Global patterns of nuclear and mitochondrial genetic diversity in marine fishes

Rene Clark[1] and Malin Pinsky[2]

[1]Drexel University
[2]Rutgers University

October 23, 2023

Abstract

Genetic diversity is a fundamental component of biodiversity. Examination of global patterns of genetic diversity can help highlight mechanisms underlying species diversity, though patterns may differ across the genome. Here, we compiled 6862 observations of genetic diversity from 492 species of marine fish, assessed their associations with macroecological drivers, and tested among hypotheses for diversity gradients: the founder effect hypothesis, the kinetic energy hypothesis, and the productivity-diversity hypothesis. We found that mitochondrial genetic diversity followed geographic gradients similar to those of species diversity, being highest near the equator, particularly in the Coral Triangle, while nuclear genetic diversity did not follow clear global patterns. Despite these differences, all genetic diversity metrics were strongly correlated with chlorophyll-a concentration, while mitochondrial diversity was also positively associated with sea surface temperature. Our results provide support for the kinetic energy hypothesis, which predicts that elevated mutation rates at higher temperatures increase mitochondrial diversity, and the productivity-diversity hypothesis, which posits that resource-rich regions support larger populations with greater genetic diversity. Overall, these findings reveal how environmental variables can influence mutation rates and drift in the ocean, caution against using mitochondrial macrogenetic patterns as proxies for nuclear DNA, and aid in defining global gradients of genetic diversity.

Introduction

At its core, genetic diversity is the foundation upon which biodiversity flourishes. Intraspecific genetic diversity can help drive speciation events by enabling adaptation to novel environments and reduce extinction risk by providing a genomic reservoir during periods of environmental change (Vellend & Geber, 2005). Exploring global trends in genetic diversity can shed light on the mechanisms, or combinations of mechanisms, that drive species diversity. Similarly, elucidating the processes that generate genetic diversity helps create a common ground for evolutionary biology and community ecology around topics of diversity and patterns of speciation (Vellend & Geber 2005). Despite this importance, the general patterns of genetic diversity across species remain poorly understood at global scales (De Kort et al. 2021; Manel et al. 2020; Messmer et al. 2012; Miraldo et al. 2016). Efforts to understand such trends are vital for identifying the factors creating and maintaining biodiversity and for pinpointing high priority areas and taxa for conservation.

Much of our knowledge on intraspecific genetic diversity, including local and regional estimates in various taxa, has only been collected in recent decades. Recent macrogenetic studies have compiled these data to construct broad-scale geographic patterns and better understand global distributions of genetic diversity (Li et al. 2021; Manel et al. 2020; Miraldo et al. 2016). For example, Martin & McKay (2004) revealed greater genetic divergence among vertebrate populations at lower latitudes, while Schmidt et al. (2022) found that environmental heterogeneity was an important predictor of genetic diversity in North American mammals. Large knowledge gaps still exist, however, as the strength and direction of latitudinal gradients in genetic diversity appear to vary across taxa and ecological systems (De Kort et al. 2021). In particular,

it remains unclear how universal such patterns are and how influential underlying ecological drivers may be. This is particularly true of marine communities, as most macrogenetic studies to date have focused on either terrestrial or freshwater systems (but see Manel et al. 2020).

While the same evolutionary processes occur in all taxa, the strength of these forces (particularly selection, drift, and gene flow) differ substantially across terrestrial and marine realms (Steele et al. 2019; Strathmann 1990). Marine species tend to exhibit larger populations, higher gene flow, and wider species ranges (Steele et al. 2019). Alleles may be more easily transported throughout species ranges and latitudes in marine systems, muting the effects of the local environment and weakening the consequences of genetic drift. Such patterns have previously been documented within individual species, including evidence that strong dispersal helped maintain high diversity in range edge populations of Senegal seabream, *Diplodus bellottii* , and erased typical core-periphery patterns of genetic variation (Robalo et al. 2020). Moreover, global patterns of species richness tend to differ between land and sea. Marine taxa commonly display bimodal latitudinal gradients of species richness (Chaudhary 2016; Tittensor et al. 2010), peaking at mid-latitudes instead of along the equator as is more common in their terrestrial counterparts (Davies et al 2007; Rolland et al 2014). Marine species also have strong longitudinal patterns in species diversity, with greatest species biodiversity in the Indo-Pacific Coral Triangle due in part to higher habitat availability and sea surface temperatures (Sanciangco et al. 2013). Given these differences, it remains unclear how environmental conditions and life history strategies in the ocean combine to shape macroecological patterns of genetic diversity. Recent studies have begun to investigate these questions, including Manel et al. (2020)'s finding that mitochondrial genetic diversity in marine fishes was positively correlated with sea surface temperature. However, the mitochondrial genome is a small (less than 0.01%) fraction of the genetic material in fish, and more work is needed to understand the ubiquity of these observed patterns across the genome.

Most macrogenetic studies have investigated patterns of mitochondrial genetic diversity, despite suggestions that such markers do not accurately reflect neutral nuclear genetic diversity (Bazin et al. 2006; Leigh et al. 2021). For example, many mitochondrial markers are linked without recombination to loci under strong selective constraints (Galtier et al. 2009). Mitochondrial diversity is therefore subject to selective sweeps and background selection as well as bottlenecks due to its small effective population size (N_e), which is a quarter that of nuclear DNA (Ballard & Whitlock 2003; Birky et al. 1989). Mitochondrial diversity also does not display a consistent relationship with population size, with strong variation across taxa that is not related to life history characteristics (Bazin et al. 2006; James & Eyre-Walker 2020; Nabholz et al. 2009). With these caveats in mind, macro-scale patterns of mitochondrial genetic variation may not be generalizable to nuclear diversity. To gain a more complete understanding of global distributions of genetic diversity, neutral genetic variation in the much larger nuclear genome should also be analyzed.

Here, we propose three distinct hypotheses for global genetic diversity gradients, all of which are grounded in foundational community ecology and population genetics theory. The first is the Kinetic Energy Hypothesis, which posits that, like species richness, intraspecific genetic diversity should be greater at hotter temperatures due to faster evolutionary tempos (e.g. higher metabolic and mutation rates), particularly in mitochondrial DNA that is affected by oxidative damage from metabolic processes (Allen et al. 2002; Manel et al. 2020; Wright et al. 2011). While oxidative damage should not influence nuclear DNA mutation rates (Hoffman et al. 2004), genome-wide mutation rates are negatively correlated with generation times (Thomas et al. 2010), which are shorter in organisms with smaller body sizes (Martin & Palumbi 1993), and, by Bergmann's rule, inversely related to temperature (Bergmann 1847; Fernández-Torres et al. 2018). Thus, nuclear genetic diversity may also be weakly correlated with temperature, albeit not to the degree of mitochondrial diversity (Gillooly et al. 2004). The second hypothesis, the Productivity-Diversity Hypothesis (Evans et al. 2004), suggests that population size is often constrained by resource availability, such that regions of high primary productivity should support larger populations with greater intraspecific genetic variation, since large populations lose genetic diversity to genetic drift at a slower rate (Charlesworth 2009; Wright 1983). However, this relationship may reverse in regions with particularly high levels of productivity - as more individuals and species are supported, resources become increasingly divided, causing population sizes and subsequently, genetic diversity, to decline (Lawrence & Fraser, 2020). Finally, the Founder Effect

Hypothesis proposes a negative relationship between latitude and genetic diversity, a lasting legacy from the last glacial maximum (LGM) (Hewitt 2000). As species expanded from equatorial to temperate and polar latitudes, a sequential series of founder and bottleneck events along the expansion front may have depleted standing genetic variation and left a latitudinal genetic footprint that is still apparent in many modern populations (Jenkins et al. 2018; Lessa et al. 2010; Mattingsdal et al. 2020). For marine species, this effect could be particularly pronounced in the Northern hemisphere, as many contemporary high-latitude taxa in the Southern Ocean endured the LGM in local polar refugia (Allcock & Strugnell 2012; Fraser et al. 2012). The Founder Effect Hypothesis may also apply more strongly to mitochondrial diversity, as mitochondrial DNA should be more sensitive to bottleneck events and founder effects from the LGM due to its smaller N_e (Birky et al. 1989).

To help better identify and understand global patterns in marine genetic diversity, we conducted a literature search to aggregate georeferenced data from population genetic studies in marine fish species and then used these data to evaluate these three hypotheses. We compiled environmental data on sea surface temperature (SST) and chlorophyll-a concentration (a proxy for primary productivity) and assessed the generality of these hypotheses using both mitochondrial and nuclear (microsatellite) DNA. Specifically, we tested 1) the Kinetic Energy Hypothesis that temperature and genetic diversity will be positively related most strongly in mitochondrial DNA and more weakly in nuclear DNA, 2) the Productivity-Diversity Hypothesis that genetic diversity will be highest in regions with mid-to-high levels of primary productivity (e.g. chlorophyll-a), particularly in nuclear DNA (as it is more closely related to population size, the mediating factor), and 3) the Founder Effect Hypothesis that both mitochondrial and nuclear genetic diversity will be negatively correlated with latitude and increase towards the equator, particularly in the Northern hemisphere and more strongly in mitochondrial diversity. To test among these three hypotheses, we fit generalized linear mixed effect models (GLMMs) and explored the extent to which each macroecological driver explained variation in mitochondrial or nuclear genetic diversity.

2. Materials and Methods

(a) Data collection

We conducted a literature search in Web of Science to build a comprehensive database of published genetic diversity observations in marine fishes. The following keyword search terms were used: *fish* microsatellite* (marine OR ocean OR sea)* and *fish* mtDNA* (marine OR ocean OR sea)* . Only studies published prior to 5 January 2020 were included in the dataset. This was a Class II study in the sense of Leigh et al. (2021) and had the benefits of more easily compiling nuclear diversity data, accounting for allele frequencies in genetic diversity estimates, accounting for methodological covariates that may explain substantial diversity variation, applying more precise data quality filters, and using expert-defined populations that do not inappropriately split or lump different geographic locations. Class II studies also often compile data across fewer species, in contrast to Class III studies that use existing online databases like NCBI or BOLD to download, grid, and analyze unique DNA sequences. During the literature search, we excluded anadromous, catadromous, and estuarine species from the database, as well as data from populations that were captive, farmed, or stocked. We also excluded data from studies that either did not report the corresponding latitudinal & longitudinal coordinates, or only vaguely identified the sampling location (precision less than 3°). For a more detailed explanation of further exclusion criteria see the Supplemental Materials.

We recorded expected heterozygosity (H_e) for microsatellite studies, and nucleotide diversity (π) or haplotype diversity (H_d) for mtDNA studies as reported. The standard errors of H_e , $H_d$or π were also recorded (or calculated from the standard deviations), when provided. All measures of genetic diversity were recorded at the population level. For mtDNA, marker length (in base pairs) was recorded. For microsatellite studies, we recorded whether or not the primers were originally developed in a different species, because cross-species amplification can negatively influence diversity estimates (Barbará et al. 2007). When possible, we recorded H_e on a per-marker basis, though some studies reported only average heterozygosity across markers. For these studies, we listed each locus separately and extrapolated per-marker diversity by adding a normally distributed error to the average diversity estimate (Pinsky & Palumbi 2014). This error distribution had a

standard deviation equal to that reported within the study. If a within-study standard deviation was not available, we used the average standard deviation (0.24) reported across all studies.

In addition to following global patterns, genetic diversity often declines towards a species' range margin, as populations at the edge tend to be smaller in size relative to those at the range center (Clark et al. 2021; Eckert et al. 2008). To help account for these cross-range effects, which may be distinct from latitudinal effects, we used the R package rfishbase v.3.1.6 (Boettiger et al. 2011) to download species range data from Aquamaps (Kaschner et al. 2019). We then calculated the latitudinal range position of each sampled population in our database. This value ranged from 0 to 1, with 0 indicating the population was located at the center of its species range and 1 indicating the population was located at the very northern or southern edge of its species range. Finally, we also recorded the order, family, and genus for each species.

(b) Model structure

To compare our hypotheses, we fit generalized linear mixed effect models (GLMMs). This mixed effects framework allowed us to account for other factors that could influence genetic diversity but that were not the focus of our study. For models with log-transformed π as the response variable, we ran linear GLMMs with a Gaussian error term using the lme4 package v.1.1.26 (Bates et al. 2014). For models withH_e or H_d as the response variable, we ran beta GLMMs using the glmmTMB package v.1.1.7 (Brooks et al. 2017). All beta models were run specifying the ordbeta family, which uses a logit link function and enables the incorporation of 0 and 1 values into the model (Kubinec 2022). For the mtDNA models ofH_d , the length of the marker in base pairs was included as an explanatory variable. For the microsatellite models, we included whether or not the microsatellite primer was cross-species amplified. Marker length and cross-species amplification, as well as range position, were all scaled and centered to have a mean of 0 and a standard deviation (SD) of 1. We incorporated the source (the study the data came from) as a random intercept for all models to help account for other study-specific methodological choices, while marker name (the specific mtDNA marker used) was added as a random intercept for the mtDNA models to help account for marker-specific mutation rates and selective constraints. Marker name was included as a random intercept because we recorded mtDNA genetic diversity from across the mitogenome and did not limit our dataset to COI or cyt-b markers. Finally, a nested genus/family random intercept was added to all models to account for phylogenetic relationships.

For each estimate of diversity (π, H_d , orH_e), we fit a series of five models to identify geographic patterns: (1) a baseline model with just the terms and random effects specified above, (2) a latitude model, (3) an absolute latitude model, (4) a longitude model, (4) a latitude and longitude model, and (5) an absolute latitude and longitude model. The latitude and longitude models contained the predictor variable of interest (e.g. latitude, longitude, etc.) in addition to the baseline model structure. Latitude, absolute latitude, and longitude were all scaled and centered (mean 0, SD 1). Latitude was included as a quadratic term to allow a peak in the tropics, while longitude was incorporated as a smoothing spline using the R package splines v.4.2.2 (R Core Team 2023) to account for its uniquely circular nature.

We used the same model structure to compare macroecological drivers of genetic diversity. As with the latitude and longitude models, we fit a series of models that incorporated either mean sea surface temperature (SST) (°C), mean chlorophyll-a concentration (mg/m^3), or both. SST was scaled and centered (mean 0, SD 1) and chlorophyll-a was log-transformed and included as a quadratic term. All environmental data were monthly climatologies (9.2 km^2 resolution) and were extracted from Bio-ORACLE (Assis et al. 2017; Tyberghein et al. 2012) using the R package sdmpredictors v.0.2.10 (Bosch & Fernandez 2021).

(c) Model comparisons

To compare among models, the model with the lowest Akaike's Information Criterion (AIC) was considered the best or most parsimonious model (Burnham & Anderson 2002). Marginal and conditional R^2 values were calculated with the performance package v.0.10.4 (Lüdecke et al. 2021, Nakagawa & Schielzeth 2013). Within each model, to identify which variables most influenced patterns of genetic diversity, we plotted marginal effects with the R package sjPlot (Lüdecke 2021) and additionally examined the p -values of variable coefficients. Model fits and spatial autocorrelation in the residuals were checked with DHARMa

v.0.4.3 (Hartig 2021). Moran's I was near zero for all models, and no significant spatial autocorrelation (defined as $p < 0.5$) was found (Table S1). To assess sensitivity to missing and rare data, all models were bootstrapped 1000x with the boot package v.1.3.28 (Canty & Ripley 2021). All analyses were performed in R v.4.2.3 (R Core Team 2023).

Finally, to identify whether global patterns varied across taxa, we ran all models on a subset of 10 families (Scombridae, Lutjanidae, Serranidae, Pomacentridae, Sebastidae, Engraulidae, Gadidae, Syngnathidae, Rajidae, and Carcharhinidae), 1 family at a time. These 10 families were chosen as they (1) had a large amount of data (> ˜30 observations/dataset) and (2) represented a broad range of life history traits.

Results

(a) Data collection

For our mitochondrial π dataset, we compiled 1781 population-level measurements of genetic diversity, while for $H_d$we compiled 1871 diversity measurements. Collectively, these observations came from 239 studies and represented 262 species in 82 families. For microsatellites, we recorded genetic diversity (H_e) from 3210 populations, 578 studies, and 341 species in 86 families. When recorded for the same population, nuclear H_e was not strongly correlated with either mitochondrial π or H_d(H_e -π $r_s = 0.242$;H_e -H_d $r_s = 0.349$) although π and H_d were positively related to each other (π-H_d $r_s = 0.818$) (Fig. S1). Mean chlorophyll-a concentration and mean SST were also not strongly correlated with each other (r_s = -0.316) (Fig. S2 & S3).

These nuclear and mitochondrial datasets represented populations from across the globe, spanning all latitudes, every ocean basin, and a wide array of environmental conditions (Fig. 1, S4-S6). Coastlines in the Northern hemisphere were the most densely sampled regions in our database. However, there were also a large number of diversity estimates near the Equator, particularly in the Coral Triangle. While the number of datapoints decreased towards the poles for both mitochondrial and nuclear diversity, there were still a substantial number of diversity estimates at latitudes greater than 60 °N or S for both mitochondrial (39 observations) and nuclear (311 observations) diversity.

(b) Mitochondrial diversity

Globally, average mitochondrial genetic diversity was higher in the western Pacific Ocean and lower along North American and European coastlines (Fig. 2A & B, S4A & B). For both $H_d$and π, diversity peaked at low-to-mid latitudes and declined towards the poles, particularly in the Northern hemisphere (Fig. 2A & B, S7A & B). Diversity was also consistently higher in the Coral Triangle and elsewhere in the western Indo-Pacific (Fig. 3A & B). For mitochondrial genetic diversity (either H_d or π), we found that all latitude and longitude models performed better than the baseline (null) model (Table 1). Latitude, absolute latitude, and longitude were significant predictors of mitochondrial genetic diversity (Table 1, Fig. S8). As expected, H_d increased consistently with the length of the locus in base pairs (Fig. S9) and decreased towards species range edges (although π did not) (Fig. S10).

Both environmental drivers were significantly correlated with mitochondrial genetic diversity (H_d and π) and performed better than the null model (Table 2). Sea surface temperature was positively related with mitochondrial diversity (Fig. 4A & B), while chlorophyll-a concentration followed a quadratic relationship with diversity highest at mid-to-upper chlorophyll-a concentrations (5-10 mg/m^3) (Fig. 4D & E).

When looking at how these global patterns differed across a subset of families represented in our dataset, we found substantial variation. While the majority of the 10 families followed the same overarching patterns (e.g. reduced mitochondrial genetic diversity at higher latitudes, increased diversity at elevated SST, and a quadratic relationship with chlorophyll-a concentration), several did not (Fig. S11-S13). Gadidae (cods) and Sebastidae (rockfishes) showed elevated mitochondrial diversity at higher latitudes and lower SST for bothH_d and π, while the relationships between latitude and SST varied in Carcharhinidae (requiem sharks), Engraulidae (anchovies), and Rajidae (skates) by marker type.

(c) Nuclear diversity

In contrast to the mitochondrial results, there was no evidence for strong latitudinal or longitudinal diversity gradients in the nuclear dataset. Nuclear genetic diversity declined only weakly towards the poles and did not follow strong longitudinal patterns (Fig. 2C & 3C, Fig. S7). The null model performed better than any models that incorporated latitude or longitude predictors, and neither latitude nor longitude was a significant term in any of the models (Table 3). However, diversity was consistently lower for loci amplified with primers originally developed in another species (Fig. S14) and showed a negative, albeit non-significant, relationship with range position (Table 3, Fig. S10).

Nuclear diversity was also significantly associated with chlorophyll-a concentration; all models with chlorophyll-a as a predictor performed better than the null, and the model with only mean chlorophyll-a concentration performed best overall (Table 2). Similarly to the mitochondrial patterns, nuclear genetic diversity peaked at mid-to-upper chlorophyll-a concentrations (5-10 mg/m^3) (Fig. 4F). Mean SST was not significantly related with nuclear genetic diversity (Table 2, Fig 4C).

As with mitochondrial genetic diversity, global patterns in nuclear genetic diversity also appeared to vary across families, although to a much diminished degree (Fig. S11-13).

Discussion

Identifying global patterns in genetic diversity is a fundamental goal in ecology and evolution. Since genetic diversity is a proxy for adaptive potential and the raw material for speciation events, determining its spatial distribution can help us better understand which species are most vulnerable to anthropogenic change and help explain global patterns in species diversity. Here, we outlined and tested three distinct macroecological drivers of intraspecific genetic diversity, identified global patterns, and assessed the congruence of these relationships across the genome. Overall, we found that nuclear genetic diversity was most strongly correlated with chlorophyll-a concentration, a proxy for primary productivity and resource availability, while mitochondrial diversity was tightly associated with chlorophyll-a concentration, sea surface temperature, latitude, and longitude. Taken together, these results provide support for our original hypotheses to varying degrees. The quadratic relationship between chlorophyll-a concentration and genetic diversity across the genome provides compelling evidence for the Productivity-Diversity Hypothesis and suggests that regions of higher productivity facilitate larger population sizes, and in turn, higher levels of genetic variation. However, our results suggest a tipping point may exist in this relationship, after which larger carrying capacities may result in reduced population sizes and declining genetic diversity (Lawrence & Fraser, 2020). Furthermore, temperature was positively correlated with mitochondrial genetic diversity, lending support to the Kinetic Energy Hypothesis and the relationship between temperature, metabolism, and mutation rates. The lack of a significant correlation with nuclear diversity further affirmed this theory, as oxidative damage is not expected to impact nuclear DNA and increase nuclear mutation rates in the same manner (Hoffman et al. 2004).

Interestingly, the Founder Effect Hypothesis was the only one of our three initial hypotheses the three that we did not find full support for, although the observed decline in mitochondrial genetic diversity towards the poles is in line with the hypothesis' predictions. This decline was particularly pronounced near the Arctic, congruent with the outsized impact of glacial expansion on Northern hemisphere species, relative to their Southern Ocean counterparts (Fraser et al. 2012). Furthermore, the smaller N_e of mitochondrial DNA makes it more sensitive to LGM-induced bottlenecks (Birkey et al. 1989); strengthening any LGM signal in mitochondrial genetic diversity. Alternatively, the high levels of dispersal and admixture often observed in marine systems, along with high N_e s, may explain why a poleward decline was not observed in nuclear diversity, as elevated dispersal across the species range may help transport genetic diversity from the center to the poleward edge and replenish depleted gene pools. In fact, many temperate marine species harbor consistent levels of genetic diversity across their species range (Almada et al. 2012; Francisco et al. 2014; Martínez et al. 2015). Furthermore, microrefugia during the LGM that are uncoupled from historical climatic gradients may provide "re-seeding" opportunities for formerly glaciated regions and help buffer northern populations from extirpation, similar to previously documented patterns in the Antarctic (Suggitt et al. 2018). Given that some of these past refugia are close to modern northern range limits, expansion waves

out of these locations would have been less susceptible to diversity loss from bottlenecks or serial founder events (Bringloe et al. 2020; Maggs et al. 2008).

Previous studies have also found latitudinal gradients in mitochondrial genetic diversity, including Manel et al. (2020), another prominent macrogenetic study that analyzed global patterns in marine fish genetic diversity. However, the methods and statistical analyses frequently employed by macrogenetic studies have come under recent criticism (Gratton et al. 2017, Paz-Vinas et al. 2021). Most earlier macrogenetic studies fall into the category of Class III (Leigh et al. 2021) - pooling samples and sequences into predefined grid cells or latitudinal bands, calculating diversity at the species level, then averaging all species estimates together (Manel et al. 2020; Miraldo et al. 2016; Theodoridis et al. 2020). While informative, studies of this design often fail to account for genetic variation within species (such as from the range center to edge), for the relative frequency of individual haplotypes within each population, for study-specific methodological choices, or for the unbalanced sampling of species across grid cells (Gratton et al. 2017; Paz-Vinas et al. 2021). As population size is the mediating factor in many hypotheses aimed at explaining global patterns of genetic diversity, including those assessed here, such distinctions are important. Genetic diversity may follow different spatial patterns at different scales, given that environmental gradients, ecosystem processes, and biogeography collectively influence how population-level genetic diversity is shaped into community-wide patterns (De Kort et al. 2021). Here, we conducted a Class II macrogenetic study, which enabled us to incorporate metadata from the original populations, including sample sizes and the demarcation of local populations (Leigh et al. 2021). This approach enabled us to better account for issues of within-species geographic variation and relative haplotype abundance.

Despite these differing techniques, our findings also show that mitochondrial diversity follows clear latitudinal and longitudinal gradients - peaking at lower latitudes and in the Indo-Pacific - and reaffirm patterns previously established in Manel et al. (2020). Interestingly, the Coral Triangle has been designated as the center of species biodiversity, and our models suggest it could play a similar role for genetic diversity, especially within the mitochondria. These results are unsurprising, as several of the predictors we found to be strongly associated with mitochondrial diversity (e.g. sea surface temperature) have also been linked with higher species richness (Tittensor et al. 2010). Furthermore, heightened habitat availability and coastline length have been suggested as specific drivers of species richness in the Coral Triangle and could also increase genetic diversity through their positive influence on population size (Sanciangco et al. 2013). However, our models suggest that other regions in the Indo-Pacific show elevated mitochondrial genetic diversity as well, including the coastline of the Indian subcontinent and Sri Lanka, suggesting other macroecological factors may also play an important role in creating and maintaining genetic diversity.

Importantly, compared to mitochondrial diversity, nuclear genetic diversity did not follow clear geographic gradients across either latitude or longitude. These results are similar to previous studies that saw no strong latitudinal patterns in the nuclear diversity of mammals (Schmidt et al. 2022), freshwater fish (Lawrence et al. 2023), plants (De Kort et al. 2021), or habitat-forming species (Figuerola-Ferrando et al. 2023). As nuclear diversity is more tightly coupled with population size than is mitochondrial diversity (Bazin et al. 2006), recent demographic processes or changes in population size may disrupt any pre-existing geographic patterns and result in no clear latitudinal gradients in diversity. When compared to the spatial gradients in mitochondrial genetic diversity, the inconsistency in global patterns across the genome reinforces the narrative that mitochondrial and nuclear DNA are distinct entities that are separately impacted by divergent evolutionary forces, like drift (via population size) and mutation rates (via kinetic energy). While useful in many circumstances, mitochondrial DNA should be employed with care, and not as a broad and convenient proxy for nuclear markers. This distinction is important because fish mitochondrial genomes are approximately 16 to 17 kb, while nuclear genomes range in size from 300 Mb to 4.5 Gb (Fan et al. 2020; Satoh et al. 2016), which means that more than 99.99% of the genome is nuclear. Thus, the nuclear genome contains the majority of standing genomic variation important for adaptation to changing conditions and for the speciation process.

Additionally, species-level variation often reduces our power to detect general macro-scale relationships, and

almost certainly contributed to the lower R^2 values reported here. Unsurprisingly, we found substantial variation in family-specific global gradients of genetic diversity for 10 families that represented a wide swath of life history traits. While most of the families followed the general patterns (at least for mitochondrial diversity) established in the main models, several instead showed increasing genetic diversity at higher latitudes and lower SST. Notably, most of these families (including Gadidae and Sebastidae) are traditionally found in colder, more temperate environments that also often have higher levels of primary productivity. If species at these latitudes can support consistently large populations due to higher resource availability, the relationship with other important ecological variables, like temperature, might be muted. This may be the case in our models, as all 10 families displayed either a positive or quadratic relationship with chlorophyll-a concentration. Nevertheless, this variation across families is an important reminder that global patterns are frequently complex, multifaceted, and often the result of many ecological and species-specific factors.

Generally speaking, macroecological drivers are likely to act in concert, not isolation, to shape global patterns. Variation in population size, and subsequently the strength of genetic drift, likely creates a baseline distribution of genetic diversity, upon which other evolutionary forces interact to create more complex patterns. Both mitochondrial and nuclear genetic diversity peaked in communities and ecosystems with higher resource availability, as represented by primary productivity. In addition, most models suggested genetic diversity was elevated closer to the range core, consistent with the central-marginal hypothesis that suggests population abundance—and subsequently, genetic diversity—is highest towards the range core where environmental conditions tend to be optimal (Eckert et al. 2008). Layered upon these findings, we found evidence that the higher mitochondrial substitution rates at lower latitudes may serve to replenish and accumulate diversity at lower latitudes, manifesting in a traditional latitudinal gradient for mitochondrial diversity that is highest near the tropics. As nuclear substitution rates are not as clearly elevated at higher temperatures (Hoffman et al. 2004), similar latitudinal patterns in nuclear genetic diversity were not apparent. Life history traits, anthropogenic change, phylogenetic relationships, and demographic history are also well-known determinants of genetic diversity, and it is likely these processes influenced our results. For instance, historically, tropical environments tend to be more stable, which can enable diversity at both the species and genetic level to accumulate over time and contribute to the latitudinal diversity gradients observed here (Rosenzweig 1995).

Range size is also commonly invoked as a driver of latitudinal patterns of genetic diversity (French et al. 2022; Lawrence & Fraser 2020), especially when genetic diversity increases towards higher latitudes. According to Rapoport's rule, range size grows with latitude (Rapoport 1982), and may be coupled with a rise in genetic diversity because larger range sizes can support more and larger populations, and even low levels of gene flow among these demes can increase local genetic diversity (Waples 2010). However, as access to this range-wide genetic diversity is mediated by dispersal, there is no guarantee that a particular population will acquire novel alleles from elsewhere in the range. While most oceanic taxa likely have high enough rates of gene flow to facilitate this level of genetic exchange (Palumbi 1992), studies have found that marine ranges can be much more structured than previously thought (Pringle & Wares 2011; Selkoe et al. 2016). Future work explicitly testing the roles of range size and gene flow in determining general patterns of genetic diversity would help provide further clarity.

Investigating other DNA markers may also help disentangle the relative importance of various environmental drivers. In addition to the issues with mitochondria that we previously discussed, the high mutation rate of microsatellites, as well as ascertainment bias for highly polymorphic loci during marker generation, can create extraneous statistical noise and may be one reason why it was difficult to identify clear spatial patterns in nuclear diversity. Furthermore, the limited range of heterozygosity (0-1) can also impose inferential challenges and restrict the scope of observable patterns. These issues aside, microsatellites remain one of the most widely available measures of neutral nuclear genetic diversity and are positively correlated with genome-wide diversity (Mittel et al. 2015). Moreover, expected heterozygosity is a robust diversity metric unlikely to be biased by either sampling effort (Toro et al. 2009) or inbreeding because it is calculated from allele frequencies (Ritland 1996). While nuclear DNA sequence diversity (e.g. SNPs, haplotypes) provides a promising next step for future macrogenetic analyses, standardizing such data across studies remains a substantial bioinformatic

challenge.

Overall, our results reveal disparate global gradients in mitochondrial and nuclear genetic diversity. While mitochondrial diversity peaks along the Equator and is positively associated with temperature, mirroring complementary patterns in marine species, nuclear genetic diversity shows no strong geographic patterns. Such a lack of clear gradients in nuclear diversity may be due in part to either evolutionary forces (e.g. contemporary demographic processes disrupting historical patterns, gene flow more evenly distributing alleles across species ranges, or latitudinally consistent mutation rates), analytical ones (e.g. the reduced statistical power of microsatellites), or a combination of the two. However, despite these differences, diversity across the genome was strongly correlated with chlorophyll-a concentrations and was elevated in regions of higher primary productivity and resource availability that are able to support larger population densities. Taken together, these findings enable a better understanding of the degree to which mutation rates (via elevated temperatures) and drift (via population size) work collectively to establish large-scale gradients of genetic diversity, providing a more comprehensive view of how forces interacting across the genome scale up to provide the starting material for species and ultimately community diversity.

Acknowledgments

We thank Diana Li, Kieryn Graham, Stan Piotrowki, Ash Battacharjee, Sarah Picon, Chloe Lewis, Michael Weiss, and Marial Malabag for help gathering, filtering, and recording data used in this study. We thank Boris Worm and Chloé Schmidt for discussions and Peter Smouse for manuscript revisions during the research. This work was supported by a Rutgers Institute of Earth, Ocean, and Atmospheric Sciences Fellowship, a Rutgers School of Environmental and Biological Sciences (SEBS) Excellence Fellowship, the Coronavirus Response and Relief Supplemental Appropriations Act, 2021 (CRRSAA)/Higher Education Emergency Relief Fund II (HEERF II), a Smith Conservation Research Fellowship, an Alfred P. Sloan Ocean Sciences Fellowship (BR2014-044), and US National Science Foundation grants #OCE-1426891, #DEB-2129351, and #OISE-1743711.

References

Allcock, A.L. & Strugnell, J.M. (2012). Southern Ocean diversity: new paradigms from molecular ecology. *Trends in Ecology & Evolution* , 27, 520-528.

Allen, A.P., Brown, J.H. & Gillooly, J.F. (2002). Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. *Science* , 297, 1545-1548.

Almada, V.C., Almada, F., Francisco, S.M., Castilho, R. & Robalo, J.I. (2012). Unexpected high genetic diversity at the extreme northern geographic limit of *Taurulus bubalis* (Euphrasen, 1786). *PLoS ONE*, 7, e44404

Assis, J., Tyberghein, L., Bosch, S., Verbruggen, H., Serrão, E.A. & De Clerck, O. (2018). Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling. *Global Ecology and Biogeography*, 27, 277-284.

Ballard, J.W.O. & Whitlock, M.C. (2003). The incomplete natural history of mitochondria. *Molecular Ecology* , 13, 729-744.

Barbará, T., Palma-Silva, C., Paggi, C.M., Bered, F., Fay, M.F. & Lexer, C. (2007). Cross-species transfer of nuclear microsatellite markers: potential and limitations. *Molecular Ecology*, 16, 3759-3767.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv: 1406.5823* .

Bazin, E., Glémin, S. & Galtier, N. (2006). Population size does not influence mitochondrial genetic

diversity in animals. *Science* , 312, 570-572.

Bergmann, C. (1847). Ueber die verhältnisse der wärmeökonomie der thiere zu ihrer grösse. *Gottinger studien* , 3, 595-708.

Birky, C.W., Fuerst, P. & Maruyama, T. (1989). Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics*, 121, 613-627.

Bosch, S. & Fernandez, S. (2021). sdmpredictors: species distribution modelling predictor datasets. R package version 0.2.10. See*https://CRAN.R-project.org/package=sdmpredictors*

Bringloe, T.T., Verbruggen, H. & Saunders, G.W. (2020). Unique biodiversity in Arctic marine forests is shaped by diverse recolonization pathways and far northern glacial refugia. *Proceedings of the National Academy of Sciences*, 117, 22590-22596.

Brooks, M.E., Kristensen, K., van Benthem, K.J., Magnusson, A., Berg, C.W., Nielsen, A., Skaug, H.J., Mächler, M. & Bolker B.M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, 9, 378-400.

Burnham, K.P. & Anderson, D.R. (2002). Model selection and multimodal inference: a practical information-theoretic approach. Springer New York, New York, NY.

Canty, A. & Ripley, B.D. (2022) *boot: Bootstrap R (S-Plus) Functions* . R package version 1.3-28.1.

Charlesworth, B. (2009) Effective population size and patterns of molecular evolution and variation. *Nature Review Genetics* , 10, 195-205.

Chaudhary, C., Saeedi, H. & Costello, M.J. (2016). Bimodality of latitudinal gradients in marine species richness. *Trends in Ecology & Evolution* , 31, 670-676.

Clark, R.D., Aardema, M.L., Andolfatto, P., Barber, P.H., Hattori, A., Hoey, J.A., *et al* . (2021). Genomic signatures of spatially divergent selection at clownfish range margins.*Proceedings of the Royal Society B: Biological Sciences*, 288, 20210407.

Core Team R. 2023: *R a language and environment for statistical computing* . Vienna, Austria: R Foundation for Statistical Computing.

Davies, R.G., Orme, C.D.L., Storch, D., Olson, V.A., Thomas, G.H., Ross, S.G., *et al* . (2007). Topography, energy and the global distribution of bird species richness.*Proceedings of the Royal Society B: Biological Sciences*, 274, 1189-1197.

De Kort, H., Prunier, J.G., Ducatez, S., Honnay, O., Baguette, M., Stevens, V.M., *et al* . (2021). Life history, climate and biogeography interactively affect worldwide genetic diversity of plant and animal populations. *Nature Communications*, 12, 516.

Eckert, C.G., Samis, K.E. & Lougheed, S.C. (2008). Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Molecular Ecology* , **17** , 1170-1188.

Evans, K.L., Warren, P.H. & Gaston, K.J. (2005). Species-energy relationships at the macroecological scale: a review of the mechanisms. *Biological Reviews* , 80, 1-25.

Fan, G., Song, Y., Yang, L., Huang, X., Zhang, S., Zhang, M., *et al* . (2020) Initial data release and announcement of the 10,000 Fish Genomes Project (Fish10K). *GigaScience* , **9** , giaa080.

Fernández-Torres, F., Martínez, P.A. & Olalla-Tárraga, M.Á. (2018). Shallow water ray-finned marine fishes follow Bergmann's rule. *Basic and Applied Ecology* , 33, 99-110.

Figuerola-Ferrando, L., Barreiro, A., Montero-Serra, I., Pagès-Escolà, M., Garrabou, J., Linares, C. & Ledoux, J.-B. (2023). Global patterns and drivers of genetic diversity among marine habitat-forming species. *Global Ecology and Biogeography*, 32, 1218-1229.

Francisco, S.M., Robalo, J.I., Levy, A. & Almada, V.C. (2014). In search of phylogeographic patterns in the Northeastern Atlantic and adjacent seas. *Evolutionary Biology: Genome Evolution, Speciation, Coevolution and Origin of Life*, 323-338.

Fraser, C.I., Nikula, R., Ruzzante, D.E. & Waters, J.M. (2012). Polewarde bound: biological impacts of Southern Hemisphere glaciation. *Trends in Ecology & Evolution* , 27, 462-471.

French, C.M., Bertola, L.D., Carnaval, A.C., Economo, E.P., Kass, J.M., Lohman, D.J., *et al* . (2022). Global determinants of insect genetic diversity. *bioRxiv* .

Galtier, N., Nabholz, B., Glémin, S. & Hurst, G.D.D. (2009). Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology* , 18, 4541-4550.

Gillooly, J.F., Allen, A.P., West, G.B. & Brown, J.H. (2004). The rate of DNA evolution: effects of body size and temperature on the molecular clock. *Proceedings of the National Academy of Sciences*, 102, 140-145.

Gratton, P., Marta, S., Bocksberger, G., Winter, M., Keil, P., Trucchi, E. *et al* . (2017). Which latitudinal gradients for genetic diversity? *Trends in Ecology & Evolution* , 32, 724-726.

Hartig, F. (2021). DHARMa: Residual diagnostics for hierarchical (multi-level/mixed) regression models. R package version 0.4.3. See*https://CRAN.R-project.org/package=DHARMa*.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature* , 405, 907-913.

Hoffmann, S., Spitkovsky, D., Radicella, J.P., Epe, B. & Wiesner, R.J. (2004). Reactive oxygen species derived from the mitochondrial respiratory chain are not responsible for the basal levels of oxidative base modifications observed in nuclear DNA of mammalian cells. *Free Radical Biology and Medicine*, 36, 765-773.

James, J. & Eyre-Walker, A. (2020). Mitochondrial DNA sequence diversity in mammals: a correlation between the effective and census population sizes. *Genome Biology and Evolution*, **12**, 2441-2449.

Jenkins, T.L., Castilho, R. & Stevens, J.R. (2018). Meta-analysis of northeast Atlantic marine taxa shows contrasting phylogeographic patterns following post-LGM expansions. *PeerJ* , 6, e5684.

Kaschner, K., Kesner-Reyes, K., Garilao, C., Rius-Barile, J., Rees, T. & Froese, R. (2019). *AquaMaps: Predicted range maps for aquatic species, version 08/2016c* .

Kubinec, R. (2022). Ordered beta regression: a parsimonious, well-fitting model for continuous data with lower and upper bounds. *Political Analysis* , 31, 519-536.

Lawrence, E.R. & Fraser, D.J. (2020). Latitudinal biodiversity gradients at three levels: Linking species

richness, population richness and genetic diversity. *Global Ecology and Biogeography*, 29, 770-778.

Lawrence, E.R., Pedersen, E.J. & Fraser, D.J. (2023). Macrogenetics reveals multifaceted influences of environmental variation on vertebrate population genetic diversity across the Americas. *Molecular Ecology*, 32, 4557-4569.

Leigh, D.M., van Rees, C.B., Millette, K.L., Breed, M.F., Schmidt, C., Bertola, L.D., *et al* . (2021). Opportunities and challenges of macrogenetic studies. *Nature Review Genetics* , 22, 791-807.

Lessa, E.P., D'Elía, G., Pardiñas, U.F.J. (2010). Genetic footprints of late Quaternary climate change in the diversity of Patagonian-Fueguian rodents. *Molecular Ecology* , 19, 3031-3037.

Li, Y., Wang, S., Cheng, C., Zhang, J., Wang, S., Hou, X., *et al* . (2021) Latitudinal gradients in genetic diversity and natural selection at a highly adaptive gene in terrestrial mammals. *Ecography*, 44, 206-218.

Lüdecke, D. (2021). sjPlot: data visualization for statistics in social science. R package version 2.8.10. See*https://CRAN.R-project.org/package=sjPlot*.

Lüdecke, D., Ben-Shachar, M.S., Patil, I., Waggoner, P. & Makowski, D. (2021). performance: an R package for assessment, comparison and testing of statistical models. *The Journal of Open Source Software*, 60, 3139.

Maggs, C.A., Castilho, R., Foltz, D., Henzler, C., Jolly, M.T., Kelly, J., *et al* . (2008) Evaluating signatures of glacial refugia for North Atlantic benthic marine taxa. *Ecology* , 89, S108-S122.

Manel, S., Guerin P.-E., Mouillot, D., Blanchet, S., Velez, L., Albouy, C., *et al* . (2020). Global determinants of freshwater and marine fish genetic diversity. *Nature Communications* , 11, 692.

Martin, P.R. & McKay, J.K. (2004). Latitudinal variation in genetic divergence of populations and the potential for future speciation. *Evolution* , 58, 938-945.

Martin, A.P. & Palumbi, S.R. (1993). Body size, metabolic rate, generation time, and the molecular clock. *Proceedings of the National Academy of Sciences* , 90, 4087-4091.

Martínez, L., Freire, R., Arias-Pérez, A., Méndez, J. & Insua, A. (2015). Patterns of genetic variation across the distribution range of the cockle *Cerastoderma edule*inferred from microsatellites and mitochondrial DNA. *Marine Biology*, 162, 1393-1406.

Mattingsdal, M., Jorde, P.E., Knutsen, H., Jentoft, S., Stenseth, N.C., Sodeland, M., *et al* . (2020). Demographic history has shaped the strongly differentiated corkwing wrasse populations in Northern Europe. *Molecular Ecology*, 29, 160-171.

Messmer, V., Jones, G.P., Munday, P.L. & Planes, S. (2012). Concordance between genetic and species diversity in coral reef fishes across the Pacific Ocean biodiversity gradient. *Evolution*, 66, 3902-3917.

Miraldo, A., Li, S., Borregaard, M.K., Flórez-Rodríguez, A., Gopalakrishnan, S., Rizvanovic, M., *et al* . (2016). An Anthropocene map of genetic diversity. *Science* , 353, 1532-1535.

Mittell, E.A., Nakagawa, S. & Hadfield, J.D. (2015). Are molecular markers useful predictors of adaptive potential? *Ecology Letters* , 18, 772-778.

Nabholz, B., Glémin, S. & Galtier, N. (2009). The erratic mitochondrial clock: variations of mutation rate,

not population size, affect mtDNA diversity across birds and mammals.*BMC Evolutionary Biology*, 9, 54.

Nakagawa, S. & Schielzeth, H. (2013) A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution* , 4, 133-142.

Palumbi, S.R. (1992). Marine speciation on a small planet. *Trends in Ecology & Evolution* , 7, 114-118.

Paz-Vinas, I., Jensen, E.L., Bertola, L.D., Breed, M.F., Hand, B.K., Hunter, M.E., *et al* . (2021). Macrogenetics must not ignore limitations of genetic markers and scale.*Ecology Letters*, 24, 1282-1284.

Pinsky, M.L. & Palumbi, S.R. (2014). Meta-analysis reveals lower genetic diversity in overfished populations. *Molecular Ecology* , 23, 29-39.

Pringle, J.M. & Wares, J.P. (2007). Going against the flow: maintenance of alongshore variation in allele frequency in a coastal ocean. *Marine Ecology Progress Series* , 335, 69-84.

Rapoport, E.H. (1982). Areography: geographical strategies of species. *Trans. B. Drausal* . Vol. 1. Pergamon, New York.

Reed, D.H. & Frankham, R. (2003). Correlation between fitness and genetic diversity. *Conservation Biology* , 17, 230-237.

Ritland, A. (1996). Estimators for pairwise relatedness and individual inbreeding coefficients. *Genetics Research* , 67, 175-185.

Robalo, J.I., Francisco, S.M., Vendrell, C., Lima, C.S., Pereira, A., Brunner, B.P., *et al* . (2020). Against all odds: a tale of marine range expansion with maintenance of extremely high genetic diversity. *Scientific Reports*, 10, 12707.

Rolland, J., Condamine, F.L., Jiguet, F. & Morlon, H. (2014). Faster speciation and reduced extinction in the tropics contribute to the mammalian latitudinal diversity gradient.*PLoS Biology*, 12, e1001775.

Rosenzweig M. L. (1995). *Species Diversity in Space and Time.*Cambridge University Press, Cambridge.

Sanciangco, J.C., Carpenter, K.E., Etnoyer, P. J. & Moretzsohn F. (2013). Habitat availability and heterogeneity and the Indo-Pacific warm pool as predictors of marine species richness in the tropical Indo-Pacific. *PLoS ONE*, 8, e56245.

Satoh, T.P., Miya, M., Mabuchi, K. & Nishida, M. (2016). Structure and variation of the mitochondrial genome of fishes. *BMC Genomics* , 17, 719.

Schmidt, C., Dray, S., Garroway, C.J. (2022). Genetic and species-level biodiversity patterns are linked by demography and ecological opportunity. *Evolution* , 76, 86-100.

Selkoe, K.A., D'Aloia, C.C., Crandall, E.D., Iacchei, M., Liggins, L., Puritz, J.B., *et al* . (2016). A decade of seascape genetics: contributions to basic and applied marine connectivity. *Marine Ecology Progress Series*, 554, 1-19.

Steele, J.H., Brink, K.H. & Scott, B.E. (2019). Comparison of marine and terrestrial ecosystems: suggestions of an evolutionary perspective influenced by environmental variation. *ICES Journal of Marine Science*, 76, 50-59.

Strathmann, R.R. (1990). Why do life histories evolve differently in the sea? *American Zoologist* , 30,

197-207.

Suggitt, A.J., Wilson, R.J., Isaac, N.J.B., Beale, C.M., Auffret, A.G., August, T., *et al* . (2018). Extinction risk from climate change is reduced by microclimatic buffering.*Nature Climate Change*, 8, 713-717.

Theodoridis, S., Fordham, D.A., Brown, S.C., Li, S., Rahbek, C. & Nogues-Bravo, D. (2020). Evolutionary history and past climate change shape the distribution of genetic diversity in terrestrial mammals. *Nature Communications*, 11, 2557.

Thomas, J.A., Welch, J.J., Lanfear, R. & Bromham, L. (2010). A generation time effect on the rate of molecular evolution in invertebrates. *Molecular Biology and Evolution* , 27, 1173-1180.

Tittensor, D.P., Mora, C., Jetz, W., Lotze, H.K., Ricard, D., *et al* . (2010). Global patterns and predictors of marine biodiversity across taxa. *Nature* , 466, 1098-1101.

Toro, M.A., Fernández, J. & Caballero, A. (2009). Molecular characterization of breeds and its use in conservation. *Livestock Science* , 120, 174-195.

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F. & De Clerck, O. (2012). BioORACLE: a global environmental dataset for marine species distribution modelling. *Global Ecology and Biogeography*, 21, 272-281.

Vellend, M. & Geber, M.A. (2005). Connections between species diversity and genetic diversity. *Ecology Letters* , 8, 767-781.

Waples, R.S. (2010). Spatial-temporal stratifications in natural populations and how they affect understanding and estimation of effective population size.*Molecular Ecology Resources*, 10, 785-796.

Wright, D.H. (1983). Species-energy theory: an extension of species-area theory. *Oikos* , 41, 496-506.

Wright, S.D., Ross, H.A., Keeling, D.J., McBride, P. & Gillman, L.N. (2011). Thermal energy and the rate of genetic evolution in marine fishes. *Evolutionary Ecology* , 25, 525-530.

Figure Legends

Figure 1. Map of observation locations for mitochondrial (A: π, B: H_d) and nuclear (C: H_e) genetic diversity. Populations were binned into 500 km x 500 km equal-area grid cells and the mean species-wide genetic diversity within each cell was plotted on a Mollweide projection. Rug plots on the x- and y-axes illustrate the latitudinal and longitudinal sampling locations.

Figure 2. Relationship between absolute latitude and genetic diversity (A: mitochondrial π· B: mitochondrialH_d ; C: nuclear microsatelliteH_e). Gray line represents the predicted relationship based on the mixed effects model with shaded 95% confidence intervals. Blue-gray violin plots show the distribution of genetic diversity binned every 10°, with the dark points representing the medians in every 10° band.

Figure 3. Relationship between longitude and genetic diversity (A: mitochondrial π· B: mitochondrial H_d ; C: nuclear microsatellite H_e). Gray line represents the predicted relationship based on the mixed effects model with shaded 95% confidence intervals. Blue circles represent median diversity binned every 10° with median average deviation (MAD) error bars. Green highlighted region represents the Coral Triangle (longitudes 95 - 165).

Figure 4. A-C: Relationship between mean sea surface temperature (SST) (A, B, C) or mean chlorophyll-a concentration (D, E, F) and genetic diversity (A, D: mitochondrial π· B, E: mitochondrialH_d ; C, F: nuclear microsatelliteH_e). Black line represents the predicted relationship based on the mixed effects model with

shaded 95% confidence intervals. Rug plots on the x-axis illustrate the SST or chlorophyll-a sampling extent. Mean chlorophyll-a concentration is plotted on a common logarithm scale.

Table 1. Mitochondrial DNA (π and H_d) model results for latitude and longitude. Standardized model coefficients are reported, along with ΔAIC compared to the null model (model AIC - null AIC) and R^2 values (R^2_C = conditional R^2, considers all fixed and random effects; R^2_M = marginal R^2, considers only fixed effects). Model coefficients represent normal slopes for mtDNA π and log odds for mtDNA H_d . For the null models, AIC is also reported in parentheses. The top model (with the lowest AIC) is bolded, while the null model is highlighted in gray. For latitude, latitude and latitude^2 were included as predictors in the same model(s). For longitude, the b-spline basis function coefficients are reported (1-3), each on a different line.

Model	bp	Range Position	Abslat	Lat [Lat^2]	Lon	ΔAIC	R^2_C [R^2_M]
π							
Null		-0.001				0 (645.0)	0.844 [0.000]
Absolute Latitude		0.012	-0.047**			-5.3	0.846 [0.006]
Latitude		0.002		-0.039* [-0.009]		-0.9	0.845 [0.003]
Longitude		0.002			0.147 0.306*** 0.086	-17.4	0.846 [0.007]
Absolute Latitude & Longitude		**0.012**	**-0.037***		**0.156** **0.283**** **0.086**	**-20.2**	**0.847 [0.010]**
Latitude & Longitude		0.004		-0.028 [-0.007]	0.142 0.301*** 0.073	-16.1	0.846 [0.009]
H_d							
Null	0.379***	-0.062*				0 (-1748.6)	0.217 [0.012]
Absolute Latitude	0.367***	-0.028	-0.110*			-2.4	0.217 [0.013]
Latitude	0.376***	-0.046		-0.185*** [-0.019]		-9.8	0.218 [0.015]
Longitude	0.399***	-0.054*			0.043 1.211*** 0.139	-25.6	0.214 [0.015]
Absolute Latitude & Longitude	0.390***	-0.032	-0.074		0.026 1.168*** 0.140	-25.5	0.214 [0.016]
Latitude & Longitude	**0.395*** **	**-0.042**		**-0.151** [-0.014]**	**0.085** **1.199***** **0.047**	**-30.3**	**0.214 [0.016]**

p-value : * < 0.05; ** < 0.01; *** < 0.001

Table 2. Mitochondrial DNA (π and H_d) and nuclear (microsatellite H_e) model results for macroecological drivers (mean sea surface temperature (SST) and mean chlorophyll-a concentration). Model coefficients are reported, along with ΔAIC compared to the null model (model AIC - null AIC) and R^2 values (R^2_C =

conditional R^2, considers all fixed and random effects; R^2_M = marginal R^2, considers only fixed effects). Model coefficients represent normal slopes for mtDNA π and log odds for mtDNAH_d and nucDNA H_e . For the null models, AIC is also reported in parentheses. All model coefficients are standardized except for mean chlorophyll-a, which were log-transformed. Furthermore, for chlorophyll-a models, chlorophyll-a and chlorophyll-a^2 were included as predictors in the same model(s). The top model (with the lowest AIC) is bolded, while the null model is highlighted in gray.

Model	CrossSpp	bp	Range position	SST mean	Chlor mean [Chlor mean^2]	ΔAIC	R^2_C [R^2_M]
π (mtDNA)							
Null			-0.001			0 (645.0)	0.844 [0.000]
SST			0.006	0.054**		-7.5	0.846 [0.008]
Chlorophyll-a			-0.002		0.027 [-0.053]**	-9.5	0.847 [0.003]
SST & Chlorophyll-a			**0.007**	**0.067*** **	**0.040* [-0.056]*** **	**-21.6**	**0.849 [0.011]**
H_d (mtDNA)							
Null		0.379***	-0.062*			0 (-1748.6)	0.217 [0.012]
SST		0.355**	-0.034	0.196***		-12.1	0.219 [0.016]
Chlorophyll-a		0.379***	-0.061*		-0.023 [-0.159]**	-5.9	0.216 [0.012]
SST & Chlorophyll-a		**0.351*** **	**-0.032**	**-0.214*** **	**-0.021 [-0.171]*** **	**-19.9**	**0.217 [0.016]**
H_e (nuclear)							
Null	-0.072***		-0.007			0 (-11524.8)	0.049 [0.001]
SST	-0.072***		-0.006	0.009		1.8	0.049 [0.001]
Chlorophyll-a	**-0.072*** **		**-0.007**		**0.023 -0.041****	**-3.1**	**0.049 [0.001]**
SST & Chlorophyll-a	-0.071		-0.005	0.014	0.024 [-0.041]**	-1.5	0.049 [0.001]

p-value: * < 0.05; ** < 0.01; *** < 0.001

Table 3. Nuclear microsatellite (H_e) model results. Standardized model coefficients are reported, along with ΔAIC compared to the null model (model AIC - null AIC) and R^2 values (R^2_C = conditional R^2, considers all fixed and random effects; R^2_M = marginal R^2, considers only fixed effects). Model coefficients represent log odds. For the null models, AIC is also reported in parentheses. The top model (with the lowest AIC) is bolded, while the null model is highlighted in gray. For latitude, latitude and latitude^2 were included

as predictors in the same model(s). For longitude, the b-spline basis function coefficients are reported (1-3), each on a different line.

Model	CrossSpp	Range Position	Abslat	Lat [Lat^2]	Lon	AIC	R^2_C [R^2_M]
Null	**-0.072*** **	**-0.007**				**0 (-11524.8)**	**0.049 [0.001]**
Absolute Latitude	-0.072***	-0.004	-0.017			1.1	0.049 [0.001]
Latitude	-0.072***	-0.006		-0.025 [-0.006]		2.9	0.048 [0.001]
Longitude	-0.072***	-0.007			0.114 0.150 0.092	0.9	0.049 [0.001]
Absolute Latitude & Longitude	-0.072***	-0.003	-0.018		0.110 0.155 0.089	2.0	0.049 [0.001]
Latitude & Longitude	-0.072***	-0.005		-0.028 [-0.008]	0.107 0.168 0.082	3.7	0.049 [0.001]

p-value: * < 0.05; ** < 0.01; *** < 0.001