What, if anything, is *Galaxias vulgaris*?

20 years of research on galaxiid speciation

Graham Wallis, Lise Wallis, and Jon Waters
Department of Zoology, University of Otago, PO Box 56, Dunedin

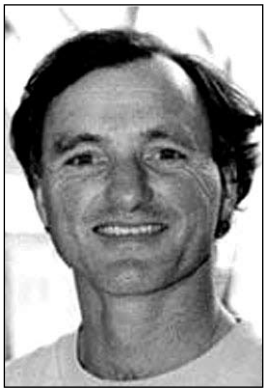
Fish of the family Galaxiidae are renowned for two particular features: their Gondwanan distribution and their movement between rivers and the sea (diadromy) (McDowall 1990). Adults live, breed and lay their eggs in streams, but larvae are flushed out to sea on hatching. Juveniles (whitebait) return to streams 5–6 months later, where they spend the rest of their lives. They are an important component of New Zealand's biodiversity and key players in freshwater ecosystems.

Some 25 species of galaxiids live and breed in New Zealand streams (McDowall 2000). Five of these species maintain the marine juvenile phase, but the others have lost it, becoming non-diadromous, completing their entire life cycle in fresh water. Many of these stream-resident species derive from a koaro (*Galaxias brevipinnis*)-like ancestor (Burridge *et al*., submitted.) This fact is in keeping with the propensity of koaro to climb waterfalls and penetrate deep into river systems, including glacial and volcanic lakes and alpine tarns. This trait may have promoted repeated propagation of stream-resident forms: *G. paucispondylus, G. prognathus, G. divergens, G. cobitinis, G. macronasus* and the *G. vulgaris* group.

Population genetic differentiation in diadromous vs non-diadromous species

Species of fish that go to sea as juveniles have the opportunity to maintain gene flow over a wide area. In support of this prediction, there is no evidence for population genetic structuring within diadromous New Zealand galaxiids (Barker & Lambert 1988; Allibone & Wallis 1993; Waters *et al.* 2000a), so gene flow among river systems is large enough to overcome genetic differentiation resulting from any natal homing that might exist.

In stark contrast, our work on non-diadromous *G. vulgaris*, a South Island endemic, revealed extensive genetic differentiation among catchments (Figure 1)(Allibone & Wallis 1993), in keeping with long-term isolation in river systems. Concordant differentiation for isozymes (Allibone *et al.* 1996), mitochondrial DNA (Waters & Wallis 2001a,b) and morphology (McDowall & Wallis 1996; McDowall 1997, McDowall & Chadderton 1999) has led to *G. vulgaris* (*sensu lato*) being replaced by a complex of at least six species (McDowall 2000) and four other evolutionarily significant units (ESUs) (Waters & Wallis



Graham Wallis is an Associate Professor in the Department of Zoology at Otago University.

Professor Wallis' research interests include population genetics, molecular evolution, molecular systematics and biogeography of endemic species. He may be contacted at graham.wallis@stonebow.otago.ac.nz

Lise Wallis is a Research Fellow in the Department of Zoology at Otago University.

Dr Wallis is a molecular geneticist, currently comparing transcriptomes of diadromous and non-diadromous galaxiid fishes.

Jon Waters is an Associate Investigator at the Allan Wilson Centre for Molecular Ecology and Evolution and an Associate Professor in the Department of Zoology at Otago University.

Professor Waters' research interests include rates of molecular evolution, freshwater phylogeography and drainage evolution, and marine biogeography.

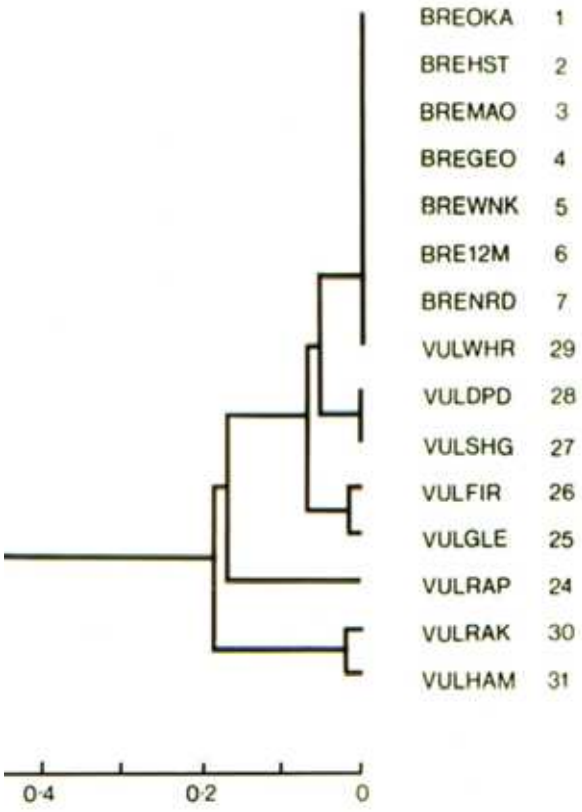


Figure 1. UPGMA phenogram based on Nei's genetic distance (Nei 1972) among seven population samples of* Galaxias brevipinnis *(BRE 1-7) and eight population samples of the* G. vulgaris *complex (VUL 24-31). (Redrawn from Allibone & Wallis 1993)

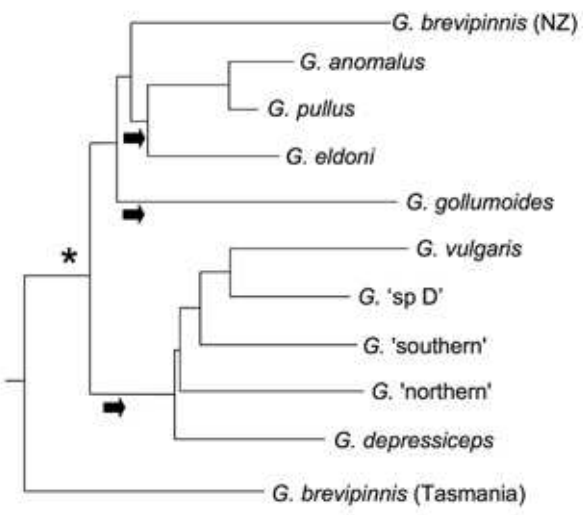


Figure 2. Summary mtDNA tree for nine lineages of the* Galaxias vulgaris *complex, showing their relationships with* G. brevipinnis *from New Zealand and Australia. The asterisk represents the common ancestor of a monophyletic New Zealand group. The arrows represent the minimum number independent losses of diadromy required, assuming that the trait is not reversible. This phylogram was derived by maximum likelihood analysis (GTR + I + Γ) of 5039 bp of mtDNA. (Redrawn from Waters & Wallis 2001)

2001a,b). In one case, broad sympatry of a species pair (Waters *et al.* 2001b) confirms species status under a biological species concept (Mayr 1942). In another case, there is long-term coexistence in parapatry, with only minimal hybridisation (Allibone *et al.* 1996). In several other cases, two or more of these species are found in the same river system with little or no evidence of hybridisation, although the opportunity for parapatry and sympatry must have been extensive before the introduction of salmonids fragmented distributions (Crowl *et al.* 1992).

How many losses of diadromy?

The resolution of a large number of closely related species begs the question: 'How did these species evolve?' One extreme scenario is that each species represents an independent loss of diadromy from a koaro-like ancestor. In contrast, maybe diadromy was only lost once, and most speciation took place within the freshwater environment. This question can be answered using a phylogenetic approach: the first scenario predicts that *G. brevipinnis* branched off at the base of the evolutionary tree leading to the *G. vulgaris* group; the latter predicts a 'comb-like' tree structure of non-diadromous species branching of a koaro like-lineage, with koaro nested inside at the crown of the tree, sister to the last species that it 'spawned'.

Mitochondrial DNA analysis suggested that the real answer lay somewhere in between – namely that the radiation of the *G. vulgaris* species complex was consistent with three losses of diadromy (Figure 2)(Waters & Wallis 2001a). This phylogeny was based on an extensive dataset of 5039 bp, and all nodes in the tree had good statistical support.

A central tenet of molecular systematics is that gene trees reflect species trees. As lineages undergo cladogenetic events, their respective genes cease to exchange genetic information, and start to accumulate new mutations independently of each other. If the rate of molecular evolution is high compared to the rate of proliferation of new lineages, then this expectation is usually met. Consequently, there is generally good agreement among gene trees tracing the evolution of orders of insects or mammalian families (Penny *et al.* 1982), for example. If, however, we are attempting to resolve numerous speciation events over a short time frame, there may be inadequate information to resolve branching patterns. Furthermore, retention of different ancestral polymorphisms by different lineages ('lineage sorting') can lead to disagreement among gene trees (Pamilo & Nei 1988). This situation can be made much worse by either hybridisation or selection. These three processes lead to a decoupling of gene histories both from each other, and from species histories (Ballard & Whitlock 2004).

More recent analysis of the molecular phylogenetics of the *G. vulgaris* group has included three nuclear genes (*S7, RAG-1, Numt*), and in contrast to the earlier mtDNA paper, now place *G. brevipinnis* sister to the *G. vulgaris* group (Figure 3; Waters *et al.* in prep). That is, if the broader nuclear gene analysis is to be believed, only a single loss of diadromy is invoked. But is this biologically plausible? Once a non-diadromous lineage has evolved in isolation, how can it spread and diversify into other freshwater systems if its young no longer migrate to sea?

Allopatric speciation by vicariant geological processes?

The answer may lie in the turbulent geological history of New Zealand, and of South Island in particular. Our position on the Pacific and Indo-Australian tectonic plate boundary has led to extensive faulting and uplift. Through rapid uplift, erosion, and

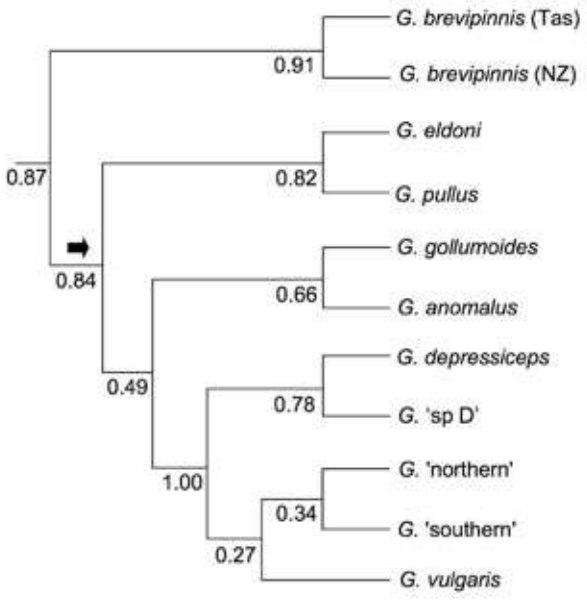


Figure 3. Summary nuclear gene tree for nine lineages of the* Galaxias vulgaris *complex, showing their relationships with* G. brevipinnis *from New Zealand and Australia. Note that in contrast to Fig. 2,* G. brevipinnis *is monophyletic and sister to the* G. vulgaris *complex, implying a single loss of diadromy. This phylogram was derived by Bayesian analysis of 4131 bp of nuclear DNA (Waters* et al.*, in prep).

wholesale tilting of surfaces (individually or in combination), one river catchment can 'capture' the headwater of another (Bishop 1995; Mortimer & Wopereis 1997; Craw *et al.* 2003, 2008). Such geological scenarios can be tested by asking whether species distributions match ancient or current connections (Mayden 1988). With DNA sequence data, one can make more fine-scale predictions and potentially apply molecular clocks to compare genetical and geological timing. We have looked at multiple locations around New Zealand where headwater capture is anticipated from geological evidence, and confirm that headwater capture in association with faunal capture has happened several times (Waters *et al.* 2001a, 2006; Burridge *et al.* 2006, 2007, 2008b; Craw *et al.* 2007).

Calibrating the molecular clock using geological timing

Studies of molecular phylogenetics are often thwarted by an absence of knowledge of rate of molecular evolution specific to the genes and species in question (Arbogast *et al.* 2002). Some classic studies have made use of major geological events to calibrate the clock, such as the emergence of the Panama Isthmus (Knowlton & Weigt 1998) or formation of the Hawaiian islands (Beverley & Wilson 1985; Fleischer *et al.* 1998). These processes are, however, quite inexact with respect to the timing of biotic splits. The isthmus arose over a long period of time: when exactly did gene flow between Pacific and Atlantic populations cease, and could it be different for different species? Although the formation of each of the islands of the Hawaiian chain may be well dated, how long was it before colonisation took place, and did gene flow continue for a while? In contrast, headwater capture events constitute fairly quick and clean severance of gene flow; a headwater does not flow into two different catchments for very long. Additionally, a galaxiid headwater population is likely to be small (compared with major ocean or Hawaiian island populations anyway), and therefore have little genetic variation. Coalescence within a population should be recent, and sorting of ancestral polymorphisms should not constitute too much of a problem.

Calibration of a molecular clock for galaxiid mtDNA in this way revealed an interesting result. In keeping with some recent observations by other researchers, the estimated rate of molecular evolution depends on the time-frame over which divergence is measured (Ho *et al.* 2005). That is, when divergence is measured using recent events (<2 Ma), inferred rates are higher than when measured over longer time periods (Lambert *et al.* 2002; Waters *et al.* 2007). One possible explanation for this 'lazy-J effect' (Penny 2005) is that over short time periods, distances among haplotypes are inflated by deleterious substitutions that tend not to be incorporated into long-term evolutionary lineages. This explanation is problematic as it requires high N_e (Woodhams 2006). Calibration from any one single event (in this case, the Kaituna river capture) is still open to considerable error, because of uncertainty of geological timing. One way to minimise such error is to use multiple dates to calibrate divergence, which can tightly constrain a line of best fit (Wilson *et al.* 1987). More recent analysis of galaxiids using nine calibration points, including six river captures/reversals, showed that this rate increase is restricted to the last 200 Ma (Burridge *et al.* 2008a). This initial apparent accelerated rate could be due to the sorting of polymorphisms after splitting (Peterson & Masel 2008). The larger effect seen with ancient DNA studies (Lambert *et al.* 2002; Ho *et al.* 2005; Hay *et al.* 2008) may reflect population replacement, technical problems recovering sequences or artifact of simulations (Emerson 2007).

Speciation at the molecular level – what makes a fish stay at home?

We now have a wealth of data on the geographic distribution of lineages of non-diadromous galaxiid fishes, the nature of the genetic differences among them, and some geological hypotheses for their range and spread. But what is the underlying genetic cause of loss of diadromy? Over thirty years ago, it was suggested that protein sequence differences between humans and chimps were likely to be too trivial to explain the differences in anatomy and behaviour, and instead, differences in gene expression may be important (King & Wilson 1975). This perceptive and prophetic view is gaining traction as our understanding of molecular genetics improves. Adaptation and speciation may be mediated more by changes in transcription factors, binding sites, promoters, enhancers and micro RNAs, than by changes in structural genes. It appears that both structural gene changes and their regulation both require consideration when searching for the molecular basis of speciation (Hoekstra & Coyne 2007).

New technologies are giving us the ability to look at both DNA sequence and gene expression across the entire genome, even in non-model organisms (Vera *et al.* 2008). Pyrosequencing of DNA allows us to assess the abundance and sequence of mRNAs from any particular tissue of any organism, giving access to any transcriptome that we want. As loss of diadromy has occurred repeatedly in galaxiids, it is our hope that comparison of related diadromous and non-diadromous pairs of species via pyrosequencing of amplified cDNAs might reveal the genetic mechanism of this life history shift. Initial data reveal some large differences in gene expression between the

Table 1. A comparison of gene expression between larval* Galaxias brevipinnis *(103,293 reads; 5748 contigs) and* G. depressiceps *(133,436 reads; 10184 contigs) by pyrosequencing. The first 16 genes are selected on the basis of the magnitude of expression difference (ratio) between species. Seven other highly expressed genes are shown for comparison.

gene	*bre %* reads	*dep %* reads	*bre : dep*
Zn finger RNA binding	0.2	-	∞
tubulin polymn-promoting	0.04	-	∞
protocadherin 19-like	0.6	0.0015	431
reticulon 4	7.8	0.09	83
sex comb on midleg-like 4	0.2	0.004	47
phosphorylase kinase β	0.06	0.002	27
ATPase Na^+/K^+ transport α1	0.3	0.02	17
Xenopus cDNA-like	0.6	0.04	17
smoothelin	0.2	0.01	14
calmodulin 2	0.08	0.36	0.22
apolipoprotein C-I precursor	0.14	0.67	0.21
CK brain	0.3	1.6	0.19
CK mt	1.6	1.1	0.15
sideroflexin 4	0.003	0.05	0.05
embryonic α-globin	-	0.7	0
β-globin	-	0.02	0
RP L4	2.90	1.90	1.3
CK muscle	1.2	1.3	0.9
protein ubiquinone	0.70	1.20	0.6
Gapdh	0.44	0.33	1.3
macropain	0.25	0.30	0.8
ATPase Na^+/K^+ β	0.15	0.08	1.9
cyt b	0.05	0.10	0.5

first pair of species sequenced (Table 1), but this work is in its early stages, and differences need to be confirmed by real-time quantitative PCR.

General implications for New Zealand endemism

Genes modulating such a change can be though of as 'speciation genes', since they cause genetic isolation of populations and ultimately lead to speciation. It is clear that most New Zealand biodiversity has not been evolving independently of other austral species since our isolation from the rest of Gondwana 85 Ma, but largely derives from migrants that have crossed the Tasman (Pole 1994; Winkworth *et al.* 2002; Waters & Craw 2006). Our biota is therefore dominated by lineages of waifs and strays that happen to be good dispersers. New Zealand's limited freshwater fish fauna is a case in point, entirely deriving from species that possess a marine life history phase (McDowall 2000; Waters *et al.* 2000b, 2002). If migration continues, then endemism will be lower; two of our diadromous galaxiids (koaro, inanga) also occur in Australia. It is only after cessation of gene flow, resulting from the types of genetic change that we are trying to identify, that speciation of local endemics ensues.

References

Allibone, R.M.; Crowl, T.A.; Holmes, J.M.; King, T.M.; McDowall, R.M.; Townsend, C.R.; Wallis, G.P. 1996. Isozyme analysis of *Galaxias* species (Teleostei: Galaxiidae) from the Taieri River, South Island, New Zealand: a species complex revealed. *Biological Journal of the Linnean Society 57*: 107–127.

Allibone, R.M.; Wallis, G.P. 1993. Genetic variation and diadromy in some native New Zealand galaxiids (Teleostei: Galaxiidae). *Biological Journal of the Linnean Society 50*: 19–33.

Arbogast, B.S.; Edwards, S.V.; Wakeley, J.; Beerli, P.; Slowinski, J.B. 2002. Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Annual Review of Ecology and Systematics 33*: 707–740.

Ballard, J.W.O.; Whitlock, M.C. 2004. The incomplete natural history of mitochondria. *Molecular Ecology 13*: 729–744.

Barker, J.R.; Lambert, D.M. 1988. A genetic analysis of populations of *Galaxias maculatus* from the Bay of Plenty: implications for natal river return. *New Zealand Journal of Marine and Freshwater Research 22*: 321–326.

Beverley, S.M.; Wilson, A.C. 1985. Ancient origin for Hawaiian Drosophilinae inferred from protein comparisons. *Proceedings of the National Academy of Sciences USA 82*: 4753–4757.

Bishop, P. 1995. Drainage rearrangement by river capture, beheading and diversion. *Progress in Physical Geography 19*: 449–473.

Burridge, C.P.; Craw, D.; Fletcher, D.; Waters, J.M. 2008a. Geological dates and molecular rates: fish DNA sheds light on time dependency. *Molecular Biology and Evolution 25*: 624–633.

Burridge, C.P.; Craw, D.; Jack, D.C.; King, T.M.; Waters, J.M. 2008b. Does fish ecology predict dispersal across a river drainage divide? *Evolution 62*: 1484–1499.

Burridge, C.P.; Craw, D.; Waters, J.M. 2006. River capture, range expansion, and cladogenesis: the genetic signature of freshwater vicariance. *Evolution 60*: 1038–1049.

Burridge, C.P.; Craw, D.; Waters, J.M. 2007. An empirical test of freshwater vicariance via river capture. *Molecular Ecology 16*: 1883–1895.

Craw, D.; Burridge, C.P.; Waters, J.M. 2007. Geological and biological evidence for drainage reorientation during uplift of alluvial basins, central Otago, New Zealand. *New Zealand Journal of Geology and Geophysics 50*: 367–376.

Craw, D.; Burridge, C.P.; Upton, P.; Rowe, D.L.; Waters, J.M. 2008. Evolution of biological dispersal corridors through a tectonically active mountain range in New Zealand. *Journal of Biogeography 35*: 1790–1802.

Craw, D.; Nelson, E.; Koons, P.O. 2003. Structure and topographic evolution of the Main Divide in the Landsborough–Hopkins area of the Southern Alps, New Zealand. *New Zealand Journal of Geology and Geophysics 46*: 553–562.

Crowl, T.A.; Townsend, C.R.; McIntosh, A.R. 1992. The impact of introduced brown and rainbow trout on native fish: the case of Australasia. *Reviews in Fish Biology and Fisheries 2*: 217–241.

Emerson, B.C. 2007. Alarm bells for the molecular clock? No support for Ho *et al.*'s model of time-dependent molecular rate estimates. *Systematic Biology 56*: 337–345.

Fleischer, R.C.; McIntosh, C.E.; Tarr, C.L. 1998. Evolution on a volcanic conveyer belt: using phylogeographic reconstructions and K–Ar-based ages of the Hawaiian Islands to estimate molecular evolutionary rates. *Molecular Ecology 7*: 533–545.

Hay, J.M.; Subramanian, S.; Millar, C.D.; Mohandesan, E.; Lambert, D.M. 2008. Rapid molecular evolution in a living fossil. *Trends in Genetics 24*: 106–109.

Ho, S.Y.W.; Phillips, M.J.; Cooper, A.; Drummond, A.J. 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution 22*: 1561–1568.

Hoekstra, H.E.; Coyne, J.A. 2007. The locus of evolution: evo devo and the genetics of adaptation. *Evolution 61*: 995–1016.

King, M.-C.; Wilson, A.C. 1975. Evolution at two levels in humans and chimpanzees. *Science 188*: 107–116.

Knowlton, N.; Weigt, L.A. 1998. New dates and new rates for divergence across the Isthmus of Panama. *Proceedings of the Royal Society of London B Biological Sciences 265*: 2257–2263.

Lambert, D.M.; Ritchie, P.A.; Millar, C.D.; Holland, B.; Drummond, A.J.; Baroni, C. 2002. Rates of evolution in ancient DNA from Adélie penguins. *Science 295*: 2270–2273.

Mayden, R.L. 1988. Vicariance biogeography, parsimony, and evolution in North American freshwater fishes. *Systematic Zoology 37*: 329–355.

Mayr, E. 1942. *Systematics and the origin of species*. New York, Columbia University Press.

McDowall, R.M. 1990. *New Zealand Freshwater Fishes: A Natural History and Guide*. Auckland, Heinemann Reed.

McDowall, R.M. 1997. Two further new species of *Galaxias* (Teleostei: Galaxiidae) from the Taieri River, southern New Zealand. *Journal of the Royal Society of New Zealand 27*: 199–217.

McDowall, R.M. 2000. *The Reed Field Guide to New Zealand Freshwater Fishes*. Auckland, Reed Publishing.

McDowall, R.M.; Chadderton, W.L. 1999. *Galaxias gollumoides* (Teleostei: Galaxiidae), a new fish species from Stewart Island, with notes on other non-migratory freshwater fishes present on the island. *Journal of the Royal Society of New Zealand 29*: 77–88.

McDowall, R.M.; Wallis, G.P. 1996. Description and redescription of *Galaxias* species (Teleostei: Galaxiidae) from Otago and Southland. *Journal of the Royal Society of New Zealand 26*: 401–427.

Mortimer, N.; Wopereis, P. 1997. Change in direction of the Pelorus River, Marlborough, New Zealand: evidence from composition of Quaternary gravels. *New Zealand Journal of Geology and Geophysics 40*: 307–313.

Nei, M. 1972. Genetic distance between populations. *American Naturalist 106*: 283–292.

Pamilo, P.; Nei, M. 1988. Relationships between gene trees and species trees. *Molecular Biology and Evolution 5*: 568–583.

Penny, D. 2005. Relativity for molecular clocks. *Nature 436*: 183–184.

Penny, D.; Foulds, L.R.; Hendy, M.D. 1982. Testing the theory of evolution by comparing phylogenetic trees constructed from five different protein sequences. *Nature 297*: 197–200.

Peterson, G.I.; Masel, J. 2008. Quantitative prediction of molecular clock acceleration at short timescales. *Molecular Biology and Evolution 25*.

Pole, M. 1994. The New Zealand flora – entirely long-distance dispersal? *Journal of Biogeography 21*: 625–635.

Vera, J.C.; Wheat, C.W.; Fescemyer, H.W.; Frilander, M.J.; Crawford, D.L.; Hanski, I.; Marden, J.H. 2008. Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. *Molecular Ecology 17*: 1636–1647.

Waters, J.M.; Allibone, R.M.; Wallis, G.P. 2006. Geological subsidence, river capture, and cladogenesis of galaxiid fish lineages in central New Zealand. *Biological Journal of the Linnean Society 88*: 367–376.

Waters, J.M.; Craw, D. 2006. Goodbye Gondwana? New Zealand biogeography, geology, and the problem of circularity. *Systematic Biology 55*: 351–356.

Waters, J.M.; Craw, D.; Youngson, J.H.; Wallis, G.P. 2001a. Genes meet geology: fish phylogeographic pattern reflects ancient, rather than modern, drainage connections. *Evolution 55*: 1844–1851.

Waters, J.M.; Dijkstra, L.H.; Wallis, G.P. 2000a. Biogeography of a southern hemisphere freshwater fish: how important is marine dispersal? *Molecular Ecology 9*: 1815–1821.

Waters, J.M.; Esa, Y.B.; Wallis, G.P. 2001b. Genetic and morphological evidence for reproductive isolation between sympatric populations of *Galaxias* (Teleostei: Galaxiidae) in South Island, New Zealand. *Biological Journal of the Linnean Society 73*: 287–298.

Waters, J.M.; López, J.A.; Wallis, G.P. 2000b. Molecular phylogenetics and biogeography of galaxiid fishes (Osteichthyes: Galaxiidae): dispersal, vicariance, and the position of *Lepidogalaxias salamandroides. Systematic Biology 49*: 777–795.

Waters, J.M.; Rowe, D.L.; Apte, S.; King, T.M.; Wallis, G.P.; Anderson, L.; Norris, R.J.; Craw, D.; Burridge, C.P. 2007. Geological dates and molecular rates: rapid divergence of rivers and their biotas. *Systematic Biology 56*: 271–282.

Waters, J.M.; Saruwatari, T.; Kobayashi, T.; Oohara, I.; McDowall, R.M.; Wallis, G.P. 2002. Phylogenetic placement of retropinnid fishes: data set incongruence can be reduced by using asymmetric character state transformation costs. *Systematic Biology 51*: 432–449.

Waters, J.M.; Wallis, G.P. 2001a. Cladogenesis and loss of the marine life history phase in freshwater galaxiid fishes (Osmeriformes: Galaxiidae). *Evolution 55*: 587–597.

Waters, J.M.; Wallis, G.P. 2001b. Mitochondrial DNA phylogenetics of the *Galaxias vulgaris* complex from South Island, New Zealand: rapid radiation of a species flock. *Journal of Fish Biology 58*: 1166–1180.

Wilson, A.C.; Ochman, H.; Prager, E.M. 1987. Molecular time scale for evolution. *Trends in Genetics 3*: 241–247.

Winkworth, R.C.; Wagstaff, S.J.; Glenny, D.; Lockhart, P.J. 2002. Plant dispersal N.E.W.S. from New Zealand. *Trends in Ecology and Evolution 17*: 514–520.

Woodhams, M. 2006. Can deleterious mutations explain the time dependency of molecular rate estimates? *Molecular Biology and Evolution 23*: 2271–2273.