Journal of Fish Biology (1991) **39** (Supplement A), 283–290

Phylogenetic relationships among the subfamily Coregoninae as revealed by mitochondrial DNA restriction analysis

L. BERNATCHEZ*, F. COLOMBANI AND J. J. DODSON†
Département de Biologie, Université Laval, Ste-Foy, Canada G1K 7P4

Mitochondrial DNA (mtDNA) restriction analysis was used to assess phylogenetic patterns among 21 taxa of the subfamily Coregoninae. The genus *Prosopium* formed a very distinct group differing by 10% (sequence divergence estimate) from other species. *Coregonus* and *Stenodus* species were closely related, diverging by sequence divergence estimates of less than 5·6%. These species split into two major sister groups. One comprised all 'true whitefish' (subgenus *Coregonus*) and four cisco species (subgenus *Leucichthys*). The most distant species within this assemblage was the Acadian whitefish (*C. huntsmani*). The other group included all other cisco species and also the Inconnu (*Stenodus leucichthys*). These results supported a polyphyletic origin of the ciscoes, and did not support *Stenodus* as a sister taxon of the genus *Coregonus*. The levels of sequence divergence observed suggested that most extant coregonines radiated during the Pleistocene.

Key words: mitochondrial DNA; phylogeny; *Coregonus*; *Stenodus*; *Prosopium*.

I. INTRODUCTION

Coregonine fishes are members of the Salmonid family. They are represented by 28 extant species divided into three genera: *Prosopium*, *Stenodus* and *Coregonus* (Reshetnikov, 1988). At present, the genus *Coregonus* is subdivided into two subgenera; *Coregonus*, the sub-terminal mouthed and usually low gill-rakered, benthic feeding species (true whitefish), and *Leucichthys*, the terminal/supra-terminal mouthed ciscoes with high-gill raker numbers and pelagic feeding habits. Phylogenetic reconstruction of coregonines has been hampered both by the scarcity of fossil records (Behnke, 1972) and their exceptional phenotypic plasticity (Mayr, 1963). The use of morphological criteria to detect their systematic relationships has led to conflicting phylogenetic hypotheses (e.g. Norden, 1961; Behnke, 1970; Reshetnikov, 1988). Genetic studies have been restricted to comparisons of few species (e.g. Ferguson *et al.*, 1978; Vuorinen, 1988; Bernatchez & Dodson, 1991). More recently, Bodaly *et al.* (1991) carried out an extensive survey of allozymic variation among *Coregonus* and *Stenodus*.

In order to increase the probability that gene trees reflect the topology of the species tree which is the true evolutionary pathway of a taxon's gene pool, phylogenetic hypotheses deduced from many unlinked and independently transmitted loci must be compared. To this end, we undertook a phylogenetic analysis based on mitochondrial DNA (mtDNA) variation within the coregonine subfamily.

II. MATERIALS AND METHODS

A total of 927 fish representing 21 taxa were sampled between 1987 and 1991 (Table I). Mitochondrial DNA was purified and analysed according to Bernatchez *et al.* (1988).

*Present address: Department of Zoology, University of Guelph, Guelph, Ontario, Canada G1K 7P4.
†Author to whom correspondence should be addressed.

0022–1112/91/39A283+08 $03.00/0

TABLE I. List of taxa sampled with their sample size (*n*) and origin

Species	*n*	Origin
Coregonus clupeaformis (Mitchill), lake whitefish	457	Various (Bernatchez & Dodson, 1991)
Coregonus clupeaformis, Beringian race	68	Various (Bernatchez & Dodson, 1991)
Coregonus lavaretus (L.) European whitefish	205	Various (Bernatchez & Dodson, unpubl. obs.)
Coregonus lavaretus pidschian (Gmelin), pidschian	4	Lapin River, Sos'va River (Siberia)
Coregonus nasus (Pallas), broad whitefish	17	Yukon River (Alaska), Mackenzie Delta (N.W.T.), Sos'va River (Siberia)
Coregonus huntsmani (Scott), Acadian whitefish	3	Hebb L. (Nova Scotia)
Coregonus albula (L.), vendace	4	Bothnian Gulf coast, Pyhäjärvi Lake (Finland);Niewlino Lake (Poland)
Coregonus sardinella (Valenciennes), least cisco	1	Yukon River (Alaska)
Coregonus peled (Gmelin), peled	1	Finland*
Coregonus tugun (Pallas.), tugun	4	Lapin River, Sos'va River (Siberia)
Coregonus autumnalis (Pallas), Arctic cisco	2	Mackenzie Delta (N.W.T.)
Coregonus autumnalis pollan, Thompson, Irish pollan	4	Lough Neagh (Ireland)
Coregonus artedii Leseur, lake cisco	141	Various (Bernatchez & Dodson, 1990)
Coregonus laurettae (Bean), Bering cisco	1	Yukon River (Alaska)
Coregonus hoyii (Gill), bloater	2	Lake Michigan (Michigan)
Stenodus leucichthys (Güldenstadt), inconnu	1	Yukon River (Alaska)
Prosopium cylindraceum (Pallas), round whitefish	3	Hudson Bay (Québec); Musquacook Lake (Maine); Dezadeash Lake (Yukon)
Prosopium williamsoni (Girard), mountain whitefish	2	Alberta
Prosopium spilonotus (Synder), Bonneville whitefish	3	Bear Lake (Utah)
Prosopium gemmiferum (Snyder), Bonneville cisco	1	Bear Lake (Utah)
Prosopium abyssicola (Snyder), Bear Lake whitefish	1	Bear Lake (Utah)

*Peled is endemic to the Soviet Union and was introduced in Finland in 1965.

Aliquots of mtDNA were digested separately with eight hexameric (*Bam* HI, *Bgl* I, *Dra* I, *Hin*d III, *Pvu* II, *Pst* I, *Sma* I, *Xmn* I), four multi-hexameric (*Ava* I, *Ban* I, *Hae* II, *Hin*c II), and one multi-pentameric (*Ava* II) restriction enzymes. Because of the large number of taxa analysed and the large number of fragments observed in most restriction profiles, it was not feasible to map all sites. Therefore, we limited most analyses to the fragments themselves. Thus, each fish was assigned an observed mtDNA composite genotype based on the restriction profiles across all restriction enzymes. A presence/absence matrix of mtDNA fragments for each composite genotype constitute the raw data for further analysis. Sequence divergence between mtDNA genotypes was estimated according to Upholt (1977). The resulting distance matrix was used to construct a phenogram using both a constant

evolutionary rate and an unconstrained branch-length clustering method (Fitch & Margoliash, 1967; programs KITSCH and FITCH in the PHYLIP package, version 3.3, provided by J. Felsenstein). A phylogenetic network was generated from the presence/absence matrix according to Wagner parsimony criteria by using MIX (from PHYLIP). All applications were run 10 times with a different taxa ordering each time. Parsimony methods require that changes of characters appear independently. This requirement is violated with restriction fragment data because the gain or loss of a fragment affects the presence of other fragments. However, the overall redundancy or loss of information is expected to be negligible for large data matrices (Ovenden *et al.*, 1987; Zink & Avise, 1990). Nevertheless, site differences among mtDNA genotypes were estimated from changes in fragment patterns where these could be accounted for by specific site gains or losses using the approach described in Bernatchez & Dodson (1991). This was done to estimate the potential bias inherent in using fragments as character states. Ten restriction enzymes and all species excluding the genus *Prosopium* and *Coregonus tugun* were analysed in this way.

III. RESULTS

LEVELS OF MTDNA SEQUENCE DIVERGENCE VARIATION

All 13 restriction enzymes used produced variable patterns and resolved 80 fragments per mtDNA genotype on average. A total of 139 haplotypes were identified. All species had a diagnostic genetic profile except *P. abyssicola* and *P. spilonotus* which shared an identical haplotype. The haplotype of *C. hoyii* fell within the intraspecfic level of variation of *C. artedi* as did *C. lavaretus pidschian* haplotypes within that of *C. lavaretus*. Averaged interspecfic divergence was highly variable and ranged from no detectable variation to 11·89% (Table II). Genus *Prosopium* was highly divergent from either *Stenodus* (9·73 ± 1·03%, range; 7·94–11·13) or *Coregonus* species (9·69 ± 1·25%, range; 7·73–11·98). Divergence between *Stenodus* and *Coregonus* species was much lower than (3·16 ± 0·6%, range; 2·30–4·40) and not significantly different from the congeneric variation within *Coregonus* (3·01 ± 1·45%, range; 0·35–5·62). Levels of intraspecific variation within the three most extensively studied species were much lower than interspecific variation; *C. artedii* (0·52 ± 0·22%), *C. clupeaformis*, including Beringian race (0·77 ± 0·44%), *C. lavaretus* (0·66 ± 0·28%).

TOPOLOGY OF PHYLOGENETIC TREES

The topologies of KITSCH and FITCH phenograms were almost identical. The concensus tree resulting from the Wagner parsimony analysis also had a concordant topology. One topological difference in species branching was observed between parsimony trees generated from site or fragment character-states. To facilitate comparisons with other studies, we present the KITSCH phenogram (Fig. 1). Cases of non-concordance between different analyses are identified in the following paragraphs.

Prosopium species were highly distinct from other species. Within this genus, *P. cylindraceum* was the most divergent, and little variation was observed among the three sympatric species of Bear Lake. All analyses demonstrated the polyphyletic origin of the subgenus *Leucichthys* (cisco species). Thus, *C. artedii*, *C. hoyii*, *C. laurettae*, *C. autumnalis* and *C. autumnalis pollan* composed a distinct clade. The other ciscoes (*C. peled*, *C. sardinella*, *C. albula* and *C. tugun*) clustered closer to members of the subgenus *Coregonus* than to the above cisco group. In FITCH and MIX trees, *C. peled* branched closer to other *Coregonus* species than to *C.*

TABLE II. Matrix of average % sequence divergence (p), (below main diagonal) with standard deviation (above main diagonal) among mtDNA genotypes observed for 21 coregonine taxa

	Cocl	*Cola*	*Cobe*	*Copi*	*Cona*	*Coal*	*Cosa*	*Cope*	*Cotu*	*Coau*	*Copo*	*Coar*	*Colr*	*Coho*	*Cohu*	*Stle*	*Prcy*	*Prwi*	*Prsp*	*Prgy*	*Prab*
Cocl		0·25	0·24	0·24	0·39	0·36	0·36	0·37	0·44	0·50	0·51	0·56	0·51	0·54	0·51	0·49	0·79	0·72	0·71	0·72	0·71
Cola	1·19		0·25	0·09	0·27	0·33	0·34	0·22	0·43	0·49	0·51	0·55	0·49	0·52	0·49	0·47	0·79	0·73	0·73	0·74	0·73
Cobe	0·85	1·11		0·25	0·37	0·39	0·39	0·34	0·46	0·53	0·54	0·56	0·52	0·55	0·52	0·50	0·79	0·72	0·72	0·72	0·72
Copi	0·71	0·42	0·72		0·28	0·33	0·22	0·42	0·42	0·49	0·50	0·54	0·49	0·52	0·49	0·46	0·78	0·72	0·72	0·73	0·72
Cona	1·98	1·41	1·15	1·15		0·36	0·36	0·35	0·44	0·52	0·54	0·57	0·52	0·55	0·52	0·49	0·80	0·73	0·73	0·74	0·74
Coal	1·81	2·00	1·89	1·89	2·42		0·23	0·28	0·44	0·42	0·43	0·50	0·43	0·48	0·47	0·41	0·83	0·70	0·74	0·74	0·74
Cosa	1·97	2·12	1·96	2·03	2·62	0·35		0·30	0·45	0·43	0·44	0·51	0·43	0·49	0·47	0·41	0·81	0·70	0·74	0·74	0·74
Cope	1·70	1·19	1·50	1·24	2·15	1·00	1·06		0·45	0·49	0·49	0·53	0·47	0·51	0·52	0·46	0·85	0·74	0·76	0·76	0·76
Cotu	2·34	2·28	2·52	1·96	2·07	2·25	2·40	2·43		0·50	0·52	0·56	0·52	0·55	0·57	0·52	0·84	0·73	0·76	0·77	0·76
Coau	3·84	4·28	4·15	3·92	4·83	2·97	3·42	3·76	3·00		0·18	0·33	0·28	0·32	0·55	0·44	0·82	0·76	0·79	0·79	0·79
Copo	4·14	4·73	4·28	4·25	5·30	2·98	3·33	3·51	3·44	0·31		0·35	0·31	0·35	0·54	0·45	0·82	0·77	0·79	0·79	0·79
Coar	4·77	4·98	4·93	4·70	5·57	3·50	3·99	4·25	4·09	1·44	1·29		0·37	0·15	0·60	0·50	0·84	0·79	0·83	0·83	0·83
Colr	3·92	4·20	4·92	3·85	4·77	2·54	2·60	3·02	3·54	1·06	1·12	1·64		0·32	0·55	0·45	0·80	0·73	0·76	0·77	0·76
Coho	4·87	5·00	5·05	4·70	5·62	3·57	4·07	4·31	4·10	1·24	1·32	0·40	1·51		0·59	0·49	0·82	0·77	0·80	0·80	0·80
Cohu	3·77	3·58	3·83	3·33	3·89	2·98	2·97	3·63	4·54	4·88	4·44	5·19	4·16	5·27		0·55	0·84	0·81	0·83	0·83	0·83
Stle	3·54	3·61	3·51	3·25	3·99	2·30	2·42	2·96	3·65	2·71	2·68	3·42	2·66	3·33	4·40		0·81	0·69	0·76	0·76	0·76
Prcy	10·48	10·29	10·15	9·83	9·84	11·19	10·43	11·98	11·46	10·79	0·78	11·37	10·55	11·24	10·43	11·13		0·62	0·61	0·63	0·63
Prwi	8·92	8·67	8·51	8·63	8·48	7·80	7·73	9·20	8·83	9·91	10·33	11·02	9·28	10·36	11·08	7·94	4·88		0·46	0·47	0·46
Prsp	8·24	8·15	8·12	7·97	7·93	8·56	8·46	9·15	8·83	10·39	10·38	11·47	9·64	10·86	11·15	9·99	4·43	2·55		0·28	0·00
Prgy	8·48	8·37	8·22	8·20	8·19	8·51	8·42	9·11	9·11	10·34	10·34	11·43	9·60	10·81	11·12	9·63	4·79	2·81	0·33		0·00
Prab	8·24	8·15	8·12	7·97	7·93	8·56	8·46	9·15	8·83	10·39	10·38	11·47	9·64	10·86	11·15	9·99	4·43	2·55	0·00	0·33	

Names are abbreviated according to the first two letters of generic and specific names, except for Beringian race whitefish (*Cobe*) and *C. laurettae* (*Colr*).

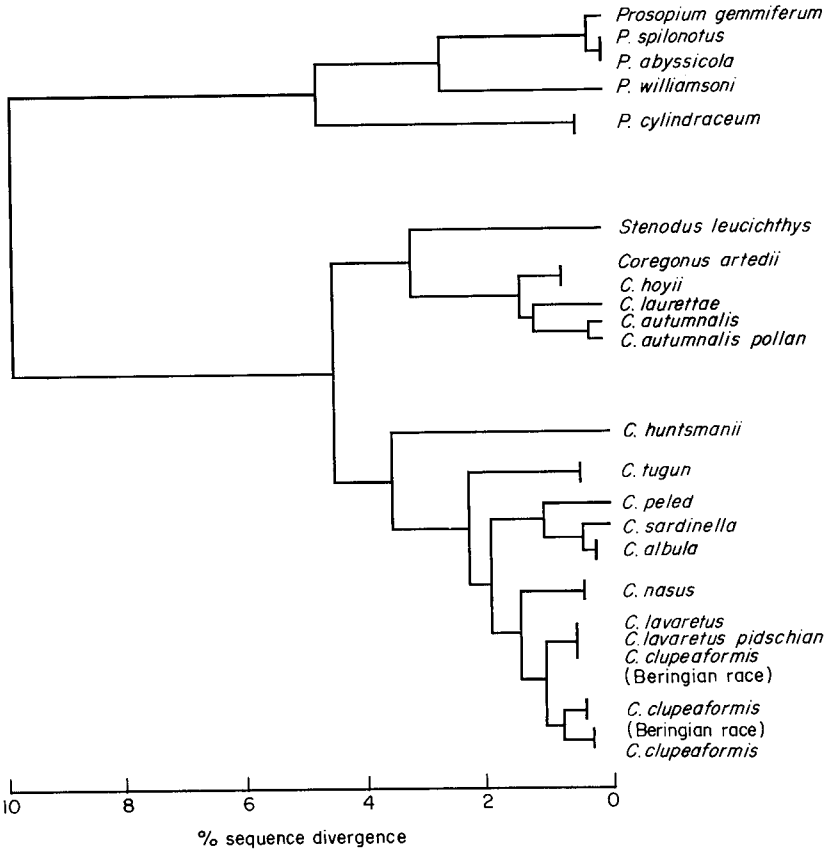

FIG. 1. KITSCH phenogram based on matrix of sequence divergence estimates (p) derived from mtDNA restriction fragment patterns observed among 21 coregonine taxa. For simplicity of presentation, details of intraspecific branchings (presented in Bernatchez & Dodson, 1990, 1991; unpubl. obs.) are omitted. Vertical bars at the tips of branches indicate the level of intraspecific variation.

sardinella and *C. albula*. Within this group, *C. huntsmani* was the most divergent species and represented a sister-taxon of a major clade composed of all other ' true whitefish ' and the four ciscoes.

Stenodus was closely related to *Coregonus* species and no indication was found that it represents a sister-taxon of this genus. In all analyses, it formed with the cisco clade *C. artedii/hoyii/autumnalis/autumnalis pollan/laurettae* a sister group relationship to the other *Coregonus* species. This association was supported in 90% of bootstrap repeats when applying BOOT (from PHYLIP package) to the tree derived from site character-states. When using a *Prosopium* species as an outgroup in MIX and FITCH, rooting was not observed on the *Stenodus* branch. In MIX, the root was located between *C. huntsmani* and all other species whereas FITCH was identical to KITSCH (Fig. 1).

Differentiation among the *C. clupeaformis* and *C. lavaretus* complex is detailed elsewhere (Bernatchez & Dodson, 1991; unpubl. obs.). Briefly, this group was distinct but closely related to *C. nasus*. *C. lavaretus* and *C. clupeaformis* outside Beringia were closely related but fixed for diagnostic mtDNA phylogenetic groupings. Beringian race whitefish were composed of a distinct endemic mtDNA group

intermixed with a mtDNA group of *C. lavaretus* origin. *Coregonus lavaretus pidschian* from Siberia had haplotypes identical or very closely related to those observed in *C. lavaretus* from Europe.

IV. DISCUSSION

LEVELS OF MTDNA DIFFERENTIATION AMONG COREGONINES

The estimate of 10% sequence divergence observed between the genus *Prosopium* and other species demonstrates their ancient separation. Yet, this value must be considered as a minimal divergence estimate. Thomas & Beckenbach (1989) demonstrated from direct sequencing that undetectable homoplasy in either restriction site or fragment analyses of salmonids lead to an increasing underestimation of divergence above the 5% level. For instance, sequence divergence between cutthroat (*Oncorhynchus clarki* Walbaum) and rainbow (*O. mykiss* Walbaum) trouts estimated from site comparisons and corrected from sequencing data were nearly identical (4·74 *v.* 4·9%). The same estimates between rainbow trout and pink salmon were 8·76 and 16·3% respectively. If such underestimation at high genetic distances applies to coregonine data, divergence between *Prosopium* and other coregonines may be as high as 20%. The application of the 2% sequence divergence per million year molecular clock (Brown *et al.*, 1979) largely used for fish mtDNA, would therefore suggest that *Prosopium* diverged in the order of 10 million years ago, during the end of Miocene–early Pliocene period.

A second major conclusion is that radiation of extant coregonine species has occurred recently in evolutionary times. Average congeneric divergence estimates varying from 2·71 to 3·16% suggest that most species radiated during the Pleistocene. Low levels of genetic variation were also observed in allozyme studies and contrast with the level of morphological differentiation observed among coregonines.

Another major finding is the demonstration of the genetic distinctiveness of the endangered Acadian whitefish (*C. huntsmani*). Morphological examination led to the suggestion that *C. huntsmani* represents either an ancestral form or a postglacial specialized phenotype of either a whitefish or cisco species (Behnke, 1972). The level of *C. huntsmani* mtDNA divergence from other species clearly demonstrates that it represents an ancient branch, at least as divergent as the genus *Stenodus* is.

COMPARISON OF PHYLOGENETIC ANALYSES

There was a high degree of congruence between different methods of analysis which suggests that the observed branching patterns are not spurious but reflect phylogenetic relationships. There was also a high degree of concordance between mtDNA trees and the allozyme tree presented by Bodaly *et al.* (1991). Eleven of 12 comparable nodes were identical between phenograms of both studies. The most significant finding of these analyses is the polyphyletic origin of the cisco species which casts doubt on the taxonomic meaning of the subgenus *Leucichthys*. These results corroborate those of Ferguson *et al.* (1978). While previous morphological studies disagreed as to whether the ciscoes were differentiated enough to warrant a subgenus designation, all studies assumed they formed a monophyletic assemblage. Conversely, genetic data demonstrated that the ' cisco phenotype ' is the result of

convergent evolution (or retention of primitive morphology) in different genetic lineages. The 'cisco-phenotype' in the distantly related *Prosopium gemmiferum* reinforces this view.

The major point of discordance between mtDNA and allozyme data concerns the phylogenetic status of *Stenodus*. The present results suggested closer phylogenetic affinity between *Stenodus* and some cisco species than between this group and other coregonines whereas allozyme data suggested that *Stenodus* is a sister taxon of all *Coregonus*, although only slightly differentiated (Bodaly *et al.*, 1991). However, as no outgroup taxon was available in the allozyme study phylogenetic patterns cannot strictly be inferred. More rigorous comparisons of both data sets must await the use of an outgroup taxon in the analysis of allozyme data. Our findings are also in opposition to the classical taxonomic distinction of *Stenodus* based on its highly specialized predator morphology. However, previous authors (e.g. Norden, 1961) recognized that morphological differences observed between *Stenodus* and *Coregonus* do not involve the most phylogenetically significant characters. A careful re-examination of morphological patterns based on cladistic analysis may lead to a reinterpretation of the classical taxonomy (G. Smith, unpubl. obs.). Clearly, much more work needs to be carried out to solve this contradiction.

In conclusion, the high congruence generally observed between mtDNA and allozymic patterns, coupled with the lack of congruence among morphological analyses and between such studies and genetic data suggest that genetic analyses are more likely to reveal phylogenetic patterns in the subfamily Coregoninae. Morphological patterns appear to be better indicators of adaptive processes than phylogenetic tracers.

The authors are indebted to the numerous collaborators who kindly provided samples. To the names already mentioned in earlier publications, we add M. Luczynski, C. Rhulé, A. Kirchkofer, D. Gerdeaux, A. Ferguson, B. Nielsen, T. Edge and the late C. Mills. We also gratefully acknowledge Suzie Boivin for laboratory assistance and Joe Felsenstein for providing the PHYLIP package. This research was funded by an NSERC strategic grant to J.J.D., H. Guderley and D. Pallota and by a grant from the Department of Fisheries and Oceans of Canada to J.J.D. L.B. was supported by NSERC and FCAR postgraduate scholarships.

References

Behnke, R. J. (1970). The application of cytogenetic and biochemical systematics to phylogenetic problems in the family Salmonidae. *Transactions of the American Fisheries Society* **99,** 237–248.

Behnke, R. J. (1972). The systematics of salmonid fishes of recently glaciated lakes. *Journal of the Fisheries Research Board of Canada* **29,** 639–671.

Bernatchez, L. & Dodson, J. J. (1990). Mitochondrial DNA variation among anadromous populations of cisco (*Coregonus artedii*) as revealed by restriction analysis. *Canadian Journal of Fisheries and Aquatic Sciences* **47,** 533–543.

Bernatchez, L. & Dodson, J. J. (1991). Phylogeographic structure in mitochondrial DNA of the lake whitefish (*Coregonus clupeaformis*) in North America and its relationships to Pleistocene glaciations. *Evolution* **45,** 1016–1035.

Bernatchez, L., Savard, L., Dodson, J. J. & Pallotta, D. (1988). Mitochondrial DNA sequence heterogeneity among James–Hudson Bay anadromous coregonines. *Finnish Fisheries Research* **9,** 17–26.

Bernatchez, L., Edge, T. A., Dodson, J. J. & Qadri, S. U. (1991). Mitochondrial DNA and isozyme electrophoretic analyses of the endangered Acadian whitefish, *Coregonus huntsmani*, Scott 1987. *Canadian Journal of Zoology* **69,** 311–316.

Bodaly, R. A., Vuorinen, J., Ward, R. D., Luczynski, M. & Reist, J. D. (1991). Genetic comparisons of New and Old World coregonid fishes. *Journal of Fish Biology* **38,** 37–51.

Brown, W. M., George, M. Jr & Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences. U.S.A.* **76,** 1967–1971.

Ferguson, A., Himberg, K.-J. M. & Svärdson, G. (1978). Systematics of the Irish pollan (*Coregonus pollan* Thompson): an electrophoretic comparison with other Holarctic Coregoninae. *Journal of Fish Biology* **12,** 221–233.

Fitch, W. M. & Margoliash, E. (1967). Construction of phylogenetic trees. *Science* **155,** 279–284.

Mayr, E. (1963). *Animal Species and Evolution.* London: Oxford University Press.

Norden, C. R. (1961). Comparative osteology of representative salmonid fishes, with particular reference to the grayling (*Thymallus arcticus*) and its phylogeny. *Journal of the Fisheries Research Board of Canada* **18,** 679–791.

Ovenden, J. R., Mackinlay, A. G. & Crozier, R. H. (1987). Systematics and mitochondrial genome evolution of the Australian rosellas (Aves: Platycerdidae). *Molecular Biology and Evolution* **4,** 526–543.

Reshetnikov, Ju. S. (1988). Coregonid fishes in recent conditions. *Finnish Fisheries Research* **9,** 11–16.

Thomas, W. K. & Beckenbach, A. T. (1989). Variation in salmonid mitochondrial DNA: evolutionary constraints and mechanisms of substitution. *Journal of Molecular Evolution* **29,** 233–345.

Upholt, W. B. (1977). Estimation of DNA sequence divergence from comparisons of restriction endonucleases digests. *Nucleic Acids Research* **4,** 1257–1265.

Vuorinen, J. (1988). Enzymes genes as hybridization probes in Coregoninae fishes. *Finnish Fisheries Research* **9,** 31–37.

Zink, R. M. & Avise, J. C. (1990). Patterns of mitochondrial DNA and allozyme evolution in the avian genus *Ammodramus. Systematic Zoology* **39,** 148–161.