doi:10.1111/j.1558-5646.2009.00870.x

NATURAL DISTRIBUTIONS OF MITOCHONDRIAL SEQUENCE DIVERSITY SUPPORT NEW NULL HYPOTHESES

John P. Wares

Department of Genetics, University of Georgia, Athens, Georgia 30602

E-mail: jpwares@uga.edu

Received April 7, 2009

Accepted September 18, 2009

A variety of forces and constraints can cause sequence data to deviate from patterns predicted under strict neutrality. Here, I present a meta-analysis of available aligned sequence data from 12 higher animal taxa to identify whether the typical null assumption for an often-used test for neutrality—Tajima's *D* statistic—is an appropriate or useful null given large numbers of empirical observations. Across 1068 cytochrome oxidase I (COI) datasets, the mean value for Tajima's *D* is −0.391, with over a sixth of these datasets representing "significant" divergence from null assumptions according to this test. These results indicate a persistent trend for mitochondrial COI data—chosen for their prevalence in population and taxonomic studies—to indicate patterns of diversity that deviate from a purely neutral description, and provide compelling support for the concept that more complex "null" hypotheses may be necessary in population genetics.

KEY WORDS: **Cytochrome oxidase I, demography, metazoans, neutrality, selection, Tajima's *D*.**

Vermeij (2003) noted that for all that we rely on molecular data for inference into the history and ecology of species in our natural world, little is known about the natural history of these gene regions themselves. There are exemplar taxa for which we understand patterns of variation and the processes that maintain these patterns very well, including the epistatic interactions maintaining deep clade structure among populations of *Tigriopus californicus* (Willett and Burton 2003) or the adaptive and demographic histories that have been reconstructed for humans (Stajich and Hahn 2005). However, at a global level we have a poor general understanding of the mechanisms underlying the diversity we observe.

At the level of gene sequences, this diversity is often used to answer questions about the divergence of populations within species (Wares and Cunningham 2005), or the effective population size of a species (Roman and Palumbi 2003). However, each case is treated as an idiosyncratic evaluation of molecular diversity that starts from the same basic assumptions about mutations. Particularly in studies of mitochondrial diversity in animals, it is convenient to follow Avise et al. (1987) and assume low recombination and effective neutrality—even though we know of clear instances to the contrary (Rand 2001; Smith and Smith 2002).

Perhaps the most crucial of these assumptions is that of effective neutrality, allowing for faithful reconstruction of the demographic history of a species. Recent work (Bazin et al. 2006) has suggested that we treat this assumption with caution; for example, persistent selective sweeps on the mitochondrial genome may limit the effectiveness of these data for estimation of ancestral parameters. Naturally, there are a variety of mechanisms that can decouple genetic diversity from census size (Wakeley 2008; Wares and Pringle 2008), but neutrality itself often remains a poorly evaluated null model (Dowling et al. 2008).

Hahn (2008) argued that now is the time to determine whether there are novel sets of assumptions that could be used for coestimating selection and demographic history. We know that a variety of forces and constraints can cause sequence data to deviate from patterns predicted under strict neutrality (Fu 1997); these forces can act on their own, or may have additive effects on the diversity found in a population. Here, I present a meta-analysis

of available aligned sequence data from 12 higher animal taxa to identify whether the typical null assumption for an often-used test for neutrality—Tajima's (1989) D statistic (D_T)—is an appropriate or useful null given large numbers of empirical observations.

Specifically, it is typically assumed that under neutrality and demographic equilibrium, D_T will be approximately zero; deviations from this value are assessed for each dataset through coalescent simulation, and when D_T is more negative (or positive) than a large proportion of parametric simulations, it is considered to reject a hypothesis of a stable population that is evolving neutrally. If, however, large numbers of species consistently deviate from this null expectation for D_T it could suggest that (1) some evolutionary mechanisms are habitually ignored or (2) we have an opportunity to establish new guidelines for whether a species presents an "interesting" or "significant" deviation in terms of molecular diversity.

Methods

Perl and Python scripts (available from author), modified from the NCBI EUTILS script (written by Oleg Khovayko, www.ncbi.nlm.nih.gov/entrez/query/static/eutils_example.pl), were developed to accept search criteria for the NCBI Popset database, then retrieve and convert these Genbank alignments into species-specific (based on appropriate text field) FASTA alignments indexed by the length of aligned sequence data. Because Popset alignments may be generated from population genetic or phylogenetic studies of single or multiple species, they are not appropriate for population diversity analysis until separated into individual files by species, that is, without outgroup. Similar databases are available (Bazin et al. 2005) but the current scripts are more flexible, based on contemporary availability in Genbank, and eliminate most heterospecific data (except for indexing or cryptic diversity problems, see Results) from resulting sequence alignments.

Searches were generated for 12 higher animal taxa for the mitochondrial cytochrome oxidase I (COI) gene region. Alignments were kept as curated by NCBI. Any resulting alignment consisting of fewer than five sequences was excluded. Remaining FASTA alignment files were analyzed using the libsequence-based COMPUTE program (Thornton 2003); this command-line software generates the entire series of summary statistics typical of population genetic analysis, including π and Tajima's D_T. Significance tests were generated from 10,000 random genealogies (using the "fixed segregating sites" method) assuming no recombination, holding sample size constant with the empirical data file.

When two or more analyses are present for a single taxon, either due to separation of the original dataset due to different sequence lengths or due to multiple published analyses of the same species, all analyses were included. If there were substantially divergent results among these replicates for a single named taxon, further consideration was given to the original datasets. For further assessment of the search method, potential curation biases to the results, and detailed analysis of one of the more well-studied taxa (in numbers of datasets), all remaining data in the Crustacea ($n = 301$ datasets) were then classified based on the NCBI Taxonomy database. Focal taxa that were not part of the higher taxon (e.g., Popset outgroups or inaccurately indexed) were excluded. Each alignment was visually assessed and analyzed for genealogical structure using the neighbor-joining algorithm implemented in PAUP*4.0b10 (Swofford 2002). For each resultant tree, the longest branch length was recorded along with the ratio of sampled individuals separated by that branch. Each dataset was bootstrapped 1000 times and the maximum bootstrap value was recorded, along with the number of branches for which this value exceeded 90. Similar results were obtained with other phylogenetic methods (not shown), but the NJ approach tended to more readily identify such phylogenetic support and so was used for a conservative classification of results.

Sequential analyses of variance (ANOVA) of these Crustacean datasets by taxonomic level were performed in the R analytical environment. Linear models were generated for associations between parameter values (π, D_T) and taxonomic level (class, subclass, order, and family); from these results a mixed-effects model ANOVA was performed and summarized. Normality of the input data were assessed with a Shapiro–Wilk test.

Results

Scans for sequence diversity at the mitochondrial COI gene region of 12 higher taxa (Table 1) included five vertebrate classes (Aves, Euteleostei, Mammalia, Otocephala, and Sauropsida) and seven invertebrate phyla, subphyla, or classes (Bivalvia, Cnidaria, Crustacea, Eleutherozoa, Gastropoda, Hexapoda, and Polychaeta). Nucleotide diversity (π) for these datasets ranged from mean values of 0.0029 (otocephalan fish) to 0.1125 (cnidarians). Across 1068 COI datasets, the mean $D_T = -0.391$, with over a sixth of these datasets representing "significant" divergence from null assumptions according to this test (corroborated by significant values for Fu and Li's D^* and F^* tests, which have very similar means and correlations of 0.78 and 0.87, respectively, with D_T across these results). There is a negative but nonsignificant correlation ($r = -0.173$) between sample size of each dataset and D_T; increased sampling recovers rare alleles more than recovering additional common alleles (Wakeley and Takahashi 2003), there could be a "pooling effect" across sample sites (Städler et al. 2009), and some datasets may suffer from excess singleton diversity due to polymerase or human-introduced error (Tindall and Kunkel 1988;

Table 1. Meta-analysis by taxon of available Popset (NCBI) datasets for the mitochondrial cytochrome oxidase I (COI) locus. Further analysis is presented subsequently for the Crustacea. Number of datasets with a significantly low (based on 10^5 simulations) value of Tajima's *D* are indicated in the last column; of these, the number that are verified with a significantly low *D or *F** (Fu 1997) are shown.**

Taxon	Datasets	Sample Size	π (mean)	D_T (mean)	$P<0.05$ (D^* or F^*)
Aves (class)	59	10.1±12.9	0.0104±0.0157	−0.103	4 (2)
Bivalvia (class)	78	16.1±16.5	0.0499±0.1281	−0.610	24 (17)
Cnidaria (phylum)	26	14.5±14.0	0.1125±0.1424	−0.633	10 (7)
Crustacea (subphylum)	301	25.7±42.8	0.0321±0.0446	−0.496	53 (52)
Eleutherozoa (subphylum)	45	45.9±92.9	0.0316±0.0925	−0.515	12 (11)
Euteleostei (class)	29	7.17±4.57	0.0163±0.0566	−0.552	3 (1)
Gastropoda (class)	40	15.9±18.3	0.0301±0.0435	−0.540	7 (7)
Hexapoda (subphylum)	107	20.0±29.0	0.0186±0.0336	−0.429	21 (19)
Mammalia (class)	86	11.7±8.80	0.0243±0.0658	−0.300	10 (9)
Otocephala (class)	200	7.94±3.28	0.0029±0.0072	−0.256	33 (24)
Polychaeta (class)	20	56.6±92.4	0.0408±0.0452	−0.577	4 (4)
Sauropsida (class)	77	11.0±13.2	0.0162±0.0280	−0.039	8 (7)
OVERALL	**1068**			−0.391	189 (17.6%)

Cariello et al. 1991). However, excluding taxa for which the mean sample size is greater than 20 still recovers a mean D_T of −0.384.

Detailed evaluation of each alignment within the Crustacea (largest dataset evaluated) revealed a screening artifact: some (23 of 301) alignments are included that are not members of the focal taxon due to inappropriate indexing or inclusion in the original published alignment, and some alignments lack sufficient diversity for estimation of π or D_T. After screening for these errors, 278 datasets from 220 distinct species (class Branchiopoda n = 36; class Malacostraca n = 192; class Maxillopoda n = 48) remained (in some cases the same species has been analyzed multiple times, see Methods). Average π is 0.033 ± 0.044 across all crustacean samples (0.030 ± 0.038 if only largest-sample dataset for any given species included); mean D_T is −0.489 ± 1.100 (−0.496 ± 1.080 for only single dataset per species). Gene tree structure had clear effects on these results (Fig. 1); when analysis was only performed on datasets for which the maximum bootstrap value was <90, D_T decreased further to −0.906 (n = 81), with substantial reduction in this mean as more stringent thresholds were considered [for maximum bootstrap of 70, D_T is −1.029 (n = 54); for no bootstrap values >50, D_T is −1.145 (n = 17)]. As datasets were included with clear phylogenetic structure, the imbalance of the tree around the longest branch became more important: in those datasets with only a single branch with strong bootstrap support, D_T was 0.048 for datasets where the length:imbalance ratio ≤5 (i.e., where similar size clades are on either side; n = 30), and −1.221 when that ratio >5 (i.e., when a minority of highly divergent sequences are included; n = 30). With increasing numbers of strongly supported branches/clades, the interpretation of D_T becomes less meaningful but the average is −0.026 (n = 118). Similar results were obtained for extended analysis of the Otocephala and Mollusca datasets (results not shown).

ANOVA of π by taxonomic rank generated a significant association across orders (Table 2A, $P < 0.001$); the same analysis for D_T suggested an association across crustacean classes (Table 2B, $P < 0.01$). Assumptions of ANOVA are violated by these data as they are not normally distributed (Shapiro–Wilk W = 0.972, $P < 0.0001$), with an excess of highly negative values of D_T from expectation under a normal distribution observed in a Q–Q plot. However, similar results were also obtained with a Kruskal–Wallis test, a nonparametric test of medians for the distributions. The variance in π at the level of orders (results not shown) is driven by some taxa (i.e., Kentrogonida, Diplostraca, Mysidacea, Sessilia, Euphausiacea, Decapoda, Notostraca, Anostraca) with quite low diversity (the top quartile of samples from each taxon is less than π = 0.05), but a wide range in π overall. Similarly, the variance in the distribution of D_T among Crustacea is driven primarily by the Maxillopoda (copepods and barnacles); data from these taxa generate more strongly negative values of D_T than the other classes, as all three classes have medians well below zero, but for the Maxillopoda the upper quartile is also negative, with a median D_T below −1.

Discussion

These results indicate a persistent trend for mitochondrial COI data to indicate patterns of diversity, as evidenced by D_T, that deviate from a purely neutral description. Although simulation of such datasets typically generates results with a slightly negative D_T (e.g., 1000 simulations of sample size n = 10 with θ = 5 using Hudson's MS will generate mean D_T of −0.0678 ± 0.9143; see

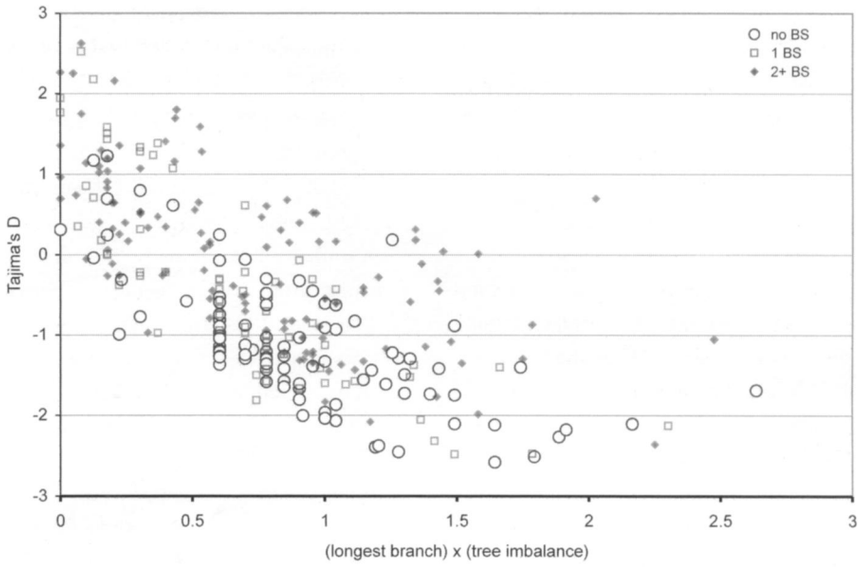

Figure 1. Datasets from Crustacea were evaluated for gene tree characteristics and structure through neighbor-joining analysis of sequence alignments. The longest branch length was recorded along with the ratio of sample sizes on either side of that branch; Tajima's *D* (D_T) is then plotted against the product of these values for datasets that exhibited no significant phylogenetic structure (no bootstrap values >70; "no BS"), or 1 branch with such structure ("1 BS"), or multiple branches with such structure ("2+ BS"). As predicted, the mean D_T for datasets with no structure is even lower than the average across all datasets whereas those datasets with phylogenetic structure (suggesting cryptic lineages or improper taxonomy) have higher means. There is also an expected effect on D_T of the product of the longest branch and tree imbalance, as datasets with only a few lineages that are genetically distant will have lower D_T; those with an even balance across the longest branch (lower product on horizontal axis) tend toward higher D_T. Similar results were obtained for other taxa.

Thornton 2005), the empirical observation is more extreme. What, then, does this additional persistent deviation mean? A negative D_T signifies an excess of low-frequency polymorphisms that may be generated by population expansion or purifying/directional selection. We must therefore assume either that most species have recently not been in demographic equilibrium, or that most genomes (in this case, mitochondrial) are affected by recurrent selective sweeps (Bazin et al. 2006), or that the data are consistent with a high mutation rate opposed by ongoing purifying selection. The last hypothesis seems to be the most likely explanation, following a comparable comparative analysis of diversity at mitochondrial and nuclear loci (Meiklejohn et al. 2007).

Table 2. ANOVA results of π and D_T against taxonomic rank in the Crustacea. Significance indicated with * (P<0.05), ** (P<0.01), *** (P<0.001). Similar results were obtained with a Kruskal–Wallis test.

Rank	df	Sum of Sq	Mean Sq	F	$Pr(>F)$
(A) ANOVA of π by taxonomic rank in the Crustacea					
Class	2	0.00396	0.00198	1.2276	0.2952
Subclass	2	0.01389	0.00694	4.3029	0.0148*
Order	11	0.05694	0.00518	3.2081	0.0004***
Family	61	0.14937	0.00245	1.5176	0.0170*
Residuals	**200**	**0.32270**	**0.00161**		
(B) ANOVA of D_T by taxonomic rank in the Crustacea					
Class	2	12.528	6.264	6.2066	0.0024**
Subclass	2	7.793	3.896	3.8608	0.0226*
Order	11	20.123	1.829	1.8127	0.0538
Family	61	91.707	1.503	1.4896	0.0214*
Residuals	**200**	**201.845**	**1.009**		

In any case, we are referring to deviations from a null hypothesis that is intended to formally describe the statistical behavior of data. If our assumptions of neutrality and demographic equilibrium are routinely violated, then development of a more appropriate set of assumptions may be necessary for evaluating empirical population genetic data. Empirical comparisons of multiple loci within a species' genome for the purpose of identifying outlier gene regions is now routine (Luikart et al. 2003; Ramos-Onsins et al. 2008) and has identified not only the prevalence of selection but also suggested mechanisms, such as codon usage bias (Haddrill et al. 2008); perhaps cross-taxa comparisons of data from individual loci, as done here, can be used for identifying species or populations that are truly extraordinary. For example, of the data plotted in Figure 1, all datapoints below D_T of -1.67 are significant ($P < 0.025$; eight points above this value are also significant), a total of 40 significantly negative results (13.2%). Taking instead the lowest 2.5% (in a two-tailed test) of results

would indicate that only those values below −2.3 (or above 1.8) are extraordinary. Such a comparison is more conservative regarding inference of either selection or expansion, and inference of processes such as diversifying selection might be more frequent, compared to typical parametric evaluation of null expectations.

A problem with any such meta-analytical treatment is the extent to which phylogenetic structure (which may have made publication of the data more likely) within any "species" in GenBank is not fully documented. In some cases, individual species are unofficially designated as distinct clades or types within GenBank; in other cases, such as the isopod *Idotea balthica*, the single undifferentiated dataset almost certainly represents multiple cryptic species (Wares 2001), and there are countless other intermediate forms of genetic structure (Avise 2000) that cannot be automatically considered in this sort of analysis without considering a range of other factors (Städler et al. 2009). To the extent that cryptic species are common and uniformly distributed among animal taxa (Pfenninger and Schwenk 2007), typical database alignments could generate inflated π and elevated D_T (Arunyawat et al. 2007). As shown, eliminating these phylogeographic cases pushes the observed mean D_T further downwards (see Results).

However, the correlation between π and D_T at the level of higher taxa shown in Table 1 is negative (−0.61), i.e., high nucleotide diversity (as might be expected if multiple demographically isolated taxa are commonly included under a single species name) tends to predict more negative values of D_T. So, the observed patterns of diversity involve more than simply our uncertainty about species delineations. These analyses do not include reproductive or dispersive life-history information, which can predict significant shifts in diversity patterns across large taxonomic groups (Foltz et al. 2004). One of many considerations for future database work is the extent to which NCBI and similar databases can begin to include ecological and life-history data that may be necessary for broad conclusions to be attained from previous work. At this point, we simply have too little information to speculate on why observed π is high in the Cnidaria (an observation opposite to typical conclusions of intraspecific studies in this group, Shearer and Coffroth 2008) or D_T tends to be lower in the Maxillopoda, the order comprised of copepods and barnacles. It is becoming clear, however, that standing diversity may vary among taxonomic groups for a variety of reasons, including variation in mutation rate (Ellegren 2009; Nabholz et al. 2009), and thus any empirical contrast for a study should be made using a sample from related taxa. Recognizing that there is an empirical negative bias to this summary statistic, and that it may vary among taxa in persistent ways, is an interesting component of understanding how variation is maintained in natural populations.

The presented results are based on a limited range of summary statistics (see Supporting Information Table S1). The statistical package used here for high-throughput reanalysis does not currently support Fu's Fs, which is a powerful test for detecting population growth and hitchhiking (Fu 1997), and comparison with other recently developed tests would be of interest for any single taxon to better understand the demographic and selective history that led to current observed diversity (Zeng et al. 2006), although these tests generally present correlated results (here, D_T and the F^* and D^* tests are significantly correlated ($P < 0.01$) and have similar means). Finally, of course, this analysis only represents data from a single, commonly analyzed gene in metazoans (Folmer et al. 1994; Hebert et al. 2003). Mitochondrial COI harbors an unusually low ratio of nonsynonymous-to-synonymous diversity (Mishmar et al. 2003; Seo et al. 2004; Ingman and Gyllensten 2007); many datasets have only a few nonsynonymous polymorphisms among many silent mutations, suggesting that removal of these few sites for analysis (more cumbersome in meta-analysis because coding frame is not directly identifiable in all datasets) of D_T and other neutrality tests would only slightly change the results (and supporting the idea that purifying selection is important at this locus). However, any single gene region may not represent general evolutionary dynamics particularly well, and such analysis should be expanded to include other appropriate gene regions (including those commonly employed in intraspecific analysis of fungi, plants, and unicellular taxa). Preliminary analysis of other common genes for population genetic studies of Crustacea, for example, generated mean D_T of −0.52 for ribosomal ITS and mean D_T of −0.23 for elongation factor 1 (J. P. Wares, unpubl. data). Eventually, of course, the increasing availability of abundant genomic data for nonmodel organisms will permit all researchers to tease apart the role of selection and demography as general forces influencing diversity in natural populations.

Nevertheless, these results provide compelling support for the concept that "alternative" null hypotheses may be necessary in evolutionary genetics (Hahn 2008). As with testing data to see if it evolves in a clock-like manner, the molecular clock model is a special case of the more general model of variable rates across lineages (Huelsenbeck and Rannala 1997); it may be that accepting the neutral model should be a special case of the more general—and more complex—model of evolution, or Bayesian approaches that consider several a priori mechanisms and their simultaneous effects on sequence data may need to be developed (Hickerson et al. 2006). It should be emphasized that the goal of this study was not to indicate that selection may affect diversity at some gene regions—-there is already ample evidence for selection in mitochondrial datasets e.g., (Rand and Kann 1996; Rand 2001; Meiklejohn et al. 2007)—nor to indicate that D_T is a summary statistic that may be biased if all underlying evolutionary mechanisms are not considered, as Thornton (2005) has shown. Instead, this study shows simply that there is a persistent effect on diversity that is unaccounted for by most tests of the "null" hypothesis.

Recognition of these problems and other shortcomings of such test statistics is not new among mathematical and theoretical geneticists (Wayne and Simonsen 1998; Thornton 2005); however, many researchers in the fields of molecular ecology and phylogeography routinely apply tests such as D_T while assuming neutrality so that demographic inferences about their populations can be made. It is clear that without full consideration of the evolutionary history of a population, such demographic inference can be inaccurate or misleading. In the case of the results shown here, the generality of negative D_T values suggests that inference of population history from such tests will require more sophisticated evaluation, and additional genomic data.

ACKNOWLEDGMENTS

I thank J. Ross-Ibarra, M. Hahn, J. DeBarry, H. Wang, K. Dyer, D. Hall, M. Hickerson, J. Meléndez, and S. Small for intellectual and technical support in the development of this article. D. Rand and an anonymous reviewer improved this article greatly with their suggestions and comments.

LITERATURE CITED

Arunyawat, U., W. Stephan, and T. Städler. 2007. Using multilocus sequence data to assess population structure, natural selection, and linkage disequilibrium in wild tomatoes. Mol. Biol. Evol. 24:2310–2322.

Avise, J. C. 2000. Phylogeography. Harvard Univ. Press, Cambridge, MA.

Avise, J. C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb, J. E. Neigel, C. A. Reeb, and N. C. Saunders. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Ann. Rev. Ecol. Syst. 18:489–522.

Bazin, E., L. Duret, S. Penel, and N. Galtier. 2005. Polymorphix, a sequence polymorphism database. Nucleic Acids Res. 33:D481–D484.

Bazin, E., S. Glémin, and N. Galtier. 2006. Population size does not influence mitochondrial genetic diversity in animals. Science 312:570–572.

Cariello, N. F., J. A. Swenberg, and T. R. Skopek. 1991. Fidelity of *Thermococcus litoralis* DNA polymerase (vent) in PCR determined by denaturing gradient gel electrophoresis. Nucleic Acids Res. 19:4193–4198.

Dowling, D. K., U. Friberg, and J. Lindell. 2008. Evolutionary implications of non-neutral mitochondrial genetic variation. Trends Ecol. Evol. 23:546–554.

Ellegren, H. 2009. Is genetic diversity really higher in large populations? J. Biol. 8:41.

Folmer, O., M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Marine Biol. Biotechnol. 3:294–299.

Foltz, D. W., A. W. Hrincevich, and A. Rocha-Olivares. 2004. Apparent selection intensity for the cytochrome oxidase subunit I gene varies with mode of reproduction in echinoderms. Genetica 122:115–125.

Fu, Y. X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915–925.

Haddrill, P. R., D. Bachtrog, and P. Andolfatto. 2008. Positive and negative selection on noncoding DNA in *Drosophila simulans*. Mol. Biol. Evol. 25:1825–1834.

Hahn, M. W. 2008. Toward a selection theory of molecular evolution. Evolution 62:255–265.

Hebert, P. D. N., S. Ratnasingham, and J. R. deWaard. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Philos. Trans. R Soc. Lond. B 270S:1–4.

Hickerson, M. J., E. Stahl, and H. Lessios. 2006. Test for simultaneous divergence using approximate Bayesian computation. Evolution 60:2435–2453.

Huelsenbeck, J. P., and B. Rannala. 1997. Phylogenetic methods come of age: testing hypotheses in an evolutionary context. Science 276:227–232.

Ingman, M., and U. Gyllensten. 2007. Rate variation between mitochondrial domains and adaptive evolution in humans. Hum. Mol. Genet. 16:2281–2287.

Luikart, G., P. R. England, D. Tallmon, S. Jordan, and P. Taberlet. 2003. The power and promise of population genomics: from genotyping to genome typing. Nat. Rev. Genet. 4:981–994.

Meiklejohn, C. D., K. L. Montooth, and D. M. Rand. 2007. Positive and negative selection on the mitochondrial genome. Trends Genet. 23:259–263.

Mishmar, D., E. Ruiz-Pesini, P. Golik, V. Macaulay, A. G. Clark, S. Hosseini, M. Brandon, K. Easley, E. Chen, M. D. Brown, et al. 2003. Natural selection shaped regional mtDNA variation in humans. Proc. Natl. Acad. Sci. USA 100:171–176.

Nabholz, B., S. Glémin, and N. Galtier. 2009. The erratic mitochondrial clock: variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. BMC Evol. Biol. 9:54.

Pfenninger, M., and K. Schwenk. 2007. Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. BMC Evol. Biol. 7:121.

Ramos-Onsins, S. E., E. Puerma, D. Balañá-Alcaide, D. Salguero, and M. Aguadé. 2008. Multilocus analysis of variation using a large empirical data set: phenylpropanoid pathway genes in Arabidopsis thaliana. Mol. Ecol. 17:1211–1223.

Rand, D. M. 2001. The units of selection on mitochondrial DNA. Annu. Rev. Ecol. Syst. 32:415–448.

Rand, D. M., and L. M. Kann. 1996. Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from Drosophila, mice, and humans. Mol. Biol. Evol. 13:735–748.

Roman, J., and S. R. Palumbi. 2003. Whales before whaling in the North Atlantic. Science 301:508–510.

Seo, T., H. Kishino, and J. L. Thorne. 2004. Estimating absolute rates of synonymous and nonsynonymous nucleotide substitution in order to characterize natural selection and date species divergences. Mol. Biol. Evol. 21:1201–1213.

Shearer, T. L., and M. A. Coffroth. 2008. Barcoding corals: limited by interspecific divergence, not intraspecific variation. Mol. Ecol. Res. 8:247–255.

Smith, J. M., and N. H. Smith. 2002. Recombination in animal mitochondrial DNA. Mol. Biol. Evol. 19:2330–2332.

Städler, T., B. Haubold, C. Merino, W. Stephan, and P. Pfaffelhuber. 2009. The impact of sampling schemes on the site frequency spectrum in nonequilibrium subdivided populations. Genetics 182:205–216.

Stajich, J. E., and M. W. Hahn. 2005. Disentangling the effects of demography and selection in human history. Mol. Biol. Evol. 22:63–73.

Swofford, D. 2002. Phylogenetic analysis using parsimony (PAUP) v. 4.0b10.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595.

Thornton, K. 2003. libsequence: a C++ class library for evolutionary genetic analysis. Bioinformatics 19:2325–2327.

———. 2005. Recombination and the properties of Tajima's D in the context of approximate-likelihood calculation. Genetics 171:2143–2148.

Tindall, K. R., and T. A. Kunkel. 1988. Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. Biochemistry 27:6008–6013.

Vermeij, G. 2003. Foreword: Molecules, Markers, and Mollusks. Pp. vii–x *in* C. Lydeard and D. L. Lindberg, eds. Molecular systematics and phylogeography of mollusks. Smithsonian Books, Washington DC.

Wakeley, J. 2008. Coalescent theory: an introduction. Roberts & Co, Greenwood Village, CO.

Wakeley, J., and T. Takahashi. 2003. Gene genealogies when the sample size exceeds the effective size of the population. Mol. Biol. Evol. 20:208–213.

Wares, J. P. 2001. Intraspecific variation and geographic isolation in *Idotea balthica* (Isopoda: Valvifera). J. Crustacean Biol. 21:1007–1013.

Wares, J. P., and C. W. Cunningham. 2005. Diversification predates glaciation in *Balanus glandula*. Biol. Bull. 208:60–68.

Wares, J. P., and J. M. Pringle. 2008. Drift by drift: effective population size is limited by advection. BMC Evol. Biol. 8:235.

Wayne, M. L., and K. L. Simonsen. 1998. Statistical tests of neutrality in the age of weak selection. Trends Ecol. Evol. 13:236–240.

Willett, C. S., and R. S. Burton. 2003. Environmental influences on epistatic interactions: viabilities of cytochrome c genotypes in interpopulation crosses. Evolution 57:2286–2292.

Zeng, K., Y.-X. Fu, S. Shi, and C.-I. Wu. 2006. Statistical tests for detecting positive selection by utilizing high-frequency variants. Genetics 174:1431–1439.

Associate Editor: T. Crease

Supporting Information

The following supporting information is available for this article:

Appendix S1. Output from population genetic analyses of species across 12 higher taxa.

Supporting Information may be found in the online version of this article.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.