Results

Bayesian linear regression recovers a reliably positive correlation between the amount of Z/W differentiation and d_N/d_S of mitochondrial genes (posterior mean = 0.48, credible interval (CI) = 0.15 - 0.80), although slightly less than estimated with ordinary least squares (OLS) regression (standardized regression coefficient or $\boldsymbol{\beta}$ = 0.035, p = .012) (**Figure 2A**). This implies that a one standard deviation change in % ZW differentiation is associated with a change of 0.48 standard deviations in d_N/d_S. To make sure that the observed variation in d_N/d_S is not due to differences in synonymous branch length rather than reflecting changes in N_e, we tested for a correlation between d_N/d_S and d_S and found no evidence of an association using Bayesian linear regression (posterior mean = -0.06, CI = -0.65 - 0.53), or with OLS ($\boldsymbol{\beta}$ = -0.01, p = 0.72). We found no correlation between mitochondrial genome size or our measure of selective efficiency, d_N/d_S, using (posterior mean = -0.06, CI = -0.49 - 0.36) or OLS ($\boldsymbol{\beta}$ = -0.07, p = 0.80; **Figure 2B**).

Looking at the relationship between % W chromosome differentiation and mitochondrial genome size, Bayesian linear regression does not recover any reliable correlation (posterior mean = 0.05, CI = -0.33 - 0.44). However, in this case, phylogenetic nonindependence due to shared ancestry creates a statistical issue that should be accounted for because any relationship between the two traits could result from shared phylogenetic history (Felsenstein 1985), so we used a Bayesian mixed model to include information about phylogenetic relationships. For comparison, but without the uncertainty in d_N/d_S, we also fit a maximum likelihood linear phylogenetic model with generalized least squares (PGLS). We found there is still no effect with Bayesian phylogenetic correction (posterior mean = 0.02, CI = -0.44 - 0.49), using PGLS ($\boldsymbol{\beta}$ = 0.04, p = 0.93), nor with OLS ($\boldsymbol{\beta}$ = 0.06, p = 0.84) **(Figure 2C)**.

Finally, we tested Lynch et al.'s (2006) alternative model for organelle genome size evolution, which states that species with higher mutation rates experience stronger selection against large, maladaptive genome sizes than those with lower mutation rates and found no support for a significant relationship using Bayesian phylogenetic regression (posterior mean = 0.12, CI = -0.25 - 0.48), PGLS ($\boldsymbol{\beta}$ = 0.15, p = 0.58), or with OLS ($\boldsymbol{\beta}$ = 0.10, p = 0.68) **(Figure 2D)**. If anything, the association is slightly positive, where species with higher mutation rates tend to have larger genomes. Furthermore, there is no evidence that the mutation rate and the amount of recombination suppression on the W together have an influence on genome size using multiple regression with a Bayesian phylogenetic mixed model (u posterior mean = 0.05, CI = -0.33 - 0.43; %W posterior mean = 0.02, CI = -0.45 - 0.50), PGLS ($u\,\boldsymbol{\beta}$ = 0.05, p = 0.87; %W $\boldsymbol{\beta}$ = 0.03, p = 0.95), or with OLS ($u\,\boldsymbol{\beta}$ = -0.002, p = 0.96; %W $\boldsymbol{\beta}$ = 0.07, p = 0.80). Therefore, this study suggests neither the strength of selection (u) nor its efficacy (N_e) significantly influences mitochondrial genome size in paleognathous birds.

Fig. 2. Correlations between the amount of genetic linkage, the rate of molecular evolution, mitochondrial genome size, and mutation rates. Black lines show the posterior mean, and the shaded regions show the 89% credible interval using Bayesian regression in 15 species of paleognathous birds. Red lines are the regression coefficients ($\boldsymbol{\beta}$) from ordinary least squares regression (OLS), and green lines are from phylogenetic least squares regression (PGLS). For d_N/d_S, blue points show estimates from HyPhy, whereas black points incorporate confidence intervals. The connecting vertical lines represent shrinkage in the posterior distribution resulting

from measurement uncertainty. All variables are standardized to have a mean of zero and a standard deviation of one, meaning a unit increase in a variable equals its standard deviation. (A) A robust association exists between the amount of genetic linkage and d_N/d_S. (B) There is no reliable correlation between the rate of molecular evolution and mitochondrial genome size. (C) There is no reliable correlation between the amount of genetic linkage and mitochondrial genome size. (D) The relationship between mitochondrial genome size and the per-generation mutation rate per nucleotide site (u) is insignificant.

Methods

We estimated the mitogenome-wide d_N/d_S ratio using HyPhy v. 2.5.36 (Pond and Muse 2005) with an MG94 model of codon evolution plus a GTR model of nucleotide substitution, including confidence intervals (https://github.com/veg/hyphy-analyses). A key component of HyPhy methods is that d_S can vary across branches, and we chose the MG94 model because it explicitly models both synonymous and nonsynonymous site variability. Additionally, the FitMG94 workflow uses a corrected empirical estimator (CF3x4) that provides improved estimates of several parameters by accounting for biases in nucleotide composition induced by stop codons (Goldman and Yang 1994). Generation times used to estimate the mutation rate were taken from Wang et al. (2021) unless otherwise noted in S1. Before proceeding, we logged all four variables, subtracted the mean, and divided them by the standard deviation to ease prior specification plus linear model fit and computation. We performed all Bayesian regressions using rstan (Stan Development Team 2023). For the phylogenetic mixed model, we used ape (Paradis et al. 2004) to calculate the distance matrix and a covariance matrix linearly related to the phylogenetic distance between the species (Brownian motion). Priors were checked via prior predictive

simulation and weakly regularized to penalize extreme parameter values: intercept ~ Normal (0, 0.2), slope ~ Normal(0,0.5), and standard deviation ~ Exponential(1). We accommodated the inferred measurement error in each observed HyPhy value (d_N/d_SOBS, i) by adding the additional parameter (d_N/d_STRUE, i): d_N/d_SOBS, i ~ Normal((d_N/d_STRUE, i), (d_N/d_SSD, i)). Maximum-likelihood PGLS regressions were implemented using nlme (Pinheiro et al. 2017). Since the mutation rate (*u*) predominately influences dS (Kimura 1983), we divided d_S by the generation time and normalized by the terminal branch length to approximate the mutation rate per generation per nucleotide site.