


Genomic signatures of convergent shifts to plunge-diving behavior in birds

Chad M. Eliason ^{1,7}, Lauren E. Mellenthin², Taylor Hains^{1,3,8}, Jenna M. McCullough⁴, Stacy Pirro⁵, Michael J. Andersen⁴ & Shannon J. Hackett ^{6,7}

Understanding the genetic basis of convergence at broad phylogenetic scales remains a key challenge in biology. Kingfishers (Aves: Alcedinidae) are a cosmopolitan avian radiation with diverse colors, diets, and feeding behaviors—including the archetypal plunge-dive into water. Given the sensory and locomotor challenges associated with air-water transitions, kingfishers offer a powerful opportunity to explore the effects of convergent behaviors on the evolution of genomes and phenotypes, as well as direct comparisons between continental and island lineages. Here, we use whole-genome sequencing of 30 diverse kingfisher species to identify the genomic signatures associated with convergent feeding behaviors. We show that species with smaller ranges (i.e., on islands) have experienced stronger demographic fluctuations than those on continents, and that these differences have influenced the dynamics of molecular evolution. Comparative genomic analyses reveal positive selection and genomic convergence in brain and dietary genes in plunge-divers. These findings enhance our understanding of the connections between genotype and phenotype in a diverse avian radiation.

¹Grainger Bioinformatics Center, The Field Museum, Chicago, IL, USA. ²Department of Ecology & Evolutionary Biology, Yale University, New Haven, CT, USA. ³Negaunee Integrative Research Center, The Field Museum, Chicago, IL, USA. ⁴Department of Biology and Museum of Southwestern Biology, University of New Mexico, Albuquerque, NM, USA. ⁵Iridian Genomes, Inc., 6213 Swords Way, Bethesda, MD, USA. ⁶Committee on Evolution Biology, University of Chicago, Chicago, IL, USA. ⁷Present address: Negaunee Integrative Research Center, The Field Museum, Chicago, IL, USA. ⁸Present address: Committee on Evolution Biology, University of Chicago, Chicago, IL, USA. ✉email: celiason@fieldmuseum.org

Novel behaviors can promote phenotypic diversity by allowing organisms to use resources in new ways, thereby causing selection towards different adaptive peaks^{1,2}, or hinder phenotypic change through functional constraints³. Despite considerable research on morphological changes associated with behavioral convergence⁴, less is known about the genomic basis of behavioral convergence⁵. Part of what makes this such a challenge is the complex mapping between genotype and phenotype⁶. Because of this many-to-one mapping, there are many ways that genomic convergence can play out in evolutionary time. For example, we might see convergence in the same amino acid sites (e.g., polar fish evolving antifreeze blood)⁷, the same genes but different amino acids (e.g., high altitude phenotypes in hummingbirds)⁸, or the same pathways but entirely different genes⁹. Furthermore, the expected number of convergent substitutions across a clade is expected to be low⁸ owing to species-specific differences in the genotype-phenotype map¹⁰.

Transitions from air to water pose major sensory and locomotor challenges that affect the evolution of both genes and morphological structures involved in these functions¹¹. For example, recent work in aquatic mammals found genomic convergence in response to selective pressures associated with diving and thermoregulation^{12,13}. Similarly, loss of function in genes important for taste perception in penguins was suggested to be a consequence of swallowing fish whole¹⁴. Although many birds eat fish, diving into water at high speeds (i.e., plunge-diving) is rare, evolving in only a few avian families: Pelicanidae (pelicans), Laridae (terns), Phaethontidae (tropicbirds), Sulidae (gannets and boobies), and Alcedinidae (kingfishers). Kingfishers in particular comprise a species-rich clade of birds (116 recognized species^{15,16}) with diverse diets—ranging from aquatic to terrestrial prey and including both invertebrates (e.g., crustaceans and insects) and vertebrates (e.g., fishes and squamates)¹⁷—and multiple evolutionary origins of plunge-diving behavior¹⁸. Evidence for morphological convergence in beak¹⁸ and brain shape¹⁹ concomitant with these repeated shifts to plunge-diving makes kingfishers an ideal group for studying the genomic basis of convergent phenotypes.

In addition to diverse foraging behaviors, many kingfishers are geographically isolated on oceanic islands recently associated with rapid rates of brain shape evolution¹⁹ and elevated speciation rates^{15,20}. Species on islands experience different evolutionary pressures than those on continents²¹. Divergent selection among islands or between the mainland and islands might therefore drive phenotypic changes (e.g., island gigantism) and genomic changes. Alternatively, genetic drift might reduce adaptive potential on islands. For example, genetic diversity of island species is positively correlated with island size²². Correspondingly, genetic drift should have an outsized effect on island populations and thus decrease overall genomic convergence. Yet, comparisons of effective population size between continental and island taxa through time, as well as patterns of genomic convergence, have not been thoroughly tested across a species-rich clade. Nearly neutral theory²³ predicts that small populations should experience higher rates of non-synonymous substitutions than large populations²⁴. This has been demonstrated empirically in island birds²⁵, as well as more recently across a broader range of taxa²⁶. Because of this, selection is expected to be less efficient in small island populations^{27,28}.

Here, we use whole-genome data from 30 kingfisher species to assess whether convergent evolution of plunge-diving behavior is associated with convergence at the genomic level. First, we compared temporal patterns of population demographics to test the prediction that island species would have more fluctuations and lower effective population sizes than continental species. Second, we compared genomic convergence associated with a phenotype (i.e., plunge-diving behavior) that has evolved in both continental and island systems to test the prediction that genetic

drift associated with small island populations should lessen the strength of genome-wide convergence. Third, we compared levels of positive selection between continental and island plunge-diving species to weigh relative support for drift and selection operating on islands. Given that plunge-diving behavior has emerged independently in both continental and island-dwelling kingfishers, we hypothesized that positive selection would act more strongly on genes related to plunge-diving behavior in continental kingfishers as compared to island species. An alternative hypothesis is that smaller body sizes or environmental heterogeneity among islands²⁹ might drive increased rates of positive selection on islands. We anticipate our results will have implications for our understanding of convergent evolution across scales and provide a blueprint for further functional genomics work on the sensory capabilities of birds.

Results

Effective population size fluctuates more on islands. Using a large-scale genomic dataset of 30 kingfisher species, including a newly annotated Collared Kingfisher (*Todiramphus chloris collaris*) genome³⁰, we fit neutral models of sequence evolution in PAML (model M0) to estimate molecular branch lengths for each annotated gene³⁰. We then used synonymous mutation rates for each species as input in a pairwise sequentially Markovian coalescent (PSMC) analysis of demographic history through time. Using a recent time-calibrated phylogeny of kingfishers²⁰, we identified four origins of plunge-diving behavior and five origins of island-dwelling (Fig. 1). Many species living only on oceanic islands had similarly shaped PSMC curves that fluctuated more than continental species (Fig. 1). Rather than focusing on absolute values of effective population sizes (N_e) that are known to be strongly affected by uncertainty in generation times and inaccurate for recent timeframes³¹, we instead focused on the overall shape of PSMC curves analyzed in a multivariate phylogenetic comparative framework³². These analyses showed no support for the idea that plunge-diving species differ in population size fluctuations through time ($P_{\text{rand}} = 0.25$), yet there was significant divergence in curve shape between continents and islands ($P_{\text{rand}} < 0.01$; Supplementary Fig. 1), with species living solely on islands showing more drastic decreases in effective population size towards the present (Fig. 1). This result echoes a recent study on population demographics in montane birds in which species with smaller ranges showed greater demographic fluctuations³³. Convergent evolution of traits that enable species to adapt to unique ecological conditions on islands (e.g., small geographic areas, reduced species diversity, low predation pressures) might be driven by convergent genomic changes, as has been recently documented in boas³⁴. However, reduced population sizes on islands might also result in lower convergence in genomes of island taxa. This is because stronger genetic drift in small populations reduces the effectiveness of selection, decreasing the probability of independently evolving the same phenotypic or genetic mutations³⁵. Based on this idea, we predicted that island kingfishers would show little evidence for genome-wide convergence.

Genome-wide convergence is associated with plunge-diving behavior. To test for genome-wide convergence, we extracted ancestral states of amino acid sequences for each node of the phylogeny from our PAML M0 model fits. For each gene, we then estimated the number of convergent substitutions in plunge-diving lineages (same amino acids evolving from a different ancestral amino acid) and the number of divergent substitutions (i.e., different amino acids evolving from the same ancestral amino acid). Bayesian phylogenetic analyses showed that plunge-diving lineages have significantly more convergent

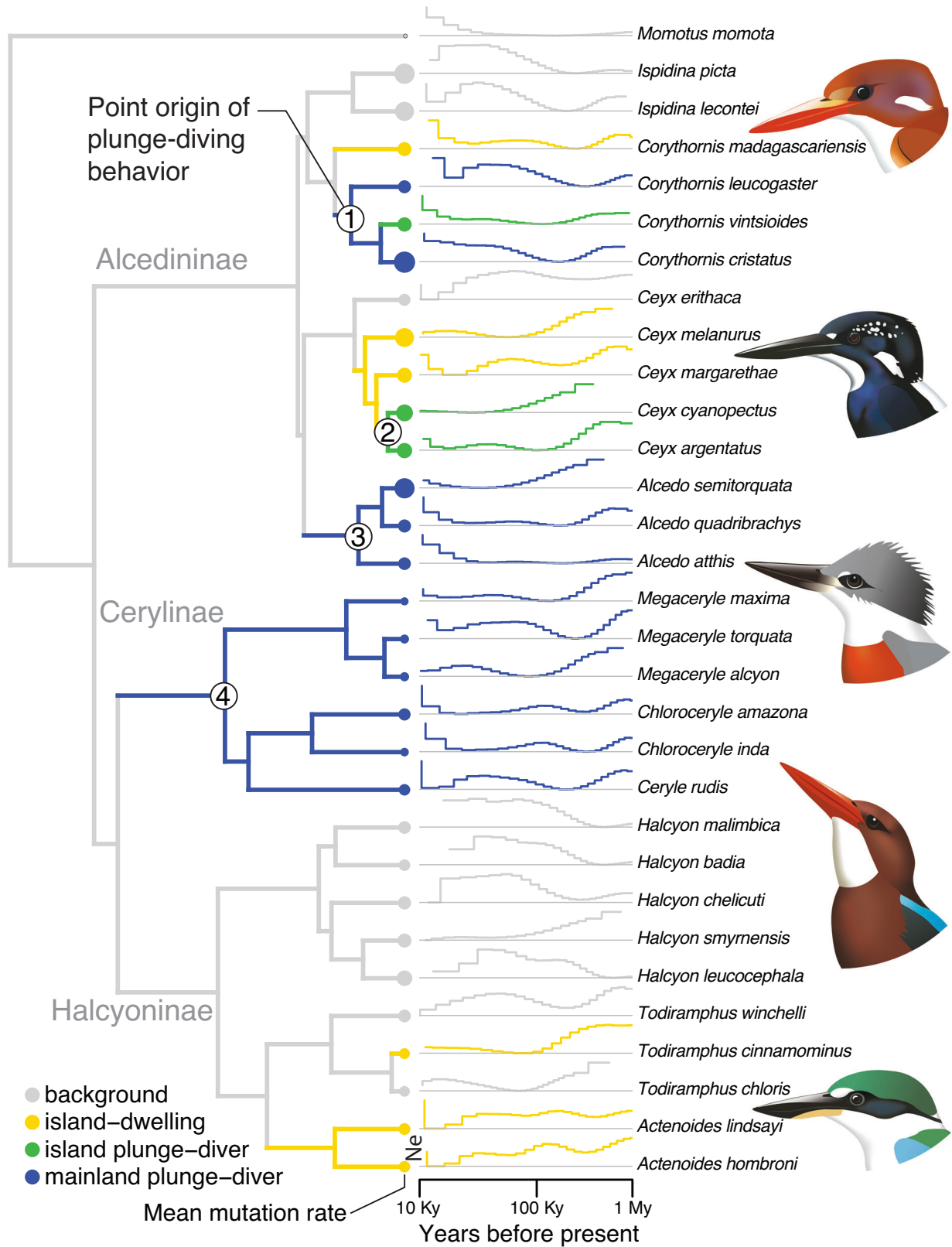


Fig. 1 Convergent feeding behavior and population demographics of kingfishers. Phylogeny of kingfisher species sampled with relationships and branch lengths from McCullough et al.²⁰. Branch colors correspond to estimated ancestral states of island-dwelling and plunge-diving behavior (see legend). Point sizes at tips correspond to mutation rates estimated from average of PAML MO trees. Four evolutionary origins of plunge-diving behavior are labeled at nodes. Step plots at right show effective population size (N_e) over time estimated using pairwise sequentially Markovian coalescent (PSMC) modeling. Representative species are (from top to bottom): Madagascar Pygmy Kingfisher (*Corythornis madagascariensis*), Southern Silvery Kingfisher (*Ceyx argentatus*), Ringed Kingfisher (*Megaceryle torquata*), Brown-breasted Kingfisher (*Halcyon gularis*), and Collared Kingfisher (*Todiramphus chloris*). Illustrations created by J. M. M.

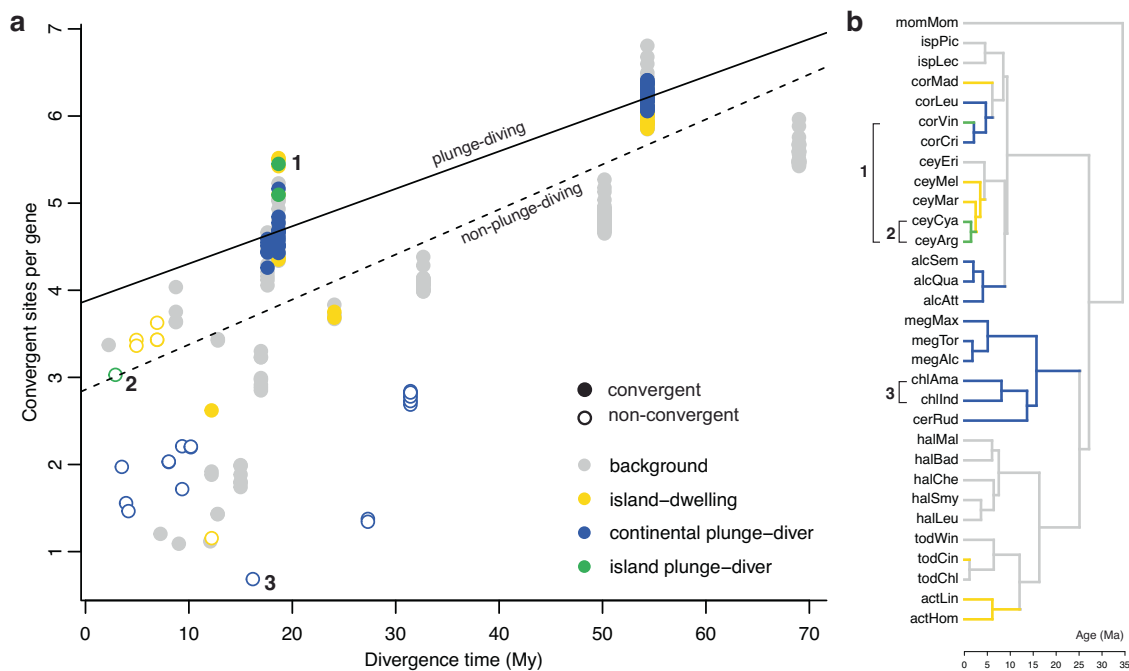


Fig. 2 Plunge-diving kingfishers are convergent across the genome. **a** Scatter plot showing number of convergent sites per gene versus divergence time for all pairs of species. Solid circles denote pairs in the focal set (i.e., those that have convergently evolved plunge-diving behavior or independently colonized islands), and hollow circles denote non-convergent pairs. Point colors correspond to plunge-diving behavior and island-dwelling (see legend). Lines are linear fits to illustrate trends. The effect size for plunge-diving behavior was $\eta^2 = 0.08$, interpreted as a medium effect size³⁶. Labels depict exemplar convergent (pair 1: *Corythornis vintioides* and *Ceyx argentatus*) and non-convergent species pairs (pair 2: *Ceyx cyanopectus* and *Ceyx argentatus*; pair 3: *Chloroceryle Americana* and *Chloroceryle inda*), illustrated on the tree in **(b)**.

Table 1 Predictors of genome-wide convergence in kingfishers.

Predictor	Estimate	(95% CI)	P_{MCMC}
Intercept	-1.411	(-6.354, 3.613)	0.560
Divergent substitutions	-0.327	(-0.374, -0.281)	<0.001
Time	0.112	(0.062, 0.162)	<0.001
Plunge-diving behavior	0.062	(0.012, 0.112)	0.015
Island	-0.007	(-0.062, 0.048)	0.805
Div. substitutions \times time	0.004	(0.003, 0.005)	<0.001

Result of MCMCglmm analysis with phylogenetic relationship incorporated. Response variable is the number of convergent substitutions per gene. Phylogenetic signal of the residuals was 0.81 (95% CI 0.72-0.88). Effect size (η^2) of plunge-diving behavior was 0.08 (calculated from an ordinary least squares regression fit using `lm` in R), indicative of a medium-sized effect³⁶.

amino acid changes than would be expected under a neutral model of sequence evolution in which convergent substitutions scale with divergent substitutions similarly for both plunge-diving and non-plunge-diving lineages (Fig. 2b, Table 1). The effect size (η^2) for plunge-diving behavior was 0.08 (Table 1), suggestive of a medium effect size³⁶. We found no support for genome-wide convergence on islands (Table 1). Given the genome-wide convergence in plunge-divers, we next asked what genes were convergently evolving.

Several genes are adaptively convergent in plunge-divers.

Identifying genes adaptively converging in plunge-diving lineages requires evidence of both amino acid convergence and positive selection. We looked for signatures of positive selection in plunge-diving lineages using branch-site models^{37,38} and protein convergence tests³⁹ (see Supplementary Fig. 2 for workflow). We identified 6091 convergent genes (Fig. 3b) out of 11938 total genes (Supplementary Fig. 3). Focusing on the smaller subset of genes both convergently evolved and under positive selection in

plunge-diving kingfishers, we identified 93 adaptively convergent genes (Fig. 3b) distributed uniformly across the genome (Fig. 3c). These genes are primarily involved in vasculature development (*ANPEP*, *CD34*, *COL3A1*, *HDAC7*, *SLC2A10*, *MIA3*), cognition (*MAPT*, *AGT*, *HRH1*, *DOP1B*, *SPECC1*, *SPG11*), chromosome segregation (*BRCA1*, *BUB1*, *SMC4*, *MLH3*, *NUSAP1*, *SMAR-CAD1*, *SPICE1*), retina homeostasis (*ALB*, *BBS10*, *NPHP4*), and microtubule-based movement (*MAPT*, *HYDIN*, *SPG11*, *DNAI3*, *KIF6*, *NPHP4*) (Fig. 3a; see Supplementary Data 1 for complete gene list).

Positive selection associated with plunge-diving is stronger on islands.

To test whether island plunge-divers have experienced weaker positive selection than continental plunge-divers, we determined lineage-specific positive selection for each gene and branch on the kingfisher phylogeny (see Supplementary Data 2 for gene list). Multiple list GO term enrichment with metaspice⁴⁰ revealed island plunge-divers clustered together in GO term co-occurrence space (Fig. 4, Supplementary Data 2). When we analyzed the number of positively selected genes (PSGs), we found that the number of PSGs was significantly related to body size, feeding behavior, and insularity, with a significant interaction between body mass and island-dwelling kingfishers (Fig. 5a, Table 2). Overlap between terms was also significantly higher in island plunge-divers compared to continental plunge-divers (Fig. 5b). Body size is expected to covary with rates of molecular evolution⁴¹, thus the result is not surprising. Given the decreases in effective population sizes on islands (Fig. 1), the elevated positively selected genes in island plunge-diving species are not consistent with nearly neutral theory²⁴ or recent empirical evidence^{27,42}, but instead suggest potential shifts in the strength or direction of selection on islands.

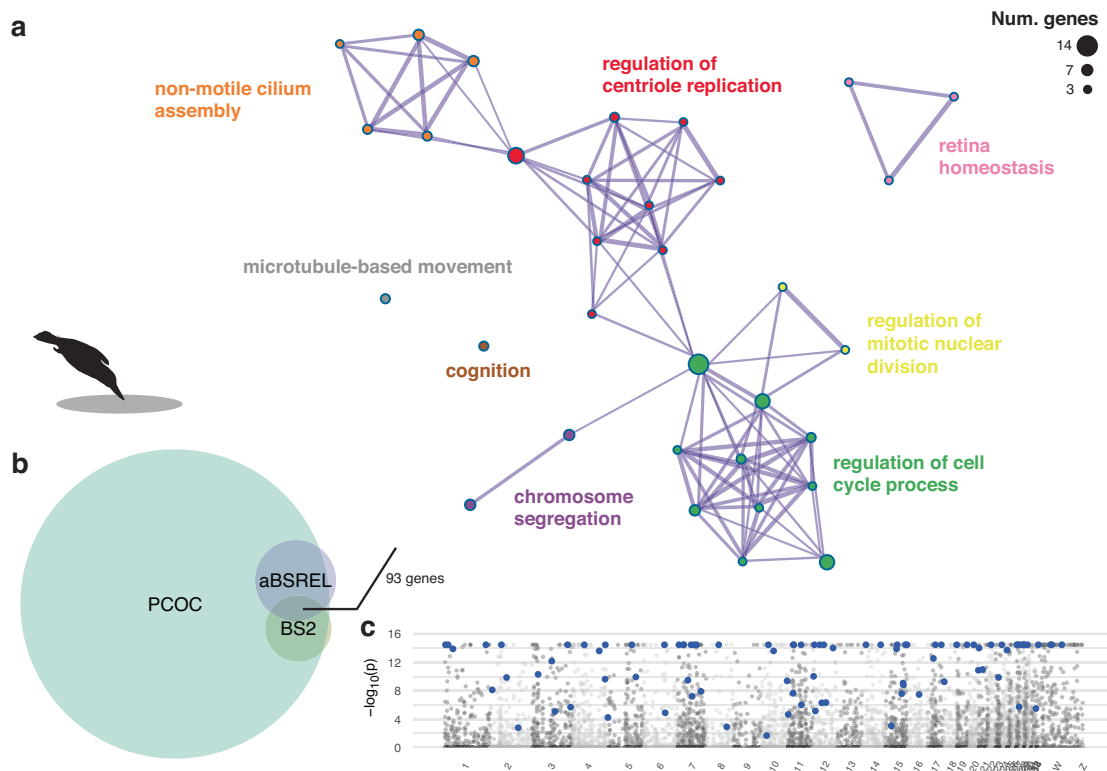


Fig. 3 Evidence for adaptive genomic convergence in plunge-diving kingfishers. Euler diagram (**b**) shows overlapping sets of genes that were identified as significant by positive selection tests (BS2, aBSREL) and genomic convergence tests (PCOC). Branch-site (BS2) tests were further filtered to only include those genes also identified as positively selected in plunge-divers under the more stringent Hyphy RELAX model. Size of circles corresponds to the number of genes, with 93 genes identified as significant in all models. **a** Network graph depicting significantly enriched biological functions (FDR < 0.05) in the focal set of 93 adaptively convergent genes. Node size indicates number of genes for that GO term, colors and labels correspond to most significantly enriched GO term for a given cluster. As an example, the microtubule associated protein gene (*MAPT*) that encodes for tau proteins was significant for all four tests (see Supplementary Fig. 11 for sequence details). **c** Manhattan plot showing $-\log_{10}$ P values (FDR-corrected) for positively selected genes identified across kingfishers (M2 model; see Supplementary Notes 1, 2, Supplementary Figs. 12, 13, Supplementary Table 2, and Supplementary Data 3 for detailed results), with 93 adaptively convergent genes highlighted in blue (see Supplementary Data 1). Chromosome locations of genes were determined by mapping the scaffold-level collared kingfisher (*Todiramphus chloris collaris*) to a chromosome-level assembly for the Carmine bee-eater (*Merops nubicus*) using liftOff⁹⁴. Photo credit of diving kingfisher: Jean and Fred Hort (CC BY 2.0).

Discussion

Our results suggest that plunge-diving kingfishers are convergent genome-wide (Fig. 2a) and that specific regions involved in tau protein formation in the brain and other dietary and microtubule genes are driving this pattern (Fig. 3a). This contrasts with recent evidence in anoles⁴³ and marine mammals¹² that show decoupling between phenotypic and molecular convergence.

Diving into water to forage likely involves adaptations to the eye⁴⁴, as well as changes in brain¹⁹ and beak shape¹⁸. Genetic variation in two retinal genes under positive selection in plunge-divers (*BBS10* and *ATRIP*) have been linked to lower visual acuity in mice^{45,46}. Other adaptively convergent genes include *CENPJ*, a gene that has been linked to decreased brain size in primates⁴⁷, and *MAPT*, a gene that codes for tau proteins that stabilize microtubules in the brain⁴⁸. Proper functioning of *MAPT* has been suggested to mitigate concussion-related brain injuries during foraging in woodpeckers⁴⁹. This is interesting given that kingfishers show convergent evolution in brain shape that is further tied to their diversification¹⁹, along with evidence of convergence at the genomic level (Fig. 2a). Given the importance of neck musculature and beak shape in species that perform high-speed dives into water^{18,50}, our findings are consistent with the idea that brain shape and tau protein evolution might be adaptive in allowing plunge-diving species to undergo daily, repetitive concussive forces without sustaining serious injuries. In addition to morphological adaptations, shifts to a

fish-eating diet likely involve shifts in dietary physiology⁵¹. The *AGT* gene in the cognition cluster of enriched genes (Fig. 3a) codes for an enzyme that has recently been linked to dietary variation in other bird species⁵², hinting at a potential role in dietary shifts of plunge-diving species. The relationship with chromosome segregation is interesting given some of the highest chromosome numbers have been described in small plunge-diving kingfishers in the Alcedininae subfamily⁵³. Further functional genomic work will be needed to test whether these convergent gene changes are associated with adaptive protein function in plunge-diving species.

Our finding of a high proportion of positively selected genes linked to convergent feeding behavior using branch-site tests (Figs. 3a, 5a) echoes work on genes involved with beak shape variation in birds⁵⁴, feeding behavior in pandas⁵⁵, social behavior in snapping shrimp⁵⁶, and extreme environments in cyprinoid fishes⁵⁷. Interestingly, we identified only a small fraction of genes (0.19%) showing support for convergent rate shifts in plunge-diving kingfishers (Supplementary Fig. 4). Similarly low numbers of rate shifts have been found for genes linked to convergent social behavior in spiders⁵⁸, bioelectric cues in flying fish⁵⁹, and miniaturization in wasps⁶⁰. By contrast, subterranean behavior⁶¹ and hairlessness in mammals⁶² have both been associated with high numbers of rapidly evolving genes. This could be due to stronger effects of subterranean behavior on morphology, as compared to plunge-diving behavior. Alternatively, in the case of hairlessness,

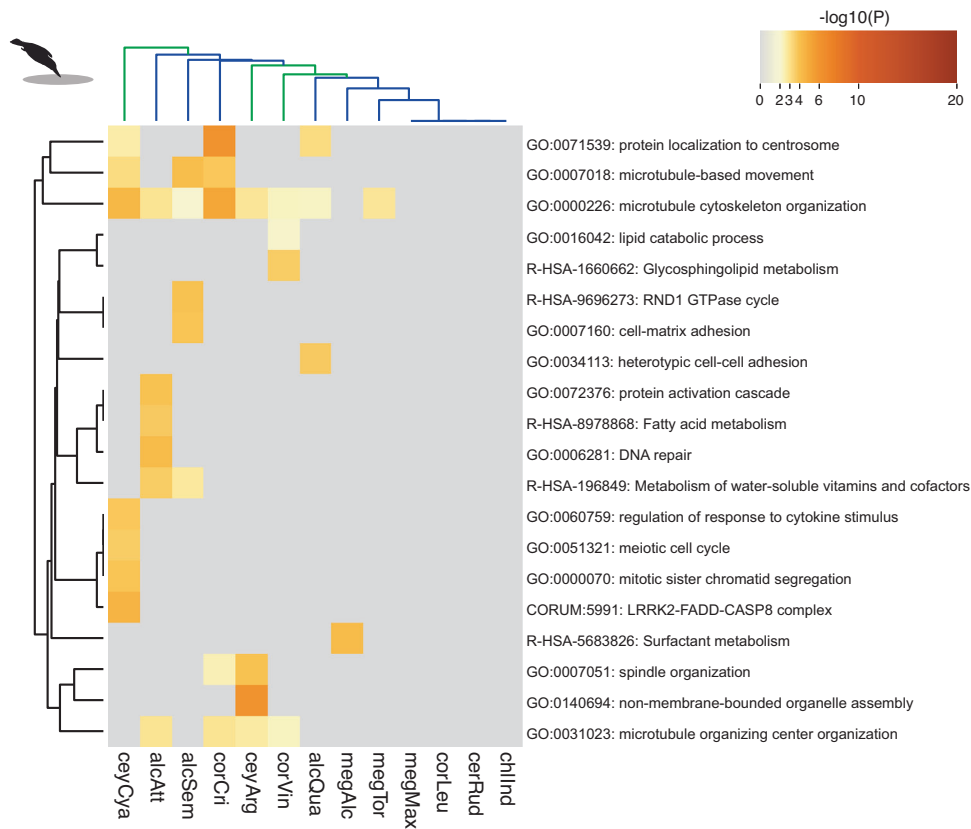


Fig. 4 Comparing enriched gene functions between continental and island plunge-diving kingfishers. Clustering analysis of top gene ontology (GO) terms associated with positively selected genes (PSGs) in plunge-diving species as estimated with the HyPhy aBSREL model⁸³ (see Supplementary Data 2). Note that trees shown are not representative of phylogenetic relationships, but rather depict semantic similarity in GO terms (left) and co-occurrence of positively selected genes associated with these terms (top). Color of cells corresponds to significance level of that GO term enrichment (see legend). Branch colors in upper part of plot correspond to continental (blue) and island-dwelling plunge-divers (green). Plot was produced using metaspice⁴⁰. Photo credit of diving kingfisher: Jean and Fred Hort (CC BY 2.0).

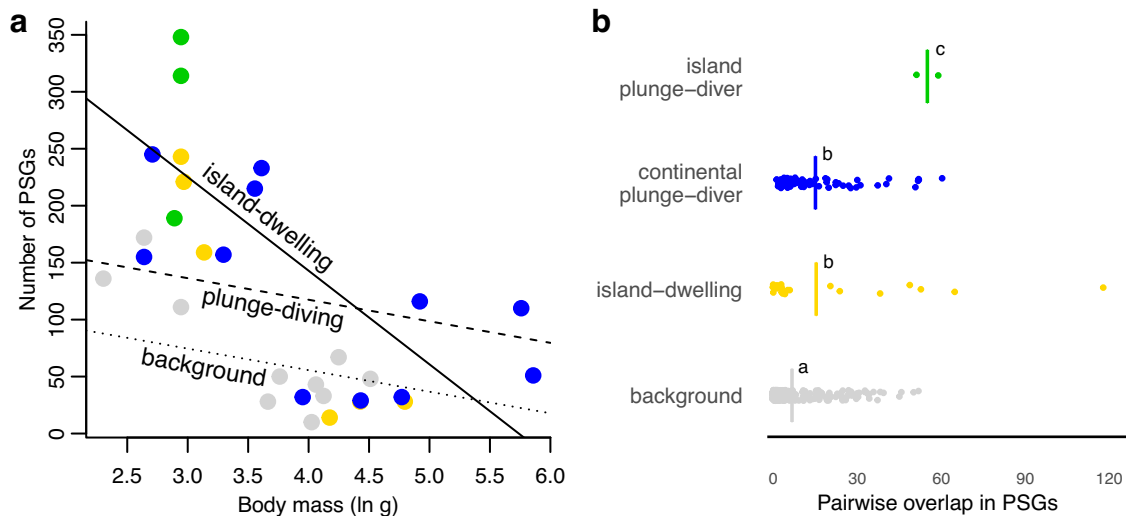


Fig. 5 Levels of positive selection are highest in small plunge-diving species living on islands. **a** Number of positively selected genes (PSGs) versus body size (see legend for point color descriptions). Lines show phylogenetic generalized least squares results of a strongly negative relationship between the number of PSGs and body mass in island kingfishers (solid line) compared to plunge-divers (dashed line; mass-by-insularity interaction, $p = 0.04$) and background (i.e., non-plunge-diving species; dotted line). Note: lines are derived from a PGLS model with moderate phylogenetic signal ($\lambda = 0.55$), thus slopes are shallower than would be expected under an ordinary least squares model without accounting for phylogeny. See Table 2 for statistical details. **b** Overlap in the number of PSGs between pairs of species in the focal (i.e., convergent) and background groups. Points are species pairs and vertical lines are average values. Groups sharing the same letter are not significantly different (i.e., $p > 0.05$).

Table 2 Predictors of the number of positively selected genes in kingfishers.

Predictor	Estimate	Std. err.	t	p
Intercept	131.3	73.8	1.78	0.09
Plunge-diving behavior	61.8	23.0	2.68	0.01
Body mass (ln g)	-18.9	18.2	-1.04	0.31
Insularity	278.7	101.8	2.74	0.01
Body mass × insularity	-63.3	28.9	-2.19	0.04

Results of a PGLS model with number of positively selected genes as the response variable. Model $n^2 = 102.6$, phylogenetic signal (λ) = 0.55, adjusted $R^2 = 0.38$.

morphological changes relate directly to changes in the genome, whereas behavioral traits tend to be more polygenic, meaning they are influenced by multiple genes with smaller individual effect sizes. Consequently, the likelihood of convergent evolution in behavioral traits is reduced, as there are numerous alternative pathways for acquiring the same phenotype. The latter hypothesis predicts a greater number of convergent rate shifts for genes associated with morphology compared to plunge-diving behavior.

Signatures of convergence might also vary depending on what part of the genome is being assessed. Indeed, we found proportionally more accelerated genes than CNEs linked to plunge-diving behavior (see Supplementary Note 3). This differs from recent work on limb loss in reptiles⁶³, flightlessness in birds⁹, and wolves⁶⁴ showing a strong effect of convergence in regulatory regions of the genome compared to coding regions. Challenges in comparing genomic convergence across taxa stem from differences in analytical approach (e.g., HyPhy versus PAML, RERconverge versus phyloACC) and in what parts of the genomes are assessed (e.g., coding versus regulatory regions). Overall, studies showing high numbers of rate shifts in genic and regulatory regions tend to be those involving either extreme morphological changes (e.g., loss of limbs or hair)^{9,62,63} or behaviors expected to have strong downstream effects on morphology (e.g., loss of flight)^{9,65}. Changes in behavior likely involve several trait systems (e.g., sensory, locomotor), which might explain the emerging trend of morphological convergence being linked to stronger signatures of molecular convergence than behavioral convergence. Consistent with past work, we find overlap but also differences in sets of genes identified by different methods (Supplementary Fig. 5). Further work will be needed to test the functions of putative genes identified as critical for feeding adaptations in kingfishers.

Our study provides evidence for adaptive convergence of brain, digestion, and microtubule-related genes associated with convergent plunge-diving feeding behavior and uncovers previously unrecognized diversity of taste receptor genes in kingfishers relative to other avian species. Genomic convergence was linked to plunge-diving behavior but not insularity, consistent with our PSMC results showing reduced population sizes in island taxa that are expected to lower the potential for genome-wide convergence. However, recovered high levels of positive selection in small plunge-diving species suggests that distinct selective pressures on islands could play a role in driving rapid molecular evolution given the clear connection between feeding behavior and island adaptation (e.g., Darwin's finches). Together, these results elucidate genomic bases of behavioral convergence and set the stage for further work on functional genomics and sensory system evolution.

Methods

Genome resequencing. We were loaned tissues from frozen samples at the Field Museum of Natural History (see Supplementary Table 1). We chose species to maximize taxonomic coverage and dietary diversity. We extracted genomic DNA with the

Qiagen DNAeasy prep kit according to the manufacturer's directions, and we prepared genomic libraries with the Illumina TruSeq kit using standard adapters. We sequenced libraries at ~70X average coverage (range: 47.7X–95.1X) with 150 bp paired-end reads on an Illumina HiSeq X Ten platform (Genewiz, New Jersey, USA). We removed duplicate reads with SUPERDEDUPER v. 1.3.2 (available at <https://github.com/ibest/HTStream>). We then used FASTP v. 0.20.1⁶⁶ to assess read quality. The majority of reads had low adapter content (<1%) and were of high quality (Q > 30).

Genome assembly. We aligned cleaned reads to the draft collared kingfisher genome³⁰ using BWA-MEM v. 0.7.17-r1188⁶⁷. This read-mapping approach gave a similar result as NextGenMap⁶⁸, suitable for highly divergent sequences⁶⁹, as well as LAST⁷⁰ in a subset of species (Supplementary Fig. 6), therefore we opted for the faster BWA-MEM. We converted alignments to BAM format with SAMtools VIEW v. 1.10 and sorted BAM alignments with SAMtools SORT. We performed INDEL realignment with GATK v. 3.8⁷¹ and finally used BCFTOOLS mpileup v. 1.8 and SAMtools vcfutils.pl vcf2fq to call genotypes and generate individual consensus sequences (i.e., pseudo-genomes). Bases with low coverage (<5X) or low mapping quality (Q < 30) were N-masked in the final consensus sequences. We used the “-I” option in BCFTOOLS mpileup to exclude INDELS because we wanted genomic coordinates to be the same as the reference genome⁷². This allowed us to use the reference genome annotations (GFF format) obtained from GeMoMa and extract coding sequences for each individual. All scripts for data analyses are available at Zenodo⁷³.

Reference bias assessment. To determine the proportion of raw reads mapped to the reference genome, we computed read-mapping statistics with the flagstats function of SAMtools. This showed that between 76.9–99.4% of reads were mapped to the reference genome, with depth of coverage ranging from 11X–20.4X (Supplementary Table 1, Supplementary Fig. 7). Aligning all species reads to the same reference genome can result in mapping bias, in which species-specific reads that are more similar to the reference map more efficiently than those from more distantly related species. To deal with this potential source of bias, we used an N-masking approach⁷⁴ in which we N-masked all variable sites and re-mapped reads of three species differing in their phylogenetic divergence to the reference collared kingfisher (Supplementary Fig. 8). If mapping bias is an issue, we would predict that more distantly related species would show a significant increase in the proportion of mapped reads after N-masking, whereas closely related species would not be affected by N-masking. Our results did not show this pattern, suggesting that reference bias is not strongly influencing our results (Supplementary Figs. 8, 9).

Population demographic history. To understand population demographic fluctuations over time, we used PSMC v. 0.6.5⁷⁵. We used PSMC settings (-N25 -t15 -r5 -p'4 + 25 * 2 + 4 + 6') from a recent study on penguin genomes that used a similar assembly approach to ours⁷². To plot effective population sizes (Ne) as a function of time, we used published R code⁷⁶ to input data into R with a generation time of 4.8 years (IUCN 2021) and mutation rates taken from the average substitution rate estimated with the PAML M0 model (see below). These rates were comparable to the value used for the common kingfisher (*Alcedo atthis*, 0.11 mutations per million years) in a larger study on the evolution of mutation rates⁷⁷. To statistically test whether the shapes of PSMC curves differ between islands and continents, we used a multivariate clustering approach. Briefly, we first determined the time range with Ne data for all 31 species. Next, we log-transformed the Ne values. Finally, after converting these data into an array,

we used the `procD.pgls` function³² in the R package `geomorph`⁷⁸ to run a distance-based phylogenetic generalized least squares analysis, with either insularity or feeding behavior as a predictor and a recently published kingfisher timetree²⁰.

Assessing genome-wide convergence. We tested the hypothesis that plunge-diving behavior results in genome-wide convergence following Sackton et al.⁹. Briefly, we determined the total number of convergent amino acids (i.e., changes in specific amino acids having evolved convergently in two species compared to the most recent common ancestor, MRCA) and divergent amino acids (i.e., amino acids that differ in two species as compared to their MRCA) for each pair of lineages. In calculating convergent and divergent sites, we opted to treat lineages as the whole path from an ancestor to tip rather than single branches⁷⁹. We then used phylogenetic linear mixed models implemented in `MCMCglmm`⁸⁰ to test whether the number of convergent changes was explained by feeding behavior, island dwelling, divergence time, number of divergent amino acid substitutions, or the interaction between divergent substitutions and divergence time. Given potential inflation of effect sizes due to non-independent samples (e.g., single lineage compared to multiple other lineages), we treated each lineage in a pairwise comparison as a random effect in the model, after shuffling lineage pairs to ensure even occurrence of each lineage in a comparison⁸¹. We also accounted for phylogenetic non-independence by treating the MRCA of each pair as a random effect, with an expected covariance structure determined by the phylogenetic tree calculated with the `ginverse` function in `MCMCglmm`. We ran the model for 10^6 generations and discarded the first 25% as burn-in. Phylogenetic signal was calculated as the posterior mean of the node random effect divided by the sum of all random effects scaled by $(1/\pi^{2/3})$ ⁸¹.

To further test which genes were convergently evolved in plunge-diving kingfishers, we used the program `PCOC`³⁹. This model tests for signatures of convergence in all residues of a protein. We set up the cutoff of the posterior probability that a given residue is evolving convergently in plunge-diving lineages to 0.8. Genes that showed support for convergence were retained and used in downstream enrichment analyses (see below).

Detecting signatures of positive selection. Positive selection in DNA sequences is typically tested by estimating ω , the ratio of the rate of nonsynonymous (d_N) nucleotide substitutions to that of synonymous substitutions (d_S) for aligned homologous protein sequences³⁷. Support for $\omega > 1$ is interpreted as evidence for positive selection, whereas $\omega < 1$ suggests purifying selection (i.e., selective constraints) and $\omega = 1$ indicates neutral evolution. For each aligned coding sequence, we used `PAML`³⁷ to fit a null model (model M0) of sequence evolution assuming a single ω across the genome and phylogeny. For this model, we used a fixed species tree previously estimated using a large UCE data set²⁰. To test whether some genes show a pattern consistent with positive selection, we compared two sets of site models that allow ω to vary across sites in a protein but not across the phylogeny. The first set of comparisons was between a nearly neutral model that assumes two site categories ($\omega_0 < 1$, $\omega_1 = 1$; model M1a) and a positive selection model that allows three site categories ($\omega_0 < 1$, $\omega_1 = 1$, $\omega_2 > 1$; model M2a). The second comparison of site models was between a model that allows ten site categories with $\omega < 1$ along a beta distribution (model M7) and a model that allows 11 site categories, including one with $\omega > 1$ (model M8). Finally, to test the hypothesis that some genes and sites show positive selection in species that plunge dive into water relative to species that do not, we fit a branch-site model that allows for $\omega > 1$ only in predefined foreground (i.e., plunge-diving) lineages (model BS2) to one that assumes $\omega = 1$ for both plunge-diving and

non-plunge-diving species (model BS1). For all genes and models, we used the molecular branch lengths estimated under the neutral model M0 to speed up computational time⁹. We compared sets of models using likelihood ratio tests and retained ancestral state estimates under the M0 model for convergent evolution tests (see below). Models were run iteratively over all annotated genes using custom R scripts⁷³. To account for multiple comparisons, we adjusted P values using false discovery rates (FDR) and retained all genes with significant evidence (FDR < 0.05) for positive selection. Sets of positively selected genes for each model comparison were retained and used in subsequent gene enrichment tests.

We are aware of recent work suggesting that branch-site models can overestimate the number of significantly converged genes⁸². This is because branch-site models do not account for potential background positive selection independent of the focal lineages (i.e., plunge-divers). To account for this effect, we further fit HyPhy RELAX models³⁸ with the same trees used in our branch-site tests. This model is more stringent in testing whether convergent shifts in amino acid dN/dS ratios are sensitive to background lineages. We then filtered branch-site results to include only those genes that also showed significant support for positive selection in foreground lineages under the RELAX model. In total, we confirmed that 308 out of 1426 genes (21.6%) that showed support for positive selection in plunge-divers under the branch-site model were also significant in the more stringent RELAX model.

We also fit the adaptive branch-site random effects likelihood (aBSREL) model⁸³ in HyPhy. This test conducts significance tests for each branch and gene. To compare the set of positively selected genes associated with plunge-diving behavior among CODEML models and HyPhy models, we retained gene-specific aBSREL tests that showed any significant shifts along a branch (Supplementary Fig. 10). For the aBSREL results, we filtered positively selected genes to include those that were convergently positively selected in at least two lineages (e.g., *Ceyx argentatus* and *Corythornis vintsioides*; see Fig. 1). This resulted in a set of 418 genes showing support for convergent positive selection. Finally, to determine the set of genes that were adaptively convergent in plunge-divers, we required that a gene be present in the PCOC, RELAX, aBSREL convergent, and branch-site (BS2) gene sets. This resulted in a final set of 93 putative adaptively convergent genes (Fig. 3b; see Supplementary Fig. 2 for further details).

Alternative approaches for identifying convergent positively selected genes. We used several approaches to identify sets of genes putatively convergent and positively selected in plunge-diving kingfishers. In addition to the previous analyses, we also used (i) a modified branch-site test that accounts for the issue of positive selection in background species in the `codeml` BS2 model and (ii) a recent method for detecting Combinatorial SUBSTITUTIONS of codon sequences (CSUBST) that identifies genes in pairs of species that are both convergent and under positive selection⁸⁴. For the latter, we filtered genes to only include those that showed support for adaptive convergence in two or more pairs of species that have convergently evolved plunge-diving. See Supplementary Note 4 and Supplementary Data 4, 5 for results.

Comparing positive selection between islands and continents. We tabulated the number of positively selected genes for each plunge-diving species branch from the aBSREL results and then used `metascape`⁴⁰ to analyze the co-occurrence of different gene functions among species. To test whether positive selection is operating more strongly in island-dwelling kingfishers, we calculated the number of PSGs for each species as well as the amount of overlap between pairs of species. Since small body sizes are associated with elevated rates of molecular evolution⁴¹, we included

body mass (In grams) as a covariate in the model, along with insularity, plunge-diving behavior, and their interactions with body size. We determined the best-fitting model using bi-directional variable removal based on AIC scores in the phylostep function⁸⁵. For gene overlap, we used Bayesian mixed models implemented in MCMCglmm⁸⁰ with number of overlapping PSGs as the response; insularity and plunge-diving behavior as predictors; and species 1, species 2, and the node of the most recent common ancestor in the phylogeny as random effects (see ref. ⁸¹ for details).

Gene enrichment. To understand whether sets of positively selected or convergent genes were associated with certain molecular functions, we used metascap⁴⁰ with default settings and a custom background of collared kingfisher genes. This analysis links genes with their putative functions (i.e., gene ontology, or GO terms) and tests whether a set of target genes with a given function is larger than expected by chance from a random subset of GO terms. We visualized overlapping target genes from different analyses using Venn diagrams and visualized gene networks again with metascap⁴⁰. Significance of enriched gene functions was estimated using the FDR metric.

Identifying conserved noncoding elements. Recent work has found that regulatory regions of the genome, rather than coding regions, are related to shifts in phenotype^{9,54}. To assess whether regulatory regions evolve faster in plunge-diving or island-dwelling kingfishers, we identified a set of avian-specific highly conserved noncoding elements (CNEs)⁸⁶. Since these elements are relative to the chicken genome, we used LAST⁷⁰ to align the draft collared kingfisher genome³⁰ to the chicken genome. We then used liftOver⁸⁷ to convert chicken genome coordinates to collared kingfisher coordinates and extracted DNA sequences with bedtools getfasta⁸⁸. In total, we located 235559 out of 265983 (88.6%) chicken CNEs in the collared kingfisher genome. After filtering out CNEs smaller than 50 bp and removing badly aligned regions with trimAL⁸⁹, we fit estimated molecular branch lengths using the BASEML free model³⁷. This model lets each branch assume its own rate of molecular evolution, a requirement of downstream RERconverge analyses. In total, we obtained a final set of 39618 trees. For those CNEs that showed significant support for a convergent rate shift in focal lineages, we used bedtools closest to identify the closest annotated gene in the collared kingfisher genome and input these annotations into GO term enrichment analyses.

Detecting convergent shifts in rates of molecular evolution. To detect significant accelerations or decelerations associated with plunge-diving behavior, we ran RERconverge⁹⁰ on sets of gene and CNE trees. Importantly, these analyses utilize overall rates of molecular evolution and do not account for variation in specific sites along a gene or CNE. To assess significance of relationships between rates of molecular evolution and phenotypic changes, we used a phylogenetic permutation approach recently shown to be more rigorous than *P* values derived from parametric estimates⁹¹. This method simulates evolutionary histories of a given phenotype (e.g., plunge-diving behavior) and then generates a null distribution of correlation values between the phenotype and rates of molecular evolution. Statistical significance is then assessed by determining the proportion of null values that are less extreme than the empirical correlation for a given gene. We did 500 permutations and corrected *P* values using FDR.

Statistics and reproducibility. Maximum likelihood approaches were used to determine support for significant positively selected genes, with correction for multiple tests using FDR ($n = 11938$ genes tested). Bayesian phylogenetic linear mixed models were used

to analyze genomic convergence and gene overlap ($n = 465$ species pairs). Phylogenetic generalized least squares analysis was used to analyze positively selected gene counts ($n = 30$ species). All statistical analyses were performed in R v. 4.2.3. Code was written in the R markdown format to allow for reproducibility of all analyses and generation of figures from raw data sets.

Reporting summary. Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Raw Illumina reads are available publicly at the Sequence Read Archive (see Supplementary Table 1 for SRR IDs). Ecological and source data underlying figures are available on Dryad (<https://doi.org/10.5061/dryad.gflvhhmyn>)⁹². All genome alignments are available on Zenodo as BAM files (<https://doi.org/10.5281/zenodo.7872534>)⁹³.

Code availability

All code needed to run analyses is available on Zenodo (<https://doi.org/10.5281/zenodo.8291004>)⁷³.

Received: 17 May 2023; Accepted: 14 September 2023;

Published online: 24 October 2023

References

- Wyles, J. S., Kunkel, J. G. & Wilson, A. C. Birds, behavior, and anatomical evolution. *Proc. Natl Acad. Sci.* **80**, 4394–4397 (1983).
- Lapedra, O., Sol, D., Carranza, S. & Beaulieu, J. M. Behavioural changes and the adaptive diversification of pigeons and doves. *Proc. R. Soc. B Biol. Sci.* **280**, 20122893 (2013).
- Collar, D. C., O'meara, B. C., Wainwright, P. C. & Near, T. J. Piscivory limits diversification of feeding morphology in centrarchid fishes. *Evolution* **63**, 1557–1573 (2009).
- Losos, J. B. Convergence, adaptation, and constraint. *Evolution* **65**, 1827–1840 (2011).
- Sackton, T. B. & Clark, N. Convergent evolution in the genomics era: new insights and directions. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **374**, 20190102 (2019).
- Wagner, G. P. & Altenberg, L. Perspective: Complex adaptations and the evolution of evolvability. *Evolution* **50**, 967–976 (1996).
- Chen, L., DeVries, A. L. & Cheng, C. H. Convergent evolution of antifreeze glycoproteins in Antarctic notothenioid fish and Arctic cod. *Proc. Natl Acad. Sci. USA* **94**, 3817–3822 (1997).
- Natarajan, C. et al. Predictable convergence in hemoglobin function has unpredictable molecular underpinnings. *Science* **354**, 336–339 (2016).
- Sackton, T. B. et al. Convergent regulatory evolution and loss of flight in paleognathous birds. *Science* **364**, 74–78 (2019).
- Storz, J. F. Causes of molecular convergence and parallelism in protein evolution. *Nat. Rev. Genet.* **17**, 239–250 (2016).
- Zhao, H., Li, J. & Zhang, J. Molecular evidence for the loss of three basic tastes in penguins. *Curr. Biol.* **25**, R141–R142 (2015).
- Elliott, K. H. et al. High flight costs, but low dive costs, in auks support the biomechanical hypothesis for flightlessness in penguins. *Proc. Natl Acad. Sci.* **110**, 9380–9384 (2013).
- Foote, A. D. et al. Convergent evolution of the genomes of marine mammals. *Nat. Genet.* **47**, 272–275 (2015).
- Yuan, Y. et al. Comparative genomics provides insights into the aquatic adaptations of mammals. *Proc. Natl. Acad. Sci. USA* **118**, e2106080118 (2021).
- Andersen, M. J., McCullough, J. M., Mauck, I. W. M., Smith, B. T. & Moyle, R. G. A phylogeny of kingfishers reveals an Indomalayan origin and elevated rates of diversification on oceanic islands. *J. Biogeogr.* **45**, 269–281 (2018).
- Gill, F. B., Donsker, D. & Rasmussen, P. C. IOC World Bird List (v 12.1). Available at <http://www.worldbirdnames.org>. <https://doi.org/10.14344/IOC.ML.12.1> (2022).
- Fry, C. H. *Kingfishers, bee-eaters & rollers*. (Princeton University Press, 1992).
- Eliason, C. M., Straker, L., Jung, S. & Hackett, S. J. Morphological innovation and biomechanical diversity in plunge-diving birds. *Evolution* **74**, 1514–1524 (2020).
- Eliason, C. M., McCullough, J. M., Andersen, M. J. & Hackett, S. J. Accelerated Brain Shape Evolution Is Associated with Rapid Diversification in an Avian Radiation. *Am. Nat.* **197**, 576–591 (2021).

20. McCullough, J. M., Moyle, R. G., Smith, B. T. & Andersen, M. J. A Laurasian origin for a pantropical bird radiation is supported by genomic and fossil data (Aves: Coraciiformes). *Proc. R. Soc. B: Biol. Sci.* **286**, 20190122 (2019).
21. Losos, J. B. & Ricklefs, R. E. Adaptation and diversification on islands. *Nature* **457**, 830–836 (2009).
22. Manthey, J. D., Oliveros, C. H., Andersen, M. J., Filardi, C. E. & Moyle, R. G. Gene flow and rapid differentiation characterize a rapid insular radiation in the southwest Pacific (Aves: Zosterops). *Evolution* **74**, 1788–1803 (2020).
23. Ohta, T. Slightly deleterious mutant substitutions in evolution. *Nature* **246**, 96–98 (1973).
24. Kimura, M. *The Neutral Theory of Molecular Evolution*. (Cambridge University Press, 1983).
25. Johnson, K. P. & Seger, J. Elevated rates of nonsynonymous substitution in island birds. *Mol. Biol. Evol.* **18**, 874–881 (2001).
26. Woolfit, M. & Bromham, L. Population size and molecular evolution on islands. *Proc. Biol. Sci.* **272**, 2277–2282 (2005).
27. James, J. E., Lanfear, R. & Eyre-Walker, A. Molecular evolutionary consequences of island colonization. *Genome Biol. Evol.* **8**, 1876–1888 (2016).
28. Kutschera, V. E. et al. Purifying selection in corvids is less efficient on islands. *Mol. Biol. Evol.* **37**, 469–474 (2020).
29. Weigelt, P., Jetz, W. & Krefft, H. Bioclimatic and physical characterization of the world's islands. *Proc. Natl Acad. Sci.* **110**, 15307–15312 (2013).
30. Eliason, C. M., McCullough, J. M., Hains, T., Andersen, M. J. & Hackett, S. J. Genome report: genomic novelty within a “great speciator” revealed by a high-quality reference genome of the collared kingfisher (*Todiramphus chloris collaris*). *G3: Genes, Genomes, Genet.* <https://doi.org/10.1093/g3journal/jkac260> (2022).
31. Nadachowska-Brzyska, K., Konczal, M. & Babik, W. Navigating the temporal continuum of effective population size. *Methods Ecol. Evol.* **13**, 22–41 (2022).
32. Adams, D. C. A method for assessing phylogenetic least squares models for shape and other high-dimensional multivariate data. *Evolution* **68**, 2675–2688 (2014).
33. Pujolar, J. M. et al. The formation of avian montane diversity across barriers and along elevational gradients. *Nat. Commun.* **13**, 268 (2022).
34. Card, D. C. et al. Genomic basis of convergent island phenotypes in *Boa* constrictors. *Genome Biol. Evol.* **11**, 3123–3143 (2019).
35. Rosenblum, E. B., Parent, C. E. & Brandt, E. E. The molecular basis of phenotypic convergence. *Annu. Rev. Ecol. Syst.* **45**, 203–226 (2014).
36. Cohen, J. *Statistical power analysis for the behavioral sciences*. (Routledge, 1988). <https://doi.org/10.4324/9780203771587>.
37. Yang, Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13**, 555–556 (1997).
38. Wertheim, J. O., Murrell, B., Smith, M. D., Kosakovsky Pond, S. L. & Scheffler, K. RELAX: detecting relaxed selection in a phylogenetic framework. *Mol. Biol. Evol.* **32**, 820–832 (2015).
39. Rey, C., Guéguen, L., Sémon, M. & Boussau, B. Accurate detection of convergent amino-acid evolution with PCOC. *Mol. Biol. Evol.* **35**, 2296–2306 (2018).
40. Zhou, Y. et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).
41. Bromham, L. Why do species vary in their rate of molecular evolution? *Biol. Lett.* **5**, 401–404 (2009).
42. Leroy, T. et al. Island songbirds as windows into evolution in small populations. *Curr. Biol.* **31**, 1303–1310.e4 (2021).
43. Corbett-Detig, R. B., Russell, S. L., Nielsen, R. & Losos, J. Phenotypic convergence is not mirrored at the protein level in a lizard adaptive radiation. *Mol. Biol. Evol.* **37**, 1604–1614 (2020).
44. Machovsky Capuska, G. E. et al. Visual accommodation and active pursuit of prey underwater in a plunge-diving bird: the Australasian gannet. *Proc. R. Soc. Lond. Ser. B-Biol. Sci.* **279**, 4118–4125 (2012).
45. Mayer, S. K. et al. Progressive retinal degeneration of rods and cones in a Bardet-Biedl syndrome type 10 mouse model. *Dis. Model. Mech.* **15**, dmm049473 (2022).
46. Matos-Rodrigues, G. E. et al. Progenitor death drives retinal dysplasia and neuronal degeneration in a mouse model of ATRIP-Seckel syndrome. *Dis. Model. Mech.* **13**, dmm045807 (2020).
47. Montgomery, S. H., Capellini, I., Venditti, C., Barton, R. A. & Mundy, N. I. Adaptive evolution of four microcephaly genes and the evolution of brain size in anthropoid primates. *Mol. Biol. Evol.* **28**, 625–638 (2011).
48. Kadavath, H. et al. Tau stabilizes microtubules by binding at the interface between tubulin heterodimers. *Proc. Natl Acad. Sci. USA.* **112**, 7501–7506 (2015).
49. Farah, G., Siwek, D. & Cummings, P. Tau accumulations in the brains of woodpeckers. *PLoS One* **13**, e0191526 (2018).
50. Chang, B. et al. How seabirds plunge-dive without injuries. *Proc. Natl Acad. Sci.* **113**, 12006–12011 (2016).
51. Hilton, G. M., Houston, D. C., Barton, N. W. H., Furness, R. W. & Ruxton, G. D. Ecological constraints on digestive physiology in carnivorous and piscivorous birds. *J. Exp. Zool.* **283**, 365–376 (1999).
52. Wang, B.-J. et al. Diet and adaptive evolution of alanine-glyoxylate aminotransferase mitochondrial targeting in birds. *Mol. Biol. Evol.* **37**, 786–798 (2020).
53. Qing-wei, B. X.-Z. L. I. Studies on The Karyotypes of Birds V. The 20 species of Climber birds.(Aves). *Zool. Res.* **10**, 309–317 (1989).
54. Yusuf, L. et al. Noncoding regions underpin avian bill shape diversification at macroevolutionary scales. *Genome Res.* **30**, 553–565 (2020).
55. Hu, Y. et al. Comparative genomics reveals convergent evolution between the bamboo-eating giant and red pandas. *Proc. Natl Acad. Sci. USA* **114**, 1081–1086 (2017).
56. Chak, S. T. C., Baeza, J. A. & Barden, P. Eusociality shapes convergent patterns of molecular evolution across mitochondrial genomes of snapping shrimps. *Mol. Biol. Evol.* **38**, 1372–1383 (2021).
57. Tong, C. & Li, M. Convergent genomic signatures of adaptation to an extreme environment in cyprinoid fishes. *bioRxiv* 2022.03.30.486405 <https://doi.org/10.1101/2022.03.30.486405> (2022).
58. Tong, C., Avilés, L., Rayor, L. S., Mikheyev, A. S. & Linksvayer, T. A. Genomic signatures of recent convergent transitions to social life in spiders. *Nat. Commun.* **13**, 6967 (2022).
59. Daane, J. M. et al. Modulation of bioelectric cues in the evolution of flying fishes. *Curr. Biol.* **31**, 5052–5061.e8 (2021).
60. Xu, H. et al. Comparative Genomics Sheds Light on the Convergent Evolution of Miniaturized Wasps. *Mol. Biol. Evol.* **38**, 5539–5554 (2021).
61. Partha, R. et al. Subterranean mammals show convergent regression in ocular genes and enhancers, along with adaptation to tunneling. *Elife* **6**, e25884 (2017).
62. Kowalczyk, A., Chikina, M. & Clark, N. Complementary evolution of coding and noncoding sequence underlies mammalian hairlessness. *eLife* **11**, e76911 (2022).
63. Roscito, J. G. et al. Convergent and lineage-specific genomic differences in limb regulatory elements in limbless reptile lineages. *Cell Rep.* **38**, 110280 (2022).
64. Feigin, C. Y., Newton, A. H. & Pask, A. J. Widespread cis-regulatory convergence between the extinct Tasmanian tiger and gray wolf. *Genome Res.* **29**, 1648–1658 (2019).
65. Burga, A. et al. A genetic signature of the evolution of loss of flight in the Galapagos cormorant. *Science* **356**, eaal3345 (2017).
66. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
67. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv [q-bio.GN]* (2013).
68. Sedlazeck, F. J., Rescheneder, P. & von Haeseler, A. NextGenMap: fast and accurate read mapping in highly polymorphic genomes. *Bioinformatics* **29**, 2790–2791 (2013).
69. Neme, R. & Tautz, D. Fast turnover of genome transcription across evolutionary time exposes entire non-coding DNA to de novo gene emergence. *Elife* **5**, e09977 (2016).
70. Kielbasa, S. M., Wan, R., Sato, K., Horton, P. & Frith, M. C. Adaptive seeds tame genomic sequence comparison. *Genome Res.* **21**, 487–493 (2011).
71. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
72. Vianna, J. A. et al. Genome-wide analyses reveal drivers of penguin diversification. *Proc. Natl Acad. Sci. USA.* **117**, 22303–22310 (2020).
73. Eliason, C. M. et al. *Code from: Genomic signatures of convergent shifts to plunge-diving behavior in birds.* <https://doi.org/10.5281/zenodo.8291004> (2023).
74. Krueger, F. & Andrews, S. R. SNPsplit: Allele-specific splitting of alignments between genomes with known SNP genotypes. *F1000Res.* **5**, 1479 (2016).
75. Li, H. & Durbin, R. Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493–496 (2011).
76. Liu, S. & Hansen, M. M. Data from: PSMC (pairwise sequentially Markovian coalescent) analysis of RAD (restriction site associated DNA) sequencing data. Preprint at <https://doi.org/10.5061/DRYAD.0618V> (2016).
77. Lanfear, R., Welch, J. J. & Bromham, L. Watching the clock: studying variation in rates of molecular evolution between species. *Trends Ecol. Evol.* **25**, 495–503 (2010).
78. Adams, D. C. & Otárola-Castillo, E. geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods Ecol. Evol.* **4**, 393–399 (2013).
79. Zhang, J. & Kumar, S. Detection of convergent and parallel evolution at the amino acid sequence level. *Mol. Biol. Evol.* **14**, 527–536 (1997).
80. Hadfield, J. D. & Nakagawa, S. General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *J. Evol. Biol.* **23**, 494–508 (2010).
81. Drury, J. P., Cowen, M. C. & Grether, G. F. Competition and hybridization drive interspecific territoriality in birds. *Proc. Natl Acad. Sci. USA.* **117**, 12923–12930 (2020).

82. Kowalczyk, A., Chikina, M. & Clark, N. L. A cautionary tale on proper use of branch-site models to detect convergent positive selection. *bioRxiv* 2021.10.26.465984 <https://doi.org/10.1101/2021.10.26.465984> (2021).
83. Smith, M. D. et al. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol. Biol. Evol.* **32**, 1342–1353 (2015).
84. Fukushima, K. & Pollock, D. D. Detecting macroevolutionary genotype-phenotype associations using error-corrected rates of protein convergence. *Nat. Ecol. Evol.* **7**, 155–170 (2023).
85. Ho, L. S. T. & Ané, C. A linear-time algorithm for gaussian and non-gaussian trait evolution models. *Syst. Biol.* **63**, 397–408 (2014).
86. Seki, R. et al. Functional roles of Aves class-specific cis-regulatory elements on macroevolution of bird-specific features. *Nat. Commun.* **8**, 14229 (2017).
87. Hinrichs, A. S. et al. The UCSC genome browser database: update 2006. *Nucleic Acids Res.* **34**, D590–D598 (2006).
88. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
89. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
90. Kowalczyk, A. et al. RERconverge: an R package for associating evolutionary rates with convergent traits. *Bioinformatics* **35**, 4815–4817 (2019).
91. Saputra, E., Kowalczyk, A., Cusick, L., Clark, N. & Chikina, M. Phylogenetic permutations: a statistically rigorous approach to measure confidence in associations in a phylogenetic context. *Mol. Biol. Evol.* **38**, 3004–3021 (2021).
92. Eliason, C. M. et al. Data from: Genomic signatures of convergent shifts to plunge-diving behavior in birds. <https://doi.org/10.5061/dryad.gflvhmvn> (2023).
93. Eliason, C. M. et al. Genomic data from: Genomic signatures of convergent shifts to plunge-diving behavior in birds. <https://doi.org/10.5281/zenodo.7872534> (2023).
94. Shumate, A. & Salzberg, S. L. Liftoff: accurate mapping of gene annotations. *Bioinformatics* **37**, 1639–1643 (2020).

Acknowledgements

We thank Isabel Distefano for helping with DNA preps. Funding for genome sequencing was provided by Iridian Genomes, grant# IRGEN_RG_2021-1345 Genomic Studies of Eukaryotic Taxa. This work was partially supported by the National Science Foundation (NSF EP 2112468 to C.M.E. and S.J.H., NSF EP 2112467 and DEB 1557051 to M.J.A.).

Author contributions

C.M.E. designed research; C.M.E., L.E.M., T.H. and S.P. performed research; C.M.E., L.E.M. and T.H. analyzed data; C.M.E., L.E.M., T.H., J.M.M., S.P., M.J.A. and S.J.H. wrote the initial manuscript; C.M.E., J.M.M., T.H. and M.J.A. revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-023-05359-z>.

Correspondence and requests for materials should be addressed to Chad M. Eliason.

Peer review information *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editor: George Inglis. A peer review file is available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023