clade des *Gadus* en raison de réversions au sein du locus *ND5* qui restaurent chez cette espèce la séquence ancestrale (changée chez le genre *Gadus*).

Received 29 October 2005. Accepted 28 June 2006. Published on the NRC Research Press Web site at http://genome.nrc.ca on 27 October 2006.

Corresponding Editor: R. Danzmann.

M.W. Coulson,[1,2] H.D. Marshall, and S.M. Carr. Genetics, Evolution, and Molecular Systematics Laboratory, Department of Biology, Memorial University of Newfoundland, St. John's, NL A1B 3X9, Canada.
P. Pepin. Northwest Atlantic Fisheries Centre, Department of Fisheries and Oceans, Canada, P.O. Box 5667, St. John's, NL A1C 5X1, Canada.

[1]Corresponding author (e-mail: mcoulson@dal.ca).
[2]Present address: Marine Gene Probe Laboratory, Department of Biology, Dalhousie University, Halifax, NS B3H 4J1, Canada.

 doi:10.1139/G06-083

Mots clés : gadinés, *Gadus*, *Theragra*, ADNmt, mitogénomique, phylogénie, phylogéographie, biogéographie, systématique.

[Traduit par la Rédaction]

Introduction

Studies of inter-specific molecular systematics have long employed the mitochondrial DNA (mtDNA) genome as a source of phylogenetic markers: the very first mtDNA restriction maps of vertebrate animals compared restriction sites for 2 species of clawed frogs (*Xenopus*) (Ramirez and Dawid 1978). Most studies have relied on DNA sequence analysis of one or a few genes, with different genes employed depending on the anticipated times of divergence under investigation. The central problem for phylogenetic inference in both morphological and molecular studies remains the identification of a sufficient number of synapomorphic characters to discern lineages with statistical confidence (Stepien and Kocher 1997). Despite large numbers of informative sites (sensu Nei 1987), deep lineages may be obscured by multiple transitions and transversions at the same nucleotide positions, which require model-specific corrections estimated from the data themselves. At the other extreme, very recent divergences are difficult to resolve when too few polymorphic positions are available to provide a significant phylogenetic or phylogeographic signal among or within species. The straightforward resolution of these problems is to increase the nucleotide data set to include the complete mtDNA genome, which provides 16–20 kb characters in vertebrate species. Another problem encountered with small mtDNA-based datasets is the inconsistency of phylogenies generated from different mitochondrial genes. The genomic approach serves to eliminate such discrepancies, and thus allows one to obtain the "true mtDNA tree''. "Genomic thinking'' has now arrived in evolutionary biology, and complete genomes from more than 800 species are now available through GenBank (as of April 2006).

Most mitogenomic phylogenetic studies have focused on resolution of ancient branches and rooting of trees. For example, Inoue et al. (2001) used complete mtDNA genome sequences from more than 100 fish species to identify the relationships among the 5 major basal teleostean lineages (Osteoglossomorpha, Elopomorpha, Clupeomorpha, Ostariophysi, and Protacanthopterygii). Their work resolved all but one internal branch with high statistical support (95%–100%) and suggested a topology that challenged a number of previously existing theories. Their work demonstrates the value of genomic systematics at the level of classes, orders, and higher taxa. More recently, mitogenomic studies have also focused on more intrafamilial and intrageneric relationships (e.g., Davis et al. 2004; Minegishi et al. 2005). The present study extends mitogenomics to the commercially and environmentally important family Gadidae, which includes cod, haddock, pollock, and whiting.

Carr et al. (1999) investigated evolutionary relationships among 14 species of gadid fishes based on 0.9 kb of 2 mitochondrial genes. These data indicated that there have likely been 3 independent invasions of the Pacific Basin following the re-opening of the Bering Strait (by Pacific tomcod (*Microgadus proximus*), Alaska pollock (*Theragra chalcogramma*), and Pacific cod (*Gadus macrocephalus*)), and that Greenland cod (*Gadus ogac*) were genetically indistinguishable from Pacific cod. Left unresolved, inter alia, were questions about the biogeographic relationships among *Gadus* spp., *Theragra*, and their sister taxa, the 2 polar basin species, *Boreogadus* and *Arctogadus*. The limits of resolution can be seen in a comparison of the bootstrap analyses by maximum-likelihood, distance, and maximum-parsimony methods (see Fig. 2 of Carr et al. 1999). Of the 10 internal nodes possible for a fully bifurcating tree with 13 terminal taxa (synonymizing Greenland cod with *G. macrocephalus*), bootstrap support greater than 50% was obtained for 7 or 8 nodes, with the 3 methods in agreement on 5 of them. Bootstrap support greater than 95% was obtained for no more than 4 of the 10 nodes, and only a single node was supported at this highest level of consensus and consistency by all 3 methods. Subsequently, Møller et al. (2002) added data on *Arctogadus* and Pogson and Mesa (2004), based on 2 nuclear loci, and Teletchea et al. (2006), based on 1530 bp of mtDNA sequence, both obtained improved bootstrap for some of these relationships (see Discussion); however, discrepancies still remain and neither study focused on the biogeographic and taxonomic questions raised by the phylogeny of this group.

We present an analysis of substantially complete sequences (less D-loop control regions) from 10 species in 8 genera of gadid fishes, as well as duplicate genomes from 3 geographically widespread species in the genus *Gadus*. The present phylogenomic study provides a practical assessment of the value of expanded data sets provided by genomic approaches to the systematics of closely-related species, and in so doing resolves biogeographic and systematic questions left unsettled by Carr et al. (1999), Pogson and Mesa (2004), and Teletchea et al. (2006).

Materials and methods

Specimens of *Boreogadus saida*, *Gadus macrocephalus ogac*, and *Gadus morhua* were collected on the Newfoundland shelf between 1991 and 1993 by the Department of Fisheries and Oceans (DFO), St. John's, Nfld. *Theragra chalcogramma*, *Microgadus proximus*, and *Gadus macrocephalus* were collected off the west coast of British Columbia (DFO, Nanaimo, B.C.). *Arctogadus glacialis* was collected off the east coast of Greenland (DFO, St. John's, Nfld.), *Melanogrammus aeglefinus* and *Pollachius virens* were collected off George's Bank (National Marine Fisheries Service, Woods Hole, Mass.), and *Merlangius merlangus* was collected in the North Sea (Centre for Environment, Fisheries and Aquaculture Science, Lowestoft, UK).

DNA was extracted from frozen muscle or heart tissue with the QIAamp DNA Mini Kit Tissue Protocol (Qiagen Inc., Valencia, Calif.). Tissue samples were digested in

180 μL of Buffer ATL (Qiagen Inc.) with 20 μL of proteinase K and incubated at 56 °C for 2–4 h or until all tissue was completely lysed. In some cases, lysis continued overnight. For removal of DNA from the lysis mixture, the manufacturer's protocol was followed. DNA was eluted in 200 μL of Buffer AE (Qiagen Inc.) and subsequently stored at –20 °C.

Appropriate primer sequences for complete mitochondrial genome amplification were identified by comparison of the complete mtDNA genome sequence of a Norwegian *G. morhua* (GenBank accession No. NC002081) to 3 outgroup orders, a pleuronectiform (*Paralichthys olivaceus*, accession No. NC002386), a salmoniform (*Salmo salar*, accession No. NC001960), and a perciform (*Trachurus japonicus*, accession No. NC002813). Strings of 25 bp with more than 96% homology among these 4 taxa were chosen as primer sites. A set of 20 primer pairs were identified that amplify the mitochondrial genome in fragments of 750–1400 bp, with overlap between adjacent fragments of 80–300 bp. Table 1 lists the 40 primers, the regions amplified, and the product size (bp) expected for each amplification. The control region of *Gadus* includes a region of short tandem repeats (Johansen and Bakke 1996) that is difficult both to amplify and sequence. Alignment of this region between even closely related species is of uncertain accuracy. We have therefore excluded the control region from this analysis (cf. Otto et al. 1996).

Polymerase chain reaction (PCR) amplifications were performed in 25 μL volumes containing the following ingredients: 2.5 μL of 10× PCR buffer (Qiagen Inc.), 0.5 μL of 10 mmol/L dNTPs (2.5 mmol/L each of dATP, dCTP, dGTP and dTTP), 1.0 μL of a 10 μmol/L stock of both forward and reverse primer (Operon Technologies Inc., Huntsville, Ala.), 0.2 μL of 5 U/μL *Taq* DNA polymerase (Qiagen Inc.), and 19 μL of ddH_2O. Following an initial denaturation at 95 °C for 3 min, the amplification profile consisted of the following steps: denaturation at 93 °C for 30 s, annealing at 48–60 °C (depending on the primer pair) for 30 s, and extension at 72 °C for 1 min. This was repeated for 40 cycles. A final incubation was carried out at 72 °C for 10 min. PCRs were performed on a GeneAmp PCR 9600 thermal cycler (Perkin Elmer, Wellesley, Mass.). Amplification products were purified with a QIAQuick PCR Purification Kit (Qiagen Inc.) as per the manufacturer's instructions. Purified DNA was eluted in 50 μL of elution buffer (10 mmol/L Tris–HCl, pH 8.5) and stored at –20 °C.

PCR products were sequenced with the ABI Prism BigDye Terminator Cycle sequencing chemistry version 2.0 (Applied Biosystems, Inc., Foster City, Calif.) using the same primers as for PCR amplification. Sequencing reactions were performed on a T1 DNA thermal cycler (Perkin Elmer) as follows: an initial denaturation at 96 °C for 2 min, followed by 35 cycles consisting of 96 °C for 30 s, annealing of primers at 50 °C for 15 s, and extension of products at 60 °C for 4 min. Sequencing reactions were purified by isopropanol precipitation, and data were collected on an ABI Prism 377 DNA sequencer.

Sequences were aligned and edited with the computer software program Sequencher™ version 4.1.2 (Gene Codes Corporation, Ann Arbor, Mich.). DNA sequences reported here are the consensus sequences of the 2 complementary strands of each region for each individual. Overlapping regions enabled the assembly of contiguous fragments. Consensus sequences of the mtDNA genome for each individual were exported for phylogenetic analysis. Mitochondrial genome sequences were submitted to GenBank and have been assigned the accession numbers DQ356935–DQ356946, respectively, for *Arctogadus glacialis*, *Boreogadus saida*, *Gadus macrocephalus* (west Pacific), *G. macrocephalus* (east Pacific), *Gadus morhua* (west Pacific), 2 individual *G. m. ogac* specimens, *Melanogrammus aeglefinus*, *Merlangius merlangus*, *Microgadus proximus*, *Pollachius virens*, and *Theragra chalcogramma*.

Maximum parsimony, neighbour-joining, and maximum-likelihood analyses were performed with the Phylogenetic Analysis Using Parsimony (PAUP*) program (version 4.0; Swofford 1998). For parsimony analysis, a heuristic search was used with transversion–transition (Tv/Ts) weightings of 1:1, 3:1, 10:1, and transversion only, and with 10 random taxon additions and the tree bisection and reconnection (TBR) branch-swapping option. Bootstrap support of parsimony trees was estimated with 1000 replicates each of 10 random taxon additions with the TBR branch-swapping option. For the neighbour-joining analysis, pairwise genetic distances were calculated with Tamura–Nei (T–N) distance measure (Tamura and Nei 1993) and the tree was constructed with a neighbour-joining search with an estimated γ parameter (0.971). Bootstrap support was estimated with 1000 replicates of this algorithm. Maximum-likelihood analysis used the TrN+I+G model of nucleotide substitution with a correction for among-site rate variation (0.971) and the proportion of assumed invariable sites set at 0.662. This model was chosen based on log-likelihood ratio tests as implemented in ModelTest version 3.06 (Posada and Crandall 1998). The starting tree was obtained with a heuristic search with 10 random taxon additions and the TBR branch-swapping option. The tree was bootstrapped with a heuristic search via stepwise addition with a single random taxon addition, for each of 1000 replicates. As a test of power, a series of bootstraps were performed for each method of tree reconstruction that resampled successively larger portions of the complete dataset. For this analysis, bootstrapping analysis was performed by resampling 1000, 2000, 4000, and 8000 base pairs of the complete data set. At each of these levels of resampling, the number of nodes resolved with both 50% and 95% bootstrap support was determined.

Bayesian analysis of phylogeny with Markov chain Monte Carlo sampling was conducted with MrBayes 3.12 (Ronquist and Huelsenbeck 2003).We used the general time-reversible (GTR), Felsenstein81 (F81), and Hasegawa–Kishino–Yano (HKY) models of DNA substitution, γ-distributed rate variation across sites, and a proportion of invariant sites. The default search method (1 cold and 3 heated chains) was run simultaneously for 1 000 000 generations, and 1 in 10 trees were sampled. The first 25 000 trees were discarded, and Bayesian posterior probabilities were estimated from a 50% majority-rule consensus of the remaining 75 000 trees.

Each protein-coding gene was translated into its amino acid sequence, and the concatenated protein sequences (3830 amino acids) were subjected to both parsimony and neighbour-joining analyses based on absolute character differences. This allowed comparison of phylogenetic resolu-

Table 1. Sequences of the 20 primer pairs used to amplify entire mitochondrial genomes in overlapping regions, the expected size (bp) of the region, bounds of the region amplified as numbered in the *Gadus morhua* reference sequence (NCBI NC002081), and the gene regions included in the amplification product.

Primer designation	Primer sequence (5′→3′)	Expected size of PCR product (bp)	Primer bounds	Gene region amplified
g01	F: 5′-CTGAAGATATTAGGATGGACCCTAG-3′ R: 5′-CTAGTCCCTACTTACTGCTAAATCC-3′	841	29–875	*tPhe / 12S*
g02	F: 5′-CCAAAAACGTCAGGTCGAGGTGTAG-3′ R: 5′-CTATTCATTTCACAGGCAACCAGCT-3′	750	743–1515	*12S / 16S*
g03	F: 5′-ACCCCGAAACTGAGCGAGCTACTCC-3′ R: 5′-TAAGCCCTCGTGATGCCATTCATAC-3′	800	1373–2199	*16S*
g04	F: 5′-TTTACCAAAAACATCGCCTCTTG-3′ R: 5′-TGAACCTCTGTAGAAAGGGCTTAGG-3′	801	2068–2896	*16S / tLeu*
g05	F: 5′-GGAGTAATCCAGGTCAGTTTCTATCTATG-3′ R: 5′-ATGTTCGGGGTATGGGCCCAAGAGC-3′	1357	2711–4080	*16S / tMet*
g06	F: 5′-GGTTAAAGTCCCTTCAACTCCTTAG-3′ R: 5′-AGCTTAATTAAAGTATTTGTTTTGC-3′	1337	3945–5293	*tIle / tAla*
g07	F: 5′-AAACTAGACCAAGGGCCTTCAAAGC-3′ R: 5′-GCTAATCAGCTAAAGACTTTTACACC-3′	1341	5171–6526	*tTrp / COI*
g08	F: 5′-ATGGGTATAGTCTGAGCTATGATGG-3′ R: 5′-TAACCCACAATTCTGCCTTGACAAG-3′	865	6365–7244	*COI / tAsp*
g09	F: 5′-TAACTAACGTTGAGTGACTCCAC-3′ R: 5′-ACCCATATTAGCTTCTTAGTGAGG-3′	954	7019–7989	*COI / tLys*
g10	F: 5′-TCCCGGAGTTTTCTACGGACAATG-3′ R: 5′-AGAGGGCGAATGAATAAACTAATTG-3′	827	7840–8666	*COII / Atp6*
g11	F: 5′-TAGCAACTGTCCTTATCGGCATACG-3′ R: 5′-TAATACTGTGGTGAGCTCAGGTTAC-3′	805	8534–9339	*Atp6 / COIII*
g12	F: 5′-CCCCGAACTAGGTGGCTGCTGACC-3′ R: 5′-AGGACAAGTAGAGGTGTTGATAGGG-3′	1441	9214–10658	*COIII / ND4*
g13	F: 5′-CTTTCTCCGCTTGTGAAGCAAG-3 R: 5′-CAATTAGAGATTTCAGGTCAGTTTG-3′	961	10353–11314	*ND4L / ND4*
g14	F: 5′-CTGTTGCAGGCTCAATAGTTCTTGC-3′ R: 5′-TTCGAGGGAGCCTTGGGGTCTAACC-3′	847	11126–11976	*ND4 / tSer*
g15	F: 5′-TAACCAAGACATTAGATTGTGATTC-3′ R: 5′-TGGTAGTCATGGGTGAAGTCCAAAC-3′	923	11846–12772	*tHis / ND5*
g16	F: 5′-GGTGATGACACGGCCGAGCAGATG-3′ R: 5′-AATAATTGCATCTTTGGAGAAGAAGC-3′	723	12533–13256	*ND5*
g17	F: 5′-ATTCATAGCCTAAACGATGAACAAG-3′ R: 5′-GTCGTTTTTCATATCATTAGTCCTG-3′	1331	13109–14440	*ND5 / tGlu*
g18	F: 5′-GCTACTAAGACCAGTCCTAAAGCAG-3′ R: 5′-CTGTGGGATTATTTGAGCCTGTTTC-3′	806	14304–15110	*ND6 / Cytb*
g19	F: 5′-GAGGAGGTTTCTCAGTAGATAATGC R: 5′-GTTTAATTTAGAATTCTAGCTTTGG	851	14976–15828	*Cytb / tPro*
g20	F: 5′-GAATGAAACTGCCCTAGTAGCTCAG-3′ R: 5′-GGCAGGACATTAAGGGCATTCTCAC-3′	1383	15614–160	*tThr / 12S*

tion between nucleotide and protein data. As with the nucleotide data, a power analysis was performed by resampling random subsets of 250, 500, 1000, 2000, and all 3830 amino acids.

Results

Based on alignment to the complete mtDNA genome sequence of *Gadus morhua* (16 696 bp) (Johansen and Bakke 1996) as a reference, and excluding the D-loop control region, the consensus alignment of mtDNA genome sequences from 13 individuals in 10 nominal taxa was 15 720 bp. Excluding a further 1684 bp that were missing or ambiguous in one or more taxa, the number of sites common to all 13 genomes is 14 036 bp.

Among these, 11 290 positions were constant across all taxa and 2746 sites were variable, of which 1631 nucleotides were parsimony informative (sensu Nei 1987). Estimated base frequencies are A (28.9%), C (25.9%), G (15.7%), and T (29.5%). As is typical of mtDNA, there is a lower G+C content (41.6%), with a particular bias against G, which is avoided in 3rd codon positions (Saccone et al. 1999).

Table 2 shows the genetic distances, calculated as absolute pairwise nucleotide differences and Tamura–Nei (T–N) distances (Tamura and Nei 1993). For the 3 pair-

Table 2. Pairwise number of nucleotide differences (lower half of the matrix) and Tamura–Nei distances (upper half of the matrix) calculated from mitochondrial gene sequences for 10 gadine species.

	G. morhua 002	*G. morhua* Norway	*G. m. ogac* 001	*G. m. ogac* 002	*G. macrocephalus* 002	*G. macrocephalus* 003	*T. chalco gramma*
G. morhua 002	—	0.0036	0.042	0.042	0.043	0.042	0.040
G. morhua Norway	51	—	0.043	0.043	0.044	0.043	0.042
G. m. ogac 001	567	578	—	0.0011	0.0039	0.0044	0.043
G. m. ogac 002	570	581	16	—	0.0032	0.0040	0.043
G. macrocephalus 002	577	588	55	45	—	0.0049	0.043
G. macrocephalus 003	572	585	62	56	69	—	0.042
T. chalcogramma	545	562	576	575	579	574	—
A. glacialis	613	623	584	586	588	592	614
B. saida	729	745	697	696	695	706	693
M. aeglefinus	1166	1165	1174	1174	1183	1191	1147
M. merlangus	1088	1091	1096	1095	1098	1110	1064
P. virens	1168	1175	1155	1154	1147	1158	1146
M. proximus	1332	1345	1280	1284	1287	1299	1278
	A. glacialis	*B. saida*	*M. aeglefinus*		*M. merlangus*	*P. virens*	*M. proximus*
G. morhua 002	0.045	0.054	0.090		0.084	0.091	0.104
G. morhua Norway	0.046	0.056	0.090		0.084	0.091	0.105
G. m. ogac 001	0.043	0.052	0.091		0.084	0.089	0.099
G. m. ogac 002	0.043	0.052	0.091		0.084	0.089	0.100
G. macrocephalus 002	0.044	0.052	0.091		0.084	0.088	0.100
G. macrocephalus 003	0.044	0.053	0.092		0.085	0.089	0.101
T. chalcogramma	0.046	0.052	0.089		0.082	0.088	0.100
A. glacialis	—	0.045	0.088		0.080	0.088	0.099
B. saida	604	—	0.090		0.082	0.090	0.101
M. aeglefinus	1139	1165	—		0.074	0.097	0.113
M. merlangus	1033	1075	969		—	0.093	0.109
P. virens	1135	1168	1242		1200	—	0.100
M. proximus	1275	1296	1427		1392	1283	—

Fig. 1. Phylogenetic trees for 10 taxa of gadine fish, based on substantially complete mitochondrial DNA genomes. Trees were produced by (A) unweighted maximum parsimony (MP), (B) neighbor-joining based on Tamura–Nei (T–N) distances, and (C) maximum-likelihood (ML) analysis. Trees are based on the consensus of 14 036 bases. Results of the bootstrap analyses are indicated in bold.

wise comparisons within species, the 2 *G. m.* subsp. *ogac* individuals are most similar (16 nucleotide differences, T–N distance = 0.0011), and the two *G. morhua* individuals (one from either side of the Atlantic) next most similar (51 nucleotide differences, T–N distance = 0.0036). The 2 *G. m.* subsp. *macrocephalus* individuals (one from either side of the Pacific) were least similar (69 nucleotide differences, T–N distance = 0.0049); the range of differences within *G. macrocephalus* sensu lato was 45–62 nucleotide differences (T–N distance = 0.0032–0.0044).

The smallest intergeneric difference was between *Theragra* and *G. morhua* (545 nucleotide differences, T–N distance = 0.040), slightly less than the extreme difference within *Gadus* (588 nucleotide substitutions, T–N distance = 0.044). The maximum intergeneric differences observed were between *Microgadus* and *Melanogrammus* (1427 nucleotide substitutions, T–N distance = 0.113).

Genomic phylogenetic results

Carr et al. (1999) identified *Trisopterus* spp. as the sister group to the remaining gadines examined and *Microgadus* spp. sensu lato (including *Eleginus*) as the sister group to the remaining 6 genera (not including *Arctogadus*; but see Møller et al. 2002), though its relationship to *Pollachius* was not supported by bootstrap analysis. However, Pogson and Mesa (2004) and Teletchea et al. (2006) confirm these relationships of *Trisopterus* and *Microgadus* to the remaining genera. Accordingly, *Microgadus proximus* was designated as the outgroup. Analysis by maximum parsimony (all sites weighted equally) (Fig. 1A), neighbour-joining (Fig. 1B), maximum-likelihood (Fig. 1C), and Bayesian (not shown) methods produced similar results, which lend support to the consensus tree as the parametric, or true mitochondrial tree for these genomes (cf. Otto et al. 1996). All intergeneric branches were strongly supported, typically by 100% of bootstrap replicates in the first 3 analyses (Fig. 1), and posterior probabilities of 1.00 by the Bayesian analysis (not shown).

All approaches identified *G. morhua* and *Theragra* as sister taxa: parsimony, distance, and likelihood bootstrap supports were 95%, 93%, and 85%, respectively, and the Bayesian posterior probability was 1.00. All approaches also identified the 2 *G. m. ogac* individuals as sister OTUs, with bootstrap supports of >95% and a posterior probability of 1.00, and indicated that this pair was most closely related to the eastern Pacific *G. macrocephalus*.

The 4 modes of analysis differed only with respect to the relative branching order of *Boreogadus* and *Arctogadus* with respect to *Gadus* spp. and *Theragra*. In the equally-weighted parsimony analysis, *Boreogadus* was the sister species to *Arctogadus* and *Gadus* sensu lato (including *Theragra*) with high bootstrap support (99%). With weighted transversion–transition ratios of 3:1 and 10:1, *Boreogadus* was placed outside *Arctogadus* with strong support (99% and 87%, respectively). Transversion-only parsimony analysis identified 2 trees of identical length, one with a topology as above and the other with *Boreogadus* and *Arctogadus* paired as sister taxa. Likelihood analysis also indicated *Boreogadus* and *Arctogadus* as sister taxa; however, bootstrap analysis collapsed these 2 species at a polytomy, with equal bootstrap support for this relationship and the one implied by parsimony. The topology of the tree produced by neighbour-joining analysis also showed these 2 species as sister taxa, but with low bootstrap support (53%), and resolved *Boreogadus* outside *Arctogadus*. The Bayesian analysis with the GTR and HKY models paired *Boreogadus* and *Arctogadus* as sister taxa, with posterior probabilities of 0.73 and 0.79, respectively, in contrast to the F81 model, which placed *Boreogadus* outside *Arctogadus* with a posterior probability of 1.00.

To reduce the influence of homoplasy introduced by the inclusion of more distantly related taxa, we conducted an analysis of a restricted set of 7 OTUs, comprising *Melanogrammus* and *Merlangus* as the designated outgroup to *Boreogadus*, *Arctogadus*, eastern Pacific *G. macrocephalus*, and *Theragra*, and the western Atlantic *G. morhua*. With transversion–transition weightings of 1:1, 3:1, 10:1, and transversions only, parsimony analysis placed *Boreogadus* outside *Arctogadus* + *Gadus* sensu lato with bootstrap support of 97%, 98%, 90%, and 71%, respectively. This result seems to be due to the occurrence of a large number of transitions with high consistency indices that favour *Boreogadus* as the outgroup, as compared with much smaller sets of transversions that favour either this hypotheses or its alternative, but with low consistency. Neighbor-joining analysis with GTR likelihood distances placed *Boreogadus* outside *Arctogadus* with support from a bare majority of bootstrap replications (5010/10000), as did an analysis with Tamura–Nei distances (γ = 0.66) in 64% of replications. Likelihood analysis with a GTR model paired *Boreogadus* and *Arctogadus* as sister taxa, but with bootstrap support of only 55%. The Bayesian analysis with the GTR and HKY models paired *Boreogadus* with *Arctogadus*, as in the full analysis, but with reduced posterior probabilities of 0.63 and 0.53, respectively; the F81 model again placed *Boreogadus* outside *Arctogadus*, with a posterior probability of 1.00.

In summary, parsimony analysis always favoured *Boreogadus* as the outgroup to *Arctogadus* + *Gadus* sensu lato (including *Theragra*), typically with strong bootstrap support; likelihood analysis always favoured *Boreogadus* and *Arctogadus* as sister taxa, but with weak bootstrap support; distance analysis favoured the latter hypothesis in the full taxon set and the former hypothesis in the reduced set, both with weak bootstrap support; finally, Bayesian analysis favoured either the latter hypothesis with weak posterior probability or the former with a high posterior probability, according to the model used.

Carr et al. (1999) reviewed evidence that *G. ogac* is conspecific with *G. macrocephalus*, and should be synonymized with it as *G. macrocephalus* subsp. *ogac*. The present analysis consistently pairs the 2 individual *G. m. ogac* specimens

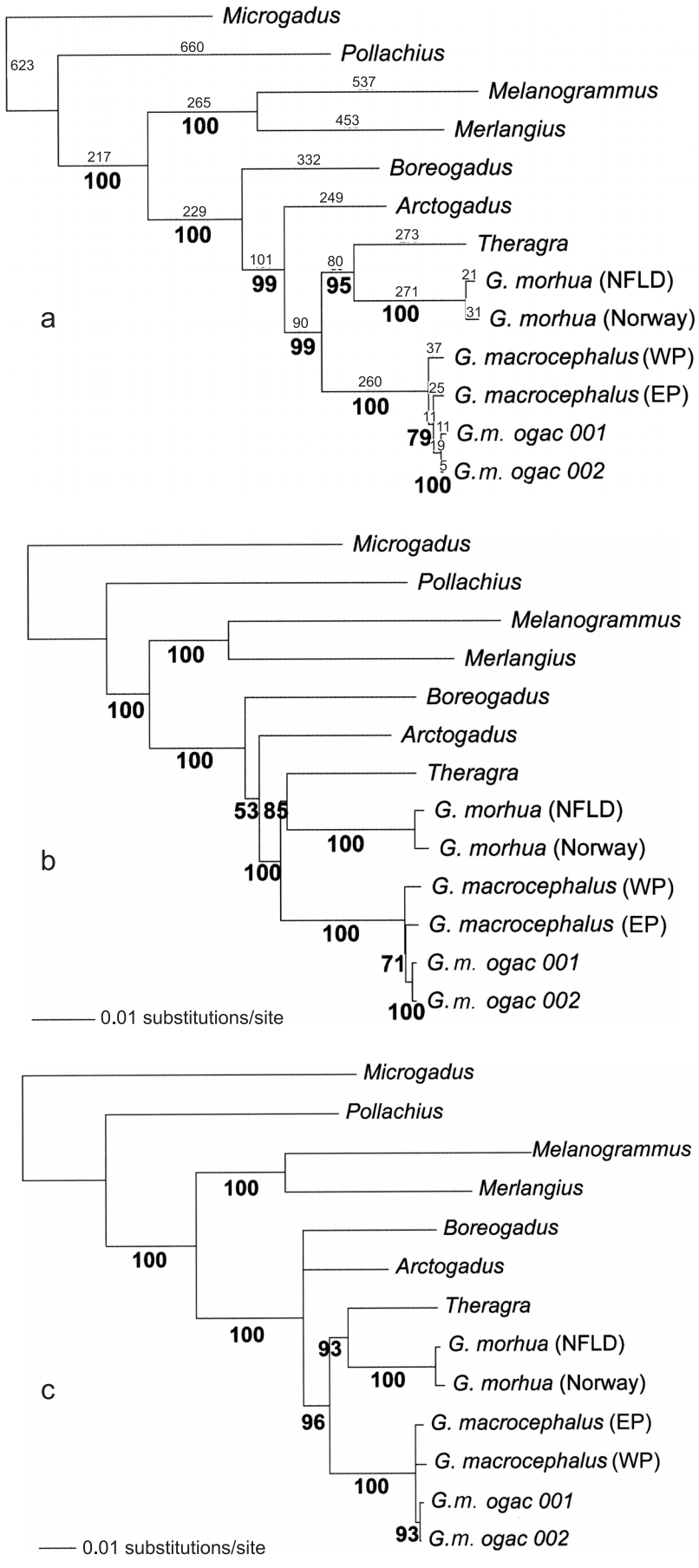


Table 3. Power analysis of mtDNA nucleotide data showing the number of nodes resolved with ≥50% and ≥95% bootstrap support (based on 1000 replicates) for successively larger resampling of the databases on 14036 bp.

		Bootstrap support	
No. of base pairs resampled	No. of nodes resolved	Nodes ≥ 50%	Nodes ≥ 95%
Maximum parsimony (MP)			
1000	6	8	3
2000	8	10	5
4000	10	10	5
8000	10	10	8
14036	10	10	9
Neighbour-joining (NJ)			
1000	7	7	3
2000	9	9	4
4000	10	10	5
8000	10	10	7
14036	10	10	7
Maximum likelihood (ML)			
1000	6	6	2
2000	8	8	3 (1 at 94%)
4000	8	8	5
8000	8	8	5
14036	8	8	6

Table 4. Power analysis of mitochondrial protein data showing the number of nodes resolved with ≥50% and ≥95% bootstrap support (based on 1000 replicates) for successively larger resampling of the databases on 3830 amino acids.

		Bootstrap support	
No. of amino acids resampled	No. of nodes resolved	Nodes ≥ 50%	Nodes ≥ 95%
Maximum parsimony (MP)			
250	0	0	0
500	0	0	0
1000	3	3	0
2000	7	7	0
3830	8	8	1
Neighbour joining (NJ)			
250	0	0	0
500	1	1	0
1000	6	6	0
2000	7	7	0
3830	10	10	3

as most closely related, and the parsimony and distance analyses support a closer relationship of this pair with the eastern Pacific *G. macrocephalus* (79% and 71%, respectively, in Figs. 1A and 1B). The likelihood analysis did not resolve relationships of *G. m. ogac* with *G. macrocephalus* (Fig. 1C). When the analysis was limited to the 7 *Gadus* and *Theragra* sequences, unweighted parsimony, distance, and likelihood analyses all showed greater than 95% bootstrap support for a closer relationship of the *G. m. ogac* pair with eastern Pacific *G. macrocephalus* (results not shown).

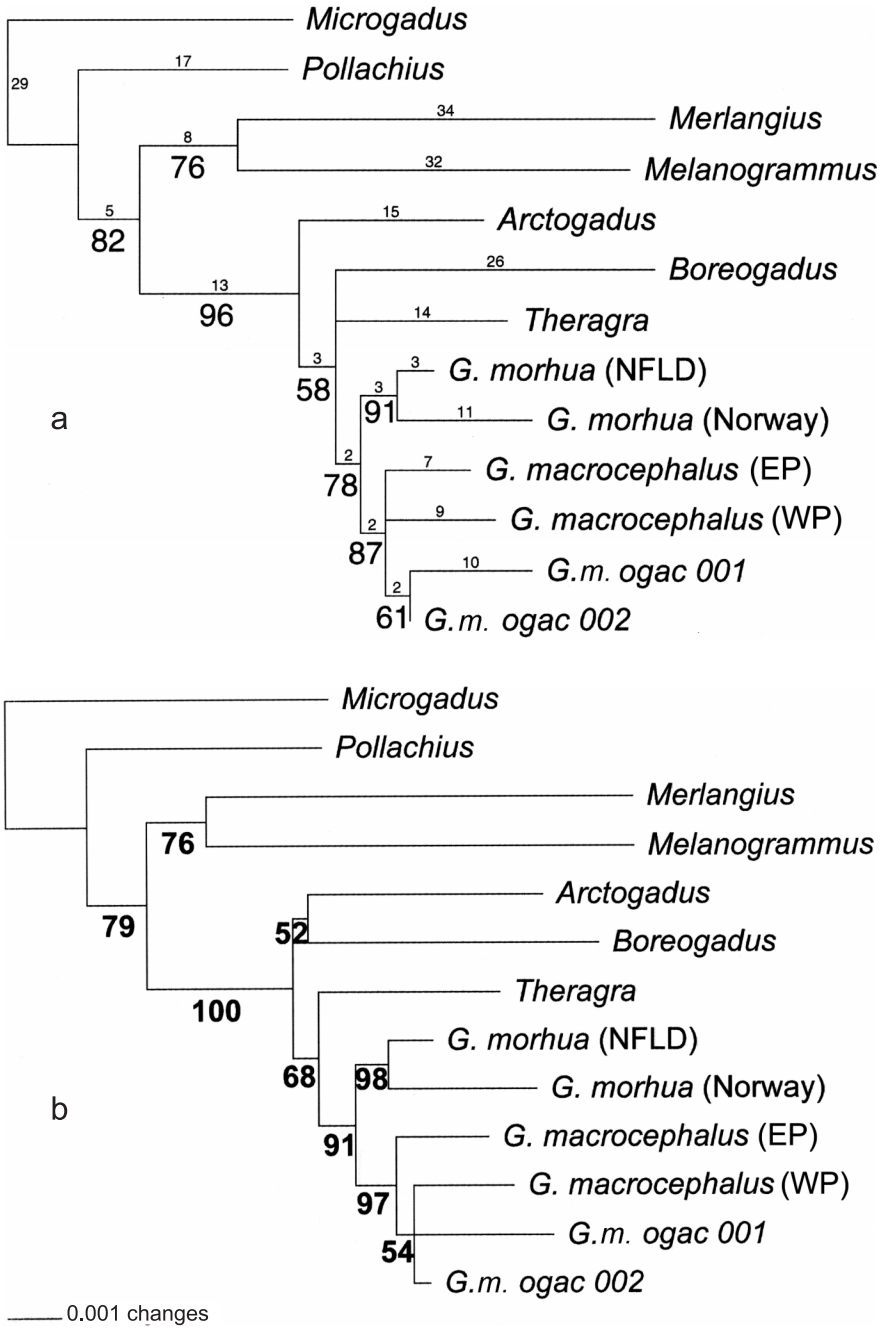
Bootstrap resampling with successively larger subsets of nucleotides resulted in increased resolution of nodes, as well as the number of nodes reaching ≥95% statistical support (Table 3). Statistical support beyond 95% did not increase after ~8000 bp, across all modes of phylogeny reconstruction and by about 4000 bp, no additional nodes were resolved. A detailed analysis of patterns of mtDNA sequence evolution, including comparisons of rates and phylogenetic efficiency among individual mtDNA genes, will be presented elsewhere. [3]

Proteomic phylogenetic analysis

The parsimony and distance analyses on 3830 amino acids gave results generally consistent with nucleotide analysis, but with lower bootstrap support. Any reasonable degree of phylogenetic resolution required at least 2000 amino acids, and even the full complement of protein sequences identified either 1 (parsimony) or 3 (distance) nodes supported in more than 95% of bootstraps (Table 4). Two differences occur between the topologies of the nucleotide and amino acid trees. In the amino acid analyses, *Arctogadus* is either the sister species to *Boreogadus* and *Gadus* sensu lato (parsimony) or else a sister species to *Boreogadus* only (distance); both alternatives are only weakly supported (58% and 52%, respectively). These same relationships were also poorly resolved in the nucleotide data. A more striking difference between the phylogenies implied by the nucleotide and amino acid data are the placement of *Theragra* (= *G. chalcogrammus*) with respect to *Gadus* sensu stricto. Whereas *Theragra* is the sister species to *G. morhua* in the nucleotide analysis, both the parsimony and distance analyses of the amino acid data place *Theragra* entirely outside *Gadus* as its sister species, with high bootstrap support (78% and 91%, respectively; Fig. 2). Inspection of the data identified an unusual pattern of amino acid substitution in the *ND5* gene. Phylogenetically informative amino acid patterns that identify *Arctogadus* – *Boreogadus* – *Gadus* spp. as a clade occur at 6 residues in the last 150 positions of this locus. However, 2 of these apparently synapomorphic sites are clearly homoplasic reversals (Pro→Ser and Ala→Thr) caused by C→T and G→A first-position transitions, respectively, that render *G. chalcogrammus* (= *Theragra*) identical to the rest of the outgroup species, rather than to the balance of *Gadus* sensu stricto, as expected from the overall analysis (Fig. 3). Other homoplasies are evident in this region. We estimated selection pressure on the *ND5* gene by a codon model of substitution (Goldman and Yang 1994) as implemented in the program PAML (Yang 1997). This analysis shows a highly biased transition–transversion rate ratio (7.01) as well as a biased usage of nucleotides at the 3 positions of the codon corresponding to unequal usage of synonymous codons in this gene. Likelihood ratio testing indicated strong support for variable selection pressures

[3] Coulson, M.W., Marshall, H.D., and Carr, S.M. Tempo and move of evolution of mtDNA genomes among gadid fishes: implications for phylogenetic accuracy and effects of selection. Manuscript in preparation.

Fig. 2. Phylogenetic trees for 10 taxa of gadine fish, based on complete amino acid proteomes. Trees were produced by (A) maximum parsimony and (B) neighbour joining. Trees are based on the consensus of 3830 amino acids. Results of the bootstrap analyses are indicated in bold.



among sites in the *ND5* gene; patterns of selection will be discussed elsewhere[3].

Discussion

The power of mitochondrial genomics

Discussions of the evolutionary implications of phylogenetic hypotheses based on smaller molecular data sets retain a certain degree of statistical ambiguity, owing both to results that vary according to the assumptions of the models employed, and to the often relatively low levels of statistical support for various branch points. The evidentiary power of mtDNA-genome-sized data sets permits such discussions to proceed with confidence that the correct tree has been identified (Otto et al. 1996). It should be remembered that this genome tree may or may not correspond to the species tree.

The topologies of the trees produced by each of the 3 methods are identical at the species level. There are 10 resolvable nodes; 6 of these among species, of which 4 were resolved in >95% of bootstrap replications by all 3 methods of analysis (3 by 100% in all methods), and one by 85%–95% depending on the method used. The remaining branch

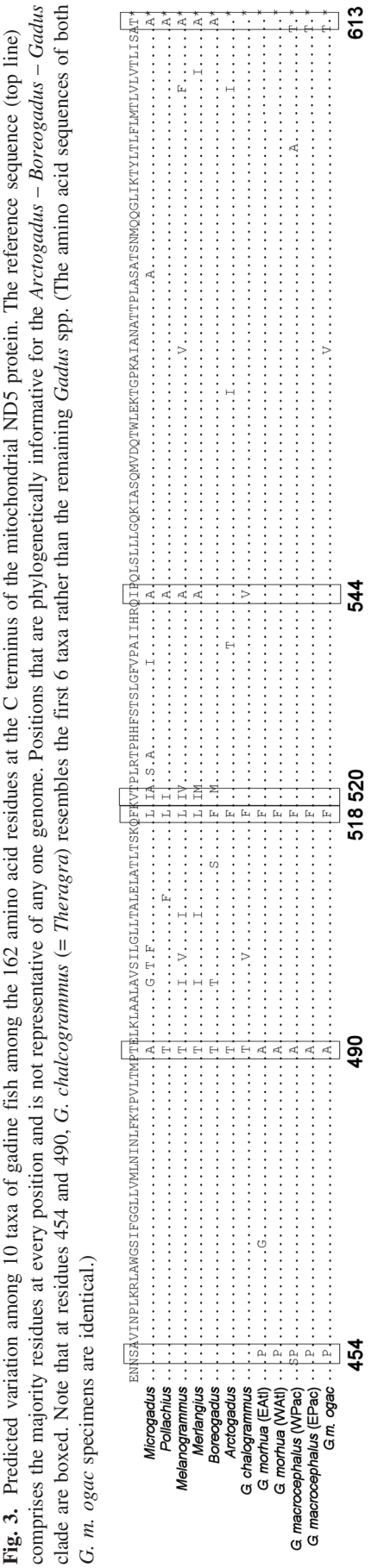


Fig. 3. Predicted variation among 10 taxa of gadine fish among the 162 amino acid residues at the C terminus of the mitochondrial ND5 protein. The reference sequence (top line) comprises the majority residues at every position and is not representative of any one genome. Positions that are phylogenetically informative for the *Arctogadus* – *Boreogadus* – *Gadus* clade are boxed. Note that at residues 454 and 490, *G. chalcogrammus* (= *Theragra*) resembles the first 6 taxa rather than the remaining *Gadus* spp. (The amino acid sequences of both *G. m. ogac* specimens are identical.)

was supported by 99% by one method, and a majority of the other 3. We therefore proceed to interpret the implications of this tree with confidence.

Systematics and biogeography of Gadinae

The subfamily Gadinae (Gadiformes: Gadidae) comprises some of the most abundant and commercially important fishes in the sea, including Atlantic and Pacific cod (*Gadus* spp.), haddock (*Melanogrammus*), whiting (*Merlangius*), and the Alaskan or walleye pollock (*Theragra*) (Cohen et al. 1990). *Gadus* and *Theragra* in particular have been the targets of large, directed fisheries. Evolutionary relationships among gadiform taxa have been analyzed with a variety of developmental and morphological approaches including early life history stages (Dunn and Matarese 1984; Fahay and Markle 1984; Dunn 1989), otoliths (Nolf and Steurbaut 1989), musculature (Howes 1989, 1991), and osteological characters (Cohen 1984; Dunn 1989; Markle 1989).

The greatest diversity of gadines (ca. 22 spp.) is found in the North Atlantic, with 2 genera found in Arctic waters (Polar and Arctic cod: *Boreogadus* and *Arctogadus*, respectively), 1 in Arcto-Pacific waters (Navaga and Saffron cod: *Eleginus* (= *Microgadus*) spp., see below), and 3 species endemic to the north Pacific (*Gadus macrocephalus*, *Theragra chalcogramma*, and the Pacific tomcod, *Microgadus proximus*). This distribution suggests an original area of endemism in the eastern North Atlantic (Svetovidov 1948), which is supported the occurrence of the earliest known fossil ascribed to gadids (*Protocodus*) in Greenland (Fedotov and Bannikov 1989 and references therein), as well as the discovery of fossil otoliths of Atlantic origin (Nolf & Steurbaut 1989). Molecular study of 0.9 kb of the mtDNA genome supported this concept (Carr et al. 1999), and identified 4 clades or clusters of species among 14 nominal species of gadines: (*i*) *Trisopterus* spp. (pouts, poutings, and poor cod), the genus that was the sister group to the remaining taxa; (*ii*) *Microgadus* spp. + *Eleginus navaga* (saffron cod), which was synonymized with the former genus as *Microgadus navaga*; (*iii*) *Melanogrammus* + *Merlangius*; and (*iv*) *Gadus* spp. + *Boreogadus* + *Theragra*. Møller et al. (2002) extended this study with the inclusion of *Arctogadus* (among other species), which they found to fall within group *iv*.

Knowledge of gadine phylogeny has since been increased by studies of 2 nuclear loci (Pogson and Mesa 2004) and a study of mtDNA from an expanded set of gadid species (Teletchea et al. 2006). Pogson and Mesa (2004) suggested that the Pacific cod *G. macrocephalus* was an outgroup to the remaining species in the genus, which they took as evidence for a Pacific origin for *Gadus*. Their study demonstrated strong positive selection at the pantophysin gene among a number of gadids; as such, the relationships implied need to be taken with caution because shared characters may not necessarily reflect common ancestry so much as similar adaptive evolutionary paths. Pogson and Mesa (2004) also suggested that lack of resolution among species was due to a high level of homoplasy in the mitochondrial data. Teletchea et al. (2006) examined the relationships of Gadidae as they relate to morphological evolution of the group, based on 1530 bp of the mitochondrial *COI* and *Cytb* loci for 32 species in 14 genera, 11 of which were in

common with Carr et al. (1999). The topology of the consensus trees and relative degrees of bootstrap support were similar. As in Carr et al. (1999), the relationships between *Gadus* spp. and *Theragra* were not resolved by either parsimony or maximum likelihood: *G. morhua* and *Theragra* were joined with a posterior probability of only 58% in Bayesian analysis. With *Microgadus* as the outgroup to the remaining gadine genera, the placement of *Pollachius* as sister group with respect to *Melanogrammus – Merlangius* and *Gadus* spp. was also not highly supported by the first 2 methods, but had a posterior Bayesian probability of 0.96. Suzuki et al. (2002) caution that reliance on the magnitude of posterior probabilities may lead to "overcredibility" of branching relationships that are not supported by other phylogenetic methods.

Intergeneric relationships among *Microgadus*, *Pollachius*, *Merlangius*, and *Melanogrammus*

Carr et al. (1999) demonstrated at least 2 invasions from the Atlantic to Pacific and Arctic waters by independent gadine lineages. The older invasion is represented by *Microgadus proximus*, a Pacific endemic, as part of *Microgadus* sensu lato, including Atlantic *Microgadus tomcod* and Arctic *Microgadus* (*Eleginus) navaga*, which was more closely related to the Pacific species. (*Microgadus (Eleginus) gracilis* is the sister species of *Microgadus navaga*, as predicted (unpublished results).) Given *Microgadus* as the outgroup, the present analysis places *Pollachius* in turn as the outgroup to the remaining genera, which comprise 2 clades (*Melanogrammus* + *Merlangius*) and (*Boreogadus – Arctogadus – Gadus*).

The monotypic genera *Merlangius merlangus* (whiting) and *Melanogrammus aeglefinus* (haddock) are sister taxa, with an estimated divergence time of 7–8.5 million years ago (Carr et al. 1999; Bakke and Johansen 2005). *Merlangius* occupies the eastern Atlantic, and *Melanogrammus* has a trans-Atlantic distribution. The distribution of *Melanogrammus* in the eastern Atlantic is almost entirely sympatric with *Merlangius*, although the former is absent from the Mediterranean sea. Dunn's (1989) morphological analysis did not identify any synapomorphies uniting these 2 taxa; the species share a dark spot above the pectoral fin in *Melanogrammus*, which is less pronounced in *Merlangius* (Cohen et al. 1990). Whiting and haddock have different ecological niches, with whiting generally feeding higher in the water column (Cohen et al. 1990). Ecological separation of these 2 species' lineages may have occurred during a period of high resource availability that permitted foraging specialization and ecological segregation. Such occurrences, though typical of anadromous or freshwater species (i.e., stickleback and various salmonids) (Ferguson and Mason 1981; McPhail 1992; Taylor and Bentzen 1993; Schluter 1996*a*, 1996*b*; Pigeon et al. 1997), are relatively rare in marine fishes (but see Rocha et al. 2005). Alternatively, the present-day distribution may have arisen after separation of their common ancestor into allopatric eastern *Merlangius* and western *Melanogrammus* lineages, followed by establishment of secondary contact by reinvasion of the latter into the eastern Atlantic.

Relationships and nomenclature for *Theragra chalcogramma*: *Gadus chalcogrammus*?

The genomic data show that *Theragra chalcogramma* and *Gadus morhua* are sister taxa, more closely related to each other than either is to *Gadus macrocephalus* (Fig. 1). Like most "codfish"-like forms, *Theragra chalcogramma* was initially described in the Linnaean genus as *Gadus chalcogrammus* Pallas 1811. It was subsequently transferred to *Pollachius*, and then to a new genus, *Theragra*, with a new species epithet, *T. fucensis* (Jordan & Evermann 1898). Svetovidov (1948) restored the name to *Theragra chalcogramma* (Svetovidov 1948). To keep the genus *Gadus* monophyletic (sensu Farris 1974), we propose restoring the original name, *Gadus chalcogrammus* Pallas 1811. We use this nomenclature throughout the remainder of this paper. Thus, the genus *Gadus* sensu lato includes only 3 nominal species, *G. morhua* in the Atlantic, *G. chalcogrammus* in the Pacific, and *G. macrocephalus* (including *G. m.* subsp. *ogac*) in the Pacific, Arctic, and Atlantic (Fig. 4A).

Walleye pollock and Atlantic cod represent similar eco-evolutionary adaptive experiments as abundant, widespread species in the 2 major oceans, which have supported what were formerly the largest commercial fisheries in the world. The ability to support such intense exploitation may be the result of shared ecological traits. For example, both species are highly fecund, with egg production on the order of 1–5 million eggs for *G. morhua*, and 0.5–15 million eggs for *G. chalcogrammus* (Cohen et al. 1990). Neither species forms localized, morphologically differentiated subspecies, and significant genetic subdivision attributable to stock separation exists primarily across ocean basins (Carr and Crutcher 1998; Arnason 2004). In contrast, the third species in the genus has given rise to a localized subspecies of low abundance (*G. m.* subsp. *ogac*) that has undergone sufficient morphological evolution to be considered previously as a separate species.

Arctic origins of *Gadus*

Boreogadus is the most northerly distributed species of bony fish (Scott and Scott 1988), and along with *Arctogadus* and *Microgadus navaga*, represents the 3 species of gadines endemic to Arctic waters. Parsimony analysis strongly supports *Boreogadus* as the sister species to *Arctogadus* and *Gadus*, and distance and likelihood analyses support this arrangement in a majority of bootstrap replicates, either in the complete or reduced taxon set. Møller et al. (2002) analyzed 0.4 kb of the mtDNA *cytochrome b* locus from *Arctogadus*, which grouped it with *Boreogadus* and *Gadus* spp., but without significant support for any particular branching arrangement. Because pairwise genetic differences between these genera were less than those within other gadine genera, such as pouts (*Trisopterus*), they suggested that *Arctogadus* should be synonymized either with *Boreogadus* or *Gadus*. However, no particular level of genetic differentiation "makes a species" (Hendry et al. 2000). *Trisopterus* is evidently an older lineage, which in consequence shows larger pairwise sequence divergences among its species. Given the phylogenetic relationships identified here, synonymization of *Arctogadus* with *Boreogadus* would create a paraphyletic genus, and synonymization with *Gadus* would violate the principle of taxonomic stability.

Fig. 4. Circumpolar distributions and 2 alternative biogeographic hypotheses of gadine species involved in the trans-Arctic exchange. (A) Distributions of gadine fish species (after Cohen et al. 1990): *Arctogadus* spp. (blue), *Boreogadus* (light blue), *Gadus morhua* (red), *G.* (= *Theragra*) *chalcogrammus* (green), *G. macrocephalus* (light green) including *G. m. ogac* (light red). (B) Origin of *Gadus* spp. from a circumpolar ancestor as an Atlantic clade (reds) followed by a secondary invasion of the northern Pacific by *macrocephalus*, and a tertiary Pacific invasion by *chalcogrammus*. (C) Origin of *Gadus* spp. from a circumpolar ancestor as a Pacific clade (greens) followed by a secondary invasion of the Atlantic by *morhua*. In either hypothesis, *G. m. ogac* is a subsequent extension into the Atlantic of *G. macrocephalus*. The first hypothesis requires an extra trans-Arctic vicariance event from Atlantic to Pacific waters; the latter implies sympatric differentiation of *macrocephalus* from *chalcogrammus* (see text for details). The scale shows the observed number of pairwise substitutions per site; the timescale is calibrated at 1.5% per million years per pair of lineages (cf. Bermingham et al. 1997).

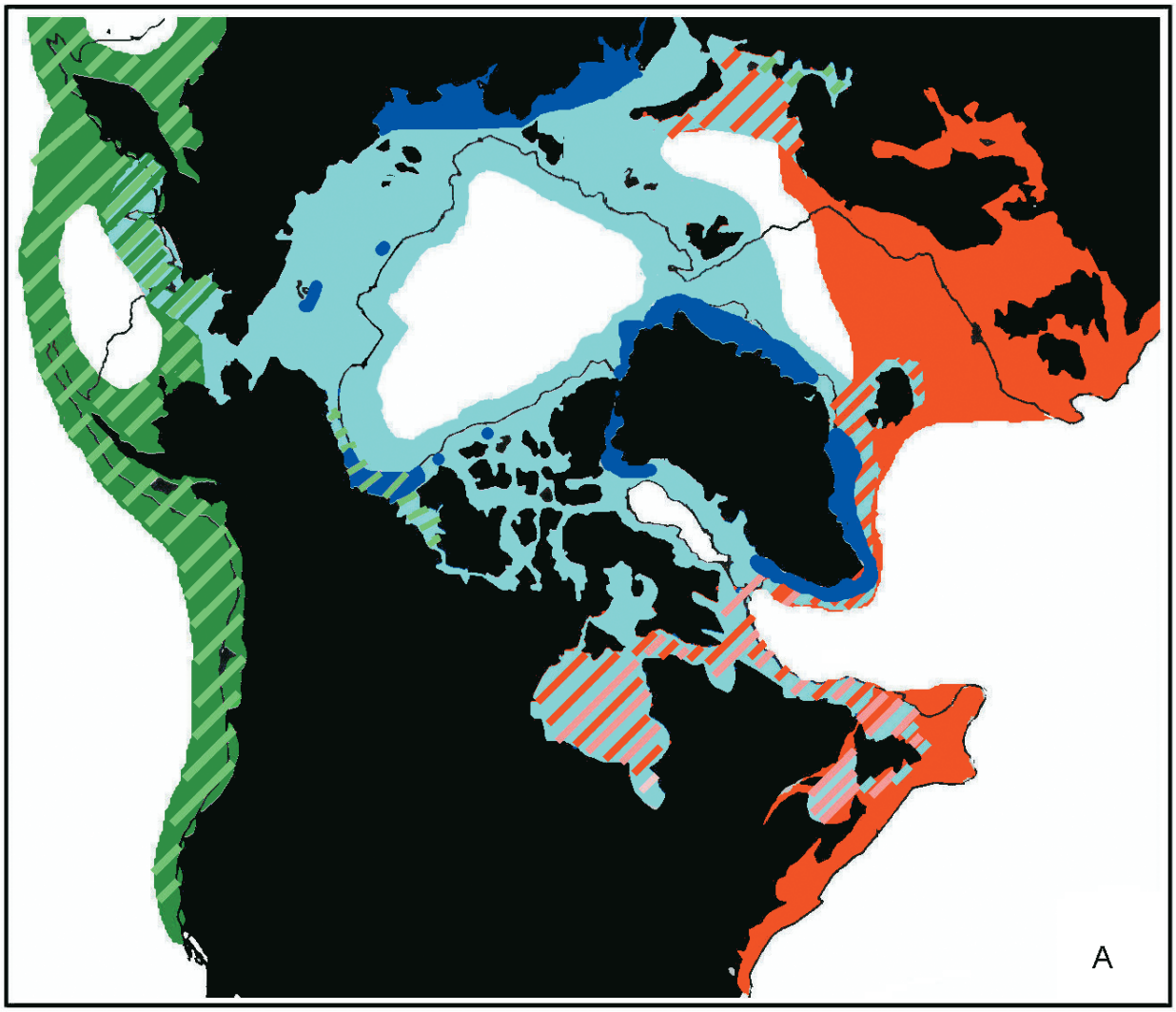
An evolutionary tree with *Boreogadus* as the sister to *Arctogadus* and *Gadus* sensu lato suggests an Arctic origin of *Gadus* (Fig. 4). Since the sister taxa to the *Boreogadus*-inclusive clade (*Pollachius* spp., and *Melanogrammus* + *Merlangius*) are all Atlantic species, the *Boreogadus*-inclusive lineage seems to have entered Arctic waters from the Atlantic. This event would almost certainly have predated significant global cooling (Herman and Hopkins 1980); subsequent separation of the *Arctogadus*-inclusive clade from the *Boreogadus* lineage is not easily explained by geographic or thermal events (cf. Møller et al. 2002). *Boreogadus* is more continuously distributed in Arctic waters than the fragmented, disjunct populations of *Arctogadus*, and the species are sympatric in parts of their range, which Møller et al. (2002) hypothesized might be a result of a secondary invasion of *Boreogadus* into the Arctic Ocean. However, given the limited distribution of *Arctogadus*, it may be that historical ranges for this species were much larger than at present. Although Arctic extinctions might explain present day-distributions of taxa in these water, paleoclimate data suggest that the Arctic was not much warmer 5–6 mya than at present (Herman and Hopkins 1980), and fossil evidence indicates that many gadids have an exclusively Miocene–Pliocene record or are most diverse and abundant during this era, during which they experienced a great deal of diversification and speciation (Nolf and Steurbaut 1989 and references therein).

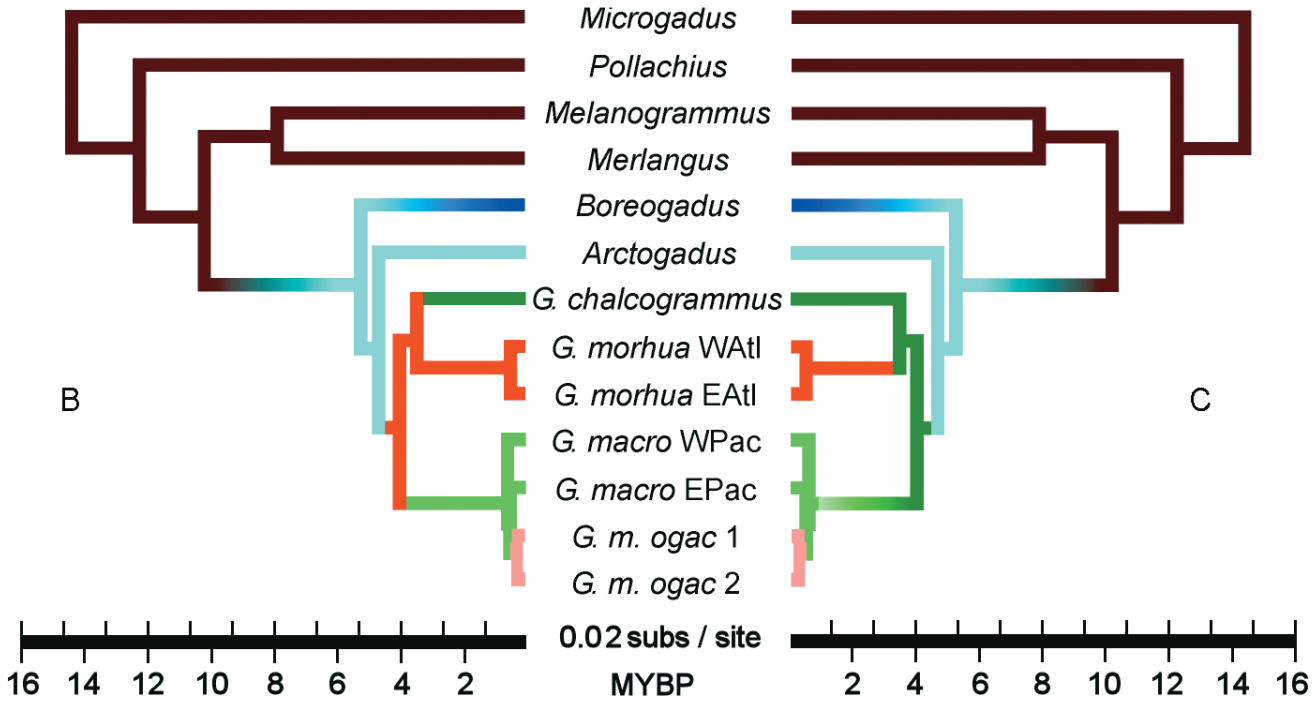
Vicariance biogeography I: Pacific–Atlantic invasions

Besides the Pacific invasion by *Microgadus*, Carr et al. (1999) considered the origins of the other 2 Pacific endemics, *G. macrocephalus* and *G.* (*Theragra*) *chalcogrammus*. Each of these 3 species shows a similar degree of sequence divergence from its closest non-Pacific relative. They therefore proposed that these species represent separate yet simultaneous invasions into the Pacific at about the time of the most recent opening of the Bering Strait 3–4 mya. However, because the exact relationships of *G. macrocephalus* and *G. chalcogrammus* with respect to *G. morhua* were not resolved, it could not be determined definitely whether their current distributions were the result of a single invasion and subsequent divergence, or 2 independent invasions by separate species. Now, the genomic evidence shows that *G. chalcogrammus* is more closely related to *G. morhua* than either is to *G. macrocephalus*. The mean number of pairwise nucleotide differences between *G. morhua* and *G. macrocephalus* (585) is slightly more than that between *G. morhua* and *G. chalcogrammus* (553) (Table 2). After an initial invasion of the Pacific by *Microgadus* (Carr et al. 1999), 2 hypotheses are then available to account for current distributions (Fig. 4A): either an invasion of the eastern Atlantic by *G. morhua* and a secondary invasion of the northern Pacific by *G. macrocephalus* from an Arctic lineage, followed by a tertiary invasion of the Pacific by *G. chalcogrammus* from the *G. morhua* lineage (Fig. 4B), or invasion of the Pacific by the common ancestor of *Gadus* sensu lato from an Arctic lineage, with subsequent re-invasion of the Atlantic by the *G. morhua* lineage (Fig. 4C).

The latter hypothesis requires one fewer dispersal event. Biogeographic movements from the Pacific into the Atlantic are generally considered more common than the reverse (Vermeij 1991; Svitoch and Taldenkova 1994). Changes in the direction of water flow through the Bering Strait are determined by the relative timing of the opening and closing of the strait and the Panamanian Isthmus. Recent studies suggest that the first opening of the Bering Strait, typically dated 3–4 mya, may in fact have occurred up to 6.4 mya (Marincovich and Gladenkov 1999, 2001; Gladenkov et al. 2002). A dominant southward flow may have persisted until sometime more than 3.5 mya (cf. Olsen et al. 2004). A critical threshold in the closure of the Isthmus then effected a reorganization of oceanic circulation in the northern hemisphere, including onset of the present northward flow through the Bering Strait (Marincovich and Gladenkov 1999). In this scenario, initial invasion by *Microgadus* on a southward current would have been followed by a second invasion of the common ancestor of the *Gadus* lineage. Upon closure of the Isthmus of Panama, the reversal of current flow would have allowed movement of *G. morhua* back through the Strait and into the Atlantic via the Arctic. This alternative theory has been advocated by Pogson and Mesa (2004). Their study of Darwinian selection on the nuclear *Pantophysin I* locus among gadids also identified *G. macrocephalus* as the sister group to *G. morhua* and *G. chalcogrammus*, as in the present data, and in contrast to the present study, *Arctogadus* and *Boreogadus* as sister taxa (see their Fig. 1). They therefore suggested that *G. macrocephalus* invaded the Pacific from an Arctic ancestor in the *Boreogadus–Arctogadus* lineage, and that *G. morhua* represents a re-invasion of the Atlantic.

Several lines of evidence argue against this alternative hypothesis. Pogson and Mesa (2004) resolved 2 alleles at the *Pantophysin I* locus in *G. morhua*, one of which is more closely related to the single allele in *G. chalcogrammus* than it was to the alternate conspecific allele, which suggest that intra-*G. morhua* polymorphism antedates the separation of *G. chalcogrammus*. We agree with them that the allelic relationships may reflect selection rather than phylogenetic affinity. Additionally, the genetic distance between *G. morhua* and *G. macrocephalus* is 50% higher than between





G. morhua and *G. chalcogrammus* (see Table 2 of Pogson and Mesa 2004). However, this larger relative divergence (as compared with our mtDNA data) of *G. morhua* and *G. macrocephalus* would not necessarily be surprising, if selection occurs at the *pantophysin* locus. Grant and Ståhl (1988) showed markedly lower heterozygosity at allozyme loci in *G. macrocephalus* relative to *G. morhua*, and proposed that Pacific cod are derived from Arcto-Atlantic ancestors that experienced a bottleneck upon entering the Pacific Ocean and have had insufficient time for a mutation-drift equilibrium to be re-established with levels of heterozygosity consistent with current population size. The same low heterozygosity is seen in *chalcogrammus* (Grant and Utter 1980), for which the same argument can be made once its phylogenetic relationships are understood. Diversity of mtDNA haplotypes within *morhua* is greater in the eastern than the western Atlantic (Carr and Crutcher 1998; Arnason 2004) and the deepest mitogenomic lineages are found there as well (Carr et al., submitted)[4], suggests an origin of *G. morhua* in the eastern Atlantic and expansion westward, rather than through a northwest passage to the western Atlantic and thence eastward. This pattern could also be the result of recolonization of the northwest Atlantic by lineages from the northeast Atlantic (Wares and Cunningham 2001). Finally, although Pogson and Mesa (2004) offer no explicit hypothesis for the Pacific distribution of *G. chalcogrammus*, this model requires separation of the *G. macrocephalus* and *G. chalcogrammus* lineages in sympatry after migration of their common ancestor southward through the strait (Fig. 4C), an intrinsically improbable event given their similar ecological requirements and distributions. We therefore favour the first hypothesis, in which the present-day occurrence of *G. macrocephalus* and *G. chalcogrammus* arise from essentially simultaneous invasions of 2 distinct evolutionary lineages southward through the Bering Strait, prior to the reversal of current patterns. The separation of the *G. macrocephalus* and *G. morhua – G. chalcogrammus* lineages is dated ~4 mya (Fig. 4), and the latter separated slighter later (~3.8 mya, Fig. 4). This slight delay may have allowed initial postzygotic isolation to evolve between *G. macrocephalus* and *G. morhua* before the second invasion, which would have facilitated the divergence of *G. chalcogrammus*.

Vicariance biogeography II: re-invasion of the Atlantic Ocean

Although it typically assumed that Greenland cod from the western Atlantic are more closely related to sympatric *G. morhua* than to allopatric *G. macrocephalus*, the genomic data show that they are most closely related to the latter. Nuclear gene data are in agreement (Pogson and Mesa 2004). Greenland cod appear to represent a re-invasion of the Atlantic, where it has come into secondary contact with *G. morhua*. The range of 45–62 pairwise interspecific nucleotide differences between Greenland cod and Pacific cod is about the same as the 51 differences observed between *G. morhua* from opposite sides of the Atlantic (Table 2).

Re-introduction and (or) continued exchange might be by way of either the continental shelf and northern Canadian archipelago or more directly through the polar basin. A major ocean current runs directly through the Bering Strait into the polar basin and down the eastern coast of Greenland, and is capable of passive transport of inanimate objects from the northern Pacific to the eastern Atlantic (Ebbesmeyer 2003). However, the Arctic Ocean is considerably deeper than the isobaths at which gadine fish are normally found, and this route would require that these demersal fish become pelagic for extended periods. It is more likely that movement from the Pacific to the Atlantic would occur as a more gradual progression through shallower areas of the northern archipelago and (or) continental shelf. Fish attributed to *G. macrocephalus* occur in the Beaufort Sea as far east as Amundsen Gulf, but there appear to be no archipelagic records (Fig. 4A).

Finally, the presence in the Barents Sea off the northern coast of Norway of a gadid referred to as *Theragra finnmarchia* (Christiansen et al. 2005), would also imply another northwestward vicariance through the Bering strait and along the north slope of Asia. Genetic affinities of this form are unknown.

Comparison of genomic versus proteomic phylogenetic signals

In the power analyses, resampling of 4000 bp subsets of the genome resolved all nodes that were identified in the complete genome, and re-sampling of 8000 bp supported almost all of these nodes at the 95% confidence threshold (Table 3). Thus, approximately one half of the mtDNA genome, sampled at random is sufficient to provide a fully resolved, statistically robust tree.

In contrast, sampling of the entire mitochondrial proteome was required to obtain the full resolution of branches. This is due at least in part to heterogeneous phylogenetic signal in different parts of the proteome. Comparison of phylogenetic structure between nucleotide and amino acid data suggests a priori evidence for positive selection on the mitochondrial genome, indicated by the heterogeneous accumulation of amino acid replacements for ND5 towards the C-terminus. A similar pattern of parallel amino acid substitution has been observed among 3 wolffish species (*Anarhichas* spp.) with respect to a more distant outgroup; there, the effect on resolution of a 3-taxon phylogenetic problem is not so pronounced (Johnstone et al., submitted)[5]. Evaluation of the significance of this pattern will be discussed elsewhere, which will have implications for functional genomic studies of nucleo-cytoplasmic interactions for the adaptive evolution of cod.

Acknowledgments

We thank J.S. Christiansen (University of Tromsö,

[4] S.M. Carr, H.D. Marshall, S.M.C. Flynn, K.A. Johnstone, A.M. Pope, S.L. Royston, and C.D. Wilkerson. 2006. The ArkChip – A multi-species, iterative mtDNA sequencing strategy for biodiversity and phylogeographic genomics. Compar. Biochem. Physiol. D Genomics Proteomics. Submitted.

[5] K.A. Johnstone, H.D. Marshall, and S.M. Carr. 2006. Biodiversity genomics for Species At Risk: patterns of DNA sequence variation within and among complete mitochondrial DNA genomes of three species of wolffish (*Anarhichas* spp.). Can. J. Zool. Submitted.

Tromsö, Norway), Julian Dodson (Université Laval, Qué.), Don Flescher and Mike Sissenwine (National Marine Fisheries Service, Woods Hole, Mass.), John Pope (Fisheries Laboratory, Lowestoft, UK), Leo Margolis (Pacific Biological Station, Nanaimo, B.C.), and the groundfish staff from the Northwest Atlantic Fisheries Centre for their assistance and willingness to provide specimens for this study. We thank Dr. Joe Bielowski (Dalhousie University, Halifax, N.S.) for assistance with the PAML analysis, and 3 anonymous reviewers for constructive comments on the original draft of the manuscript. Data in this paper were collected as part of the M.Sc. thesis research of M.W.C.: we thank Dr. Cliff Cunningham (Duke University, Durham, N.C.) for encouragement, constructive comments, and advice. The work was supported by Department of Fisheries and Oceans University Partnership Programs grants to S.M.C., a Memorial University of Newfoundland Fellowship to M.W.C., and by DFO funding to P.P.

References

Arnason, E. 2004. Mitochondrial cytochrome *b* DNA variation in the high-fecundity Atlantic cod: trans-Atlantic clines and shallow gene genealogy. Genetics, **166**: 1871–1885. doi:10.1534/genetics.166.4.1871. PMID:15126405.

Bakke, I., and Johannsen, S.D. 2005. Molecular phylogenetics of gadidae and related gadiformes based on mitochondrial DNA sequences. Mar. Biotechnol. **7**: 61–69. doi:10.1007/s10126-004-3131-0. PMID:15759085.

Bermingham, E., McAfferty, S.S., and Martin, A.P. 1997. Fish biogeography and molecular clocks: perspectives from the Panamanian isthmus. *In* Molecular systematics of fishes. *Edited by* T.D. Kocher and C.A. Stepien. Academic Press, San Diego, Calif. pp. 113–126.

Carr, S.M., and Crutcher, D.C. 1998. Population genetic structure in Atlantic cod (*Gadus morhua*) from the North Atlantic and Barents Sea: contrasting or concordant patterns in mtDNA sequence and microsatellite data? *In* The implications of localized fishery stocks. *Edited by* I. Hunt von Herbing, I. Kornfield, M. Tupper, and J. Wilson. Northeast Regional Agricultural Engineering Service, Ithaca, N.Y. pp. 91–103.

Carr, S.M., Kivlichan, D.S., Pepin, P., and Crutcher, D.C. 1999. Molecular systematics of gadid fishes: implications for the biogeographic origins of Pacific species. Can. J. Zool. **77**: 19–26. doi:10.1139/cjz-77-1-19.

Christiansen, J.S., Fevolden, S.E., and Byrkjedal, I. 2005. The occurrence of *Theragra finnimarchica* Koefoed, 1956 (Teleostei, Gadidae), 1932–2004. J. Fish Biol. **66**: 1193–1197. doi:10.1111/j.0022-1112.2005.00682.x.

Cohen, D.M. 1984. Gadiformes: overview. *In* Ontogeny and systematics of fishes. *Edited by* H.G. Moser, W.J. Richards, D.M. Cohen, M.P. Fahay, A.W. Kendall, Jr., and S.L. Richardson. American Society of Ichthyologists and Herpetologists, and Allen Press, Lawrence, Kans. pp. 259–265.

Cohen, D.M., Inada, T., Iwamoto, T., and Scialabba, N. 1990. FAO species catalogue. Vol. 10. Gadiform fishes of the world. Food and Agriculture Organization of the United Nations, Rome, Italy.

Davis, C.S., Delisle, I., Stirling, I., Siniff, D.B., and Strobeck, C. 2004. A phylogeny of the extant Phocidae inferred from complete mitochondrial DNA coding regions. Mol. Phylogenet. Evol. **33**: 363–377. PMID:15336671.

Dunn, J.R. 1989. A provisional phylogeny of Gadid fishes based on adult and early life-history characters. *In* Papers on the systematics of Gadiform fishes. *Edited by* D.M. Cohen. Scientific Publications Committee, Natural History Museum of Los Angeles County, Los Angeles, Calif. pp. 209–235.

Dunn, J.R., and Matarese, A.C. 1984. Gadidae: development and relationships. *In* Ontogeny and systematics of fishes. *Edited by* H.G. Moser, W.J. Richards, D.M. Cohen, M.P. Fahay, A.W. Kendall, Jr., and S.L. Richardson. American Society of Ichthyologists and Herpetologists, and Allen Press, Lawrence, Kans. pp. 283–299.

Ebbesmeyer, C. 2003. Far-flung bathtub toys due in New England. [online]. Available from http://www.agu.org/sci_soc/ducks.html [cited 11 July 2003].

Fahay, M.P., and Markle, D.F. 1984. Gadiformes: development and relationships. *In* Ontogeny and systematics of fishes. *Edited by* H.G. Moser, W.J. Richards, D.M. Cohen, M.P. Fahay, A.W. Kendall, Jr., and S.L. Richardson. American Society of Ichthyologists and Herpetologists, and Allen Press, Lawrence, Kans. pp. 265–283.

Farris, J.S. 1974. Formal definitions of paraphyly and polyphyly. Syst. Zool. **23**: 548–554. doi:10.2307/2412474.

Fedotov, V.F., and Bannikov, A.F. 1989. On phylogenetic relationships of fossil Gadidae. *In* Papers on the systematics of Gadiform fishes. *Edited by* D.M. Cohen. Scientific Publications Committee, Natural History Museum of Los Angeles County, Los Angeles, Calif. pp. 187–195.

Ferguson, A., and Mason, F.M. 1981. Allozyme evidence for reproductively isolated sympatric populations of brown trout *Salmo trutta* L. in Lough Melvin, Ireland. J. Fish Biol. **18**: 629–642. doi:10.1111/j.1095-8649.1981.tb03805.x.

Gladenkov, A.Y., Oleinik, A.E., Marincovich, L., Jr., and Barinov, K.B. 2002. A refined age for the earliest opening of Bering Strait. Palaeogeogr. Palaeoclimatol. Palaeoecol. **183**: 321–328. doi:10.1016/S0031-0182(02)00249-3.

Goldman, N., and Yang, Z.H. 1994. Codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol. Biol. Evol. **11**: 725–736. PMID:7968486.

Grant, S., and Ståhl, G. 1988. Evolution of Atlantic and Pacific cod: loss of genetic variation and gene expression in Pacific cod. Evolution, **42**: 138–146. doi:10.2307/2409122.

Grant, W.S., and Utter, F.M. 1980. Biochemical genetic variation in walleye pollock, *Theragra chalcogramma*: population structure in the southeastern Bering Sea and the Gulf of Alaska. Can. J. Fish. Aquat. Sci. **37**: 1093–1100.

Hendry, A.P., Vamosi, S.M., Latham, S.J., Heilbuth, J.C., and Day, T. 2000. Questioning species realities. Conserv. Genet. **1**: 67–76. doi:10.1023/A:1010133721121.

Herman, Y., and Hopkins, D.M. 1980. Arctic oceanic climate in late Cenozoic time. Science (Washington, D.C.), **209**: 557–562.

Howes, G.J. 1989. Phylogenetic relationships of Macrouroid and Gadoid fishes based on cranial myology and arthrology. *In* Papers on the systematics of Gadiform fishes. *Edited by* D.M. Cohen. Scientific Publications Committee, Natural History Museum of Los Angeles County, Los Angeles, Calif. pp. 113–128.

Howes, G.J. 1991. Biogeography of gadoid fishes. J. Biogeogr. **18**: 595–622. doi:10.2307/2845542.

Inoue, J.G., Miya, M., Tsukamoto, K., and Nishida, M. 2001. A mitogenomic perspective on the basal teleostean phylogeny: resolving higher-level relationships with longer DNA sequences. Mol. Phylogenet. Evol. **20**: 275–285. PMID:11476635.

Johansen, S., and Bakke, I. 1996. The complete mitochondrial DNA sequence of Atlantic cod (*Gadus morhua*): relevance to taxonomic studies among codfishes. Mol. Mar. Biol. Biotechnol. **5**: 203–214. PMID:8817926.

Marincovich, L., Jr., and Gladenkov, A.Y. 1999. Evidence for an early

opening of the Bering Strait. Nature (London), **397**: 149–151. doi:10.1038/16446.

Marincovich, L., Jr., and Gladenkov, A.Y. 2001. New evidence for the age of Bering Strait. Quaternary Sci. Rev. **20**: 329–335. doi:10.1016/S0277-3791(00)00113-X.

Markle, D.F. 1989. Aspects of character homology and phylogeny of the Gadiformes. *In* Papers on the sytematics of Gadiform fishes. *Edited by* D.M. Cohen. Scientific Publications Committee, Natural History Museum of Los Angeles County, Los Angeles, Calif. pp. 59–88.

McPhail, J.D. 1992. Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): morphological and genetic evidence for a species pair in Paxton Lake, British Columbia. Can. J. Zool. **70**: 361–369.

Minegishi, Y., Aoyama, J., Inoue, J.G., Miya, M., Nishida, M., and Tsukamoto, K. 2005. Molecular phylogeny and evolution of the freshwater eels genus *Anguilla* based on the whole mitochondrial genome sequences. Mol. Phylogenet. Evol. **34**: 134–146. PMID:15579387.

Møller, P., Jordan, A.D., Gravlund, P., and Steffensen, J.F. 2002. Phylogenetic position of the cryopelagic codfish genus *Arctogadus* Drjagin, 1932 based on partial mitochondrial cytochrome *b* sequences. Polar Biol. **25**: 342–349.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York, N.Y.

Nolf, D., and Steurbaut, E. 1989. Evidence from otoliths for establishing relationships between gadiforms and other groups. *In* Papers on the systematics of Gadiform fishes. *Edited by* D.M. Cohen. Scientific Publications Committee, Natural History Museum of Los Angeles County, Los Angeles, Calif. pp. 37–45.

Olsen, J.L., Stam, W.T., Coyer, J.A., Reusch, T.B.H., Billingham, M., Bostrom, C., et al. 2004. North Atlantic phylogeography and large-scale population differentiation of the seagrass *Zostera marina* L. Mol. Ecol. **13**: 1923–1941. doi:10.1111/j.1365-294X.2004.02205.x. PMID:15189214.

Otto, S.P., Cummings, M.P., and Wakeley, J. 1996. Inferring phylogenies from DNA sequence data: the effects of sampling. *In* New uses for new phylogenies. *Edited by* P.H. Harvey, A.J.L. Brown, and J. Maynard Smith. Oxford University Press, Oxford, UK. pp. 103–115.

Pigeon, D., Chouinard, A., and Bernatchez, L. 1997. Multiple modes of speciation involved in parallel evolution of sympatric morphotypes of lake whitefish (*Coregonus clupeaformis*, Salmonidae). Evolution, **51**: 196–205. doi:10.2307/2410973.

Pogson, G.H., and Mesa, K.A. 2004. Positive Darwinian selection at the pantophysin (*Pan* I) locus in marine gadid fishes. Mol. Biol. Evol. **21**: 65–75. PMID:12949133.

Posada, D., and Crandall, K.A. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics, **14**: 817–818. doi:10.1093/bioinformatics/14.9.817. PMID:9918953.

Ramirez, J.L., and Dawid, I.B. 1978. Mapping of mitochondrial DNA in *Xenopus laevis* and *X. borealis* and the postions of ribosomal genes and D-loops. J. Mol. Biol. **119**: 133–146. doi:10.1016/0022-2836(78)90273-5. PMID:633366.

Rocha, L.A., Robertson, D.R., Roman, J., and Bowen, B.W. 2005. Ecological speciation in tropical reef fishes. Proc. R. Acad. Sci. Ser. B, **272**: 573–579.

Ronquist, F., and Huelsenbeck, J.P. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics, **19**: 1572–1574. doi:10.1093/bioinformatics/btg180. PMID:12912839.

Saccone, C., De Giorgi, C., Gissi, C., Pesole, G., and Reyes, A. 1999. Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system. Gene, **238**: 195–209. doi:10.1016/S0378-1119(99)00270-X. PMID:10570997.

Schluter, D. 1996*a*. Ecological speciation in postglacial fishes. Proc. R. Soc. Lond. B. Biol. Sci. **351**: 807–814.

Schluter, D. 1996*b*. Ecological causes of adaptive radiation. Am. Nat. **148**: Suppl. 1, S40–S64. doi:10.1086/285901.

Scott, W.B., and Scott, M.G. 1988. Atlantic fishes of Canada. University of Toronto Press, Toronto, Ont.

Stepien, C.A., and Kocher, T.D. (*Editors*). 1997. Molecules and morphology in studies of fish evolution. *In* Molecular systematics of fishes. Academic Press, San Diego, Calif. pp. 1–11.

Suzuki, Y., Glazko, G.V., and Nei, M. 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. Proc. Natl. Acad. Sci. U.S.A. **99**: 16138–16143. doi:10.1073/pnas.212646199. PMID:12451182.

Svetovidov, A.N. 1948. Gadiformes. Zoological Institute of the Academy of Sciences of the USSR.

Svitoch, A.A., and Taldenkova, E.E. 1994. Recent history of the Bering Strait. Oceanology (Mosc.), **34**: 400–404.

Swofford, D. 1998. PAUP: phylogenetic analysis using parsimony. Version 4.0d63. Smithsonian Institution, Washington, D.C.

Tamura, K., and Nei, M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. **10**: 512–526. PMID:8336541.

Taylor, E.B., and Bentzen, P. 1993. Evidence for multiple origins and sympatric divergence of trophic ecotypes of smelt (*Osmerus*) in northeastern North America. Evolution, **47**: 813–832. doi:10.2307/2410186.

Teletchea, F., Laudet, V., and Hanni, C. 2006. Phylogeny of the Gadidae (sensu Svetovidov, 1948) based on their morphology and two mitochondrial genes. Mol. Phylogenet. Evol. **38**: 189–199. PMID:16311046.

Vermeij, G.J. 1991. Anatomy of an invasion: the trans-Arctic interchange. Paleobiology, **17**: 281–307.

Wares, J.P., and Cunningham, C.W. 2001. Phylogeography and historical ecology of the North Atlantic intertidal. Evolution, **55**: 2455–2469. doi:10.1554/0014-3820(2001)055[2455:PAHEOT]2.0.CO;2. PMID:11831661.

Yang, Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. **13**: 555–556. PMID:9367129.