Speciation with gene flow across an elevational gradient in New Guinea kingfishers

Ethan Linck[1]*, Benjamin G. Freeman[2], John P. Dumbacher[3]

[1]*Department of Biology and Burke Museum of Natural History & Culture, U. Washington*
[2]*Biodiversity Research Center and Department of Zoology, U. British Columbia*
[3]*Ornithology & Mammalogy, California Academy of Sciences*

*Corresponding Author: elinck@uw.edu

Abstract:

The role of environmental gradients in speciation remains contentious. Theory suggests speciation without geographic isolation is plausible if divergent natural selection is strong enough to counteract gene flow, but empirical examples are elusive. Tropical mountains provide strong and temporally stable environmental gradients that can promote local adaptation and population genetic structure. Pairs of closely related species with adjacent but divergent elevational ranges are common in these environments, a pattern that suggests parapatric speciation (i.e. speciation with moderate gene flow), but have rarely been studied in a speciation genomics framework. Here we use genomic data from modern and historical museum specimens to provide evidence for speciation with gene flow across an elevational gradient in a pair of New Guinea kingfishers. We find that the lowland species *Syma torotoro* and montane species *S. megarhyncha* form discrete genotypic clusters with bimodal variance in phenotypic traits. Despite this, demographic inference, *D*-statistics, and phylogenetic networks indicate range-wide gene flow over long time periods, with divergence concentrated in small regions of the genome shaped by positive selection. We propose these data can be explained by a "magic trait" model of ecological speciation where selection on body size also affects mate choice. Our results provide a rare validation of theoretical models of adaptive speciation and are consistent with a strong influence of tropical thermal stability on diversification. We suggest the primary role of selection across elevational gradients is the maintenance of species boundaries in the face of incomplete reproductive isolation, a mechanism for generating high tropical biodiversity.

Significance:

Speciation without geographic isolation can occur across environmental gradients when disruptive natural selection reduces effective migration between locally adapted populations. Theory suggests this process is widely plausible in a range of circumstances, but empirical examples remain rare. Strong and temporally stable thermal stratification in tropical mountains can select for narrow thermal tolerances and reduce dispersal across elevations, potentially promoting this process and contributing to globally high levels of biodiversity in these environments. We provide multiple lines of evidence for speciation across an elevational gradient in a pair of New Guinea kingfisher species that segregate by elevation, maintain species limits in the face of high levels of long-term gene flow across their range, and differ primarily in small regions of the genome shaped by positive selection. These results indicate selection across elevational gradients can maintain species boundaries in the face of incomplete reproductive isolation, a mechanism generating high tropical biodiversity.

Keywords: Syma, *speciation genomics, natural selection, tropical biodiversity, Ernst Mayr*

Introduction:

Adaptation across environmental gradients is ubiquitous in nature (1–3), but its role in promoting speciation remains contentious (4–7). Disruptive natural selection can lead to local adaptation that restricts gene flow between populations in different environments if it becomes paired with a mechanism to promote nonrandom mating such as a pleiotropic "magic trait," linkage disequilibrium between separate loci involved with local adaptation and assortative mating, or habitat preference leading to ecogeographic isolation (8–10). A robust body of theory suggests this process is possible under a range of circumstances (11–16). However, many evolutionary biologists dismiss its relevance in natural systems as it is difficult to achieve compared to models of speciation without gene flow and because there are few obvious empirical examples (4). Despite a recent shift from the traditional emphasis on the geography of diverging populations to a focus on measuring relative rates of gene flow (6, 7, 17), intermediate values of migration only imprecisely map on to conceptual models. Perhaps as a result, the most compelling studies of speciation without geographic isolation have focused on a single, relatively compact deme (18–20). Yet if speciation with some degree of gene flow is common, as Nosil has argued (21), it is likely parapatric divergence that predominates. Testing for its occurrence in a diverse range taxa will help clarify when it is likely to occur and spur the development of more accurate models.

Both the emphasis on the geography of speciation and skepticism towards speciation with gene flow were deeply seeded by the writings German-American zoologist Ernst Mayr (22–25), a prominent architect of the Modern Synthesis. Mayr's theory of 'geographic' (or allopatric) speciation was rooted in his detailed study of the birds of the island of New Guinea and nearby islands Northern Melanesia (22, 23). Mayr observed that putative sister species or geographically differentiated races nearly always had non-overlapping geographic ranges on separate islands, isolated mountains, or in lowland basins separated by mountain ranges. Consequently, he suggested geographic differentiation in isolation was nearly always a better explanation for speciation than divergence in sympatry, in part because it allowed Bateson-Dobzhansky-Muller Incompatibilities (26) or other isolating factors to accumulate in the absence of the homogenizing influence of gene flow. Initially implicating drift as the primary evolutionary force driving speciation, his views later shifted to emphasize founder effects (27) while conceding a role for natural selection (25).

Recent theory and empirical work has undermined the generality of Mayr's intuitive central assumption, which held that contemporary distribution of species can be used to infer their arrangement at the time of divergence. Models of speciation that invoke biotic interactions between diverging populations suggest it is often easier to achieve reproductive isolation in the presence of competition and an environmental gradient than in strict allopatry (14). Phylogenetic comparative approaches have been applied to evaluate the frequency of alternate modes of speciation, but are hampered by the lability of species ranges over evolutionary time scales and their limited information on underlying mechanisms (28). Demographic modeling using genomic data from multiple young species pairs has similarly found that the arrangement of current species ranges fails to predict inferred rates of ancestral gene flow (29). Moreover, genome-wide DNA sequence data has profoundly altered our understanding of the speciation process. We now understand speciation is far more dynamic than previously assumed: reticulation of lineages is common across the tree of life and may be a generative force for adaptive variation or even speciation itself, while species limits can be maintained by selection in the face of significant gene flow (30–33).

Ironically, Melanesia, the very region that was so important to Mayr's theory and ecology and evolution more broadly (34–36), has seen comparatively little modern research in the emerging dis-

cipline of speciation genomics (but see 37), in part due to logistical difficulties facing fieldworkers. While allopatric divergence following dispersal or vicariance is doubtless an important factor in the origin of species in tropical areas like New Guinea (38), the importance of abiotic and biotic ecological variables in driving evolutionary processes agnostic to geography is increasingly recognized (39). Ecologists have long known that reduced seasonal variation in temperature drives strong thermal stratification across tropical mountainsides (39–41), a pattern correlated with high beta diversity (42, 43), especially among closely related species (44, 45). Because many tropical taxa are residents that appear to show low rates of dispersal and have a "slow pace of life" (46–48), strong disruptive selection could plausibly counteract migration to drive adaptive divergence. Local adaptation across tropical elevational gradients has been documented by clines in functional genes and intraspecific population genetic structure (3, 49–51). This has been shown theoretically to "scale up" to speciation under a scenario of niche expansion (16), but empirical evidence remains mixed. Phylogenetic and phylogeographic comparative studies suggest this process has occurred in *Ithioma* butterflies (52) and Andean amphibians and reptiles (53). In Andean birds, a broad consensus holds that elevational replacements primarily form through divergence in allopatry followed by secondary contact and displacement (43, 49, 54, 55), despite equivocal results in some tests. Yet we are aware of only one study that has used population genomics to evaluate speciation across elevational gradients, which found strong evidence that adaptation to altitude drives speciation in *Senecio* ragwort plants in temperate Italy (56, 57). The question of the relative contribution of adaptation in the presence of gene flow across elevational gradients to speciation and broader patterns of tropical biodiversity remains open.

The Yellow-billed and Mountain Kingfishers *Syma torotoro* and *S. megarhyncha* (Aves: Alcedinidae) are putative sister taxa that segregate by elevation and vary only subtly in phenotype (22, 58, 59). The lowland species *S. torotoro* is reportedly smaller with a higher-pitched call, and primarily found below 700 m, above which it is replaced by the slightly larger, deeper-voiced *S. megarhyncha* (58). Among New Guinea's many elevational series, the distribution and morphological conservatism of Syma has led systematists to suggest speciation driven by colonization of montane forest (60), though Mayr discounted this possibility in New Guinea mountain birds generally (22). However, species limits and range-wide variation have never been quantitatively assessed, and observed differences might instead reflect phenotypic plasticity or clinal variation of a single widespread lineage (54).

Here we use genome-wide DNA sequences, bioacoustic data, and morphometric analyses to ask whether parapatric speciation across an elevational gradient occured in *Syma* kingfishers. Under this hypothesis (**Figure 1A**), we predicted 1) discrete genotypic and phenotypic clusters supporting a two-species model, rather than clinal variation; 2) evidence of extensive historical and contemporary gene flow resulting from long-term interbreeding at range margins; and 3) evidence of positive selection shaping highly divergent regions of the genome, independent of structural features that can inflate metrics of relative differentiation (61). If speciation occurred in allopatry by an alternate mechanism and current elevational distributions are the result of secondary contact and range displacement (**Figure 1B**), we predicted 1) discrete genotypic and phenotypic clusters; 2) limited gene flow restricted to recent timescales (as secondary contact prior to adaptive differentiation would collapse gene pools into a single population and postzygotic reproductive isolation would restrict the extent of introgression); and 3) no evidence of positive selection in highly divergent regions of the genome.

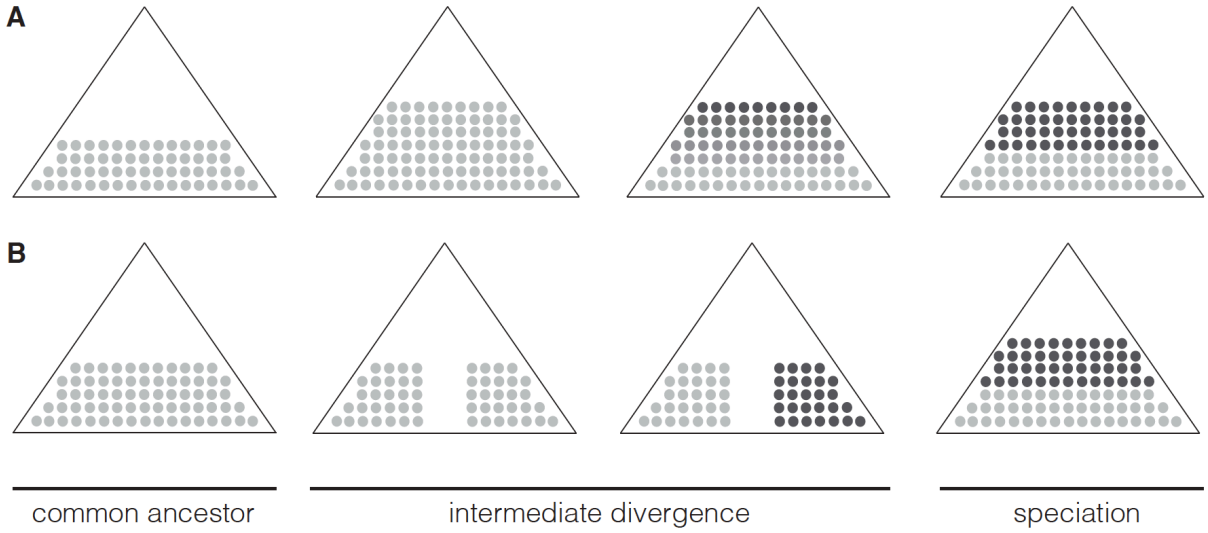

Figure 1: **Simplified hypotheses for the origin of elevational elevational series.** A) Range expansion across an elevational gradient exposes a population to disruptive natural selection, leading to local adaptation that restricts migration across the gradient until reproductive isolation develops as a byproduct. B) A single population splits following a dispersal event or vicariance and develops reproductive isolation in allopatry. Following range expansion and secondary contact, niche displacement leads to separate elevational ranges.

Results:

DNA sequencing from historical specimens. We extracted DNA and generated genome-wide sequence data from all historic toepad samples across both species' relatively inaccessible distributions, with collection dates ranging from 1896 to 1973 and including 3 individuals collected by Ernst Mayr himself in 1929. We present a detailed description of reduced representation hyRAD data elsewhere (62). On average we were able to align 83.3% of reads to the draft *Halcyon senegalensis* genome, ranging from 35.5% to 92.37% across individuals. Whole genome resequencing data had an average depth of coverage of 5.38x per sample, ranging from 1.92x to 12.12x. Following variant calling and filtering for depth of coverage and quality, a 95% complete data matrix including 37 samples and both WGS and hyRAD data had 66,917 SNPs, which was further reduced to 10,351 SNPs after trimming for linkage disequilibrium. A second matrix of whole genome resequencing data alone had 78,882,912 SNPs, which was reduced to 1,858,764 SNPs after filtering for a minimum depth of coverage of 3x and a minimum quality score of 30.

Species limits and phylogeny. Analysis of genome-wide DNA sequence data, morphometrics, and calls supported *Syma torotoro* and *S. megarhyncha* as distinct, assortatively mating lineages (**Figure 2**). Principal component analysis of genotypes identified elevation (and by extension current species limits) as the primary axis of genetic differentiation, with PC1 explaining 17.02% of total variance and separating *S. torotoro* and *S. megarhyncha* into discrete clusters (**Figure 2A**; **Figure 2B**) with little evidence of hybrid genotypes. PC1 explained 6.31% of total variance and separated insular subspecies *S. t. ochracea* into a third discrete cluster. The best-fit model from a K-means clustering analysis (as identified by BIC scores) perfectly recovered these groups (**Figure 2B**). Using assignments from this best-fit model as population priors, supervised inference

of population structure in `ADMIXTURE` (63) recovered identical clusters (log-likelihood=-104066.81; cross validation error: 0.5974) with no evidence of recent hybridization (**Figure S1**). Unsupervised inference of population structure for K=2 through K=5 assigned clusters that were discordant with both species identity and geography, with high cross-validation error (0.6822-1.113), suggesting the analysis was inappropriate due to uneven sampling (64) or small total sample size. Analyses of phenotypic data formed discrete clusters concordant with genomic results and consistent with morphological differentiation by elevation. Bill width and depth, tarsus, wing chord, and tail length were significantly larger in *Syma megarhyncha* after correcting for multiple comparisons in Welch's two sample T-tests (all comparisons $P<1\text{x}10^{-7}$), and species was a significant predictor of PC1 in a linear model ($P<1\text{x}10^{-14}$). Following principal component analysis, PC1 explained 81.47% of variance across all five traits. Bayesian Information Criterion (BIC) selected two distinct normal distributions out of normal mixture models (NMMs) fit to PC1 (Log Likelihood=-378.0914) (**Figure 2D**). Calls of *S. torotoro* had a significantly higher frequency ($P<1\text{x}10^{-5}$) but did not differ in duration. Species was a significant predictor $P<1\text{x}10^{-5}$) of PC1 in a principal component analysis of the 24 bioacoustic variables quantified in `warbleR` (65), and explained 35.72% of total variance. Because NMMs assume independence of observations, we did not perform formal model fitting for bioacoustic data, but a frequency distribution strongly suggests values are bimodally distributed by species identity (**Figure 2E**).

A maximum likelihood phylogeny (66) from near-complete mtDNA genomes further supported the distinctiveness of *S. megarhyncha* and *S. torotoro*, but unexpectedly recovered mainland *S. torotoro* as sister with a clade containing *S. megarhyncha* and insular endemic subspecies *S. t. ochracea* (**Figure 2F**). Within *S. torotoro*, a clade of four individuals with full bootstrap support from the southeast peninsula of Papua New Guinea was separated from the remaining individuals, but with little other apparent geographic structure across the tree. Because we expected the assumption of strict bifurcation underpinning many methods of phylogenetic inference to be explicitly violated by recent gene flow between species, we evaluated evolutionary relationships using resampled neighbor joining trees. Neighbor-joining trees are an empirical description of the distance matrix among samples, and their behavior under admixture has been described (67). We recovered the same four major clades as in our mtDNA phylogeny in a consensus neighbor joining tree using a 95% complete data matrix from all samples passing quality filters. Neighbor joining trees estimated from resampling a subset of these data with replacement indicated substantial discordance within species. Nodes in the consensus tree separating *S. torotoro* from *S. megarhyncha* and *S. t. ochracea*, and *S. megarhyncha* from *S. t. ochracea*, were present in all 500 resampled trees. The bipartition separating four *S. torotoro* individuals from the southeast peninsula from the remainder of the clade was present in 469/500 trees.

Demographic history and introgression: A formal test for introgression, a phylogenetic network analysis, and demographic inference indicate a long history of gene flow between *S. torotoro* and *S. megarhyncha*. We performed ABBA-BABA tests between *S. torotoro* and the three subspecies of *S. megarhyncha* in isolated mountain ranges, and between *S. t. ochracea* and these taxa. These tests can distinguish gene flow from incomplete lineage sorting in discordant genealogies (68). We found evidence of widespread introgression, with each tested pairwise comparison revealing significant median Z scores, even after correction for multiple comparisons (**Figure 3A**). Of these, the proportion of individual tests that were significant ranged from 0.74-1.0. Intriguingly, there was also evidence of introgression between *S. t. ochracea* and mainland *S. torotoro*, and *S. t. ochracea* and two out of three *S. megarhyncha* populations (**Figure 3A**). There was no evidence of gene flow between the far western population of *S. megarhyncha* and far eastern *S. t. ochracea*. These results were mirrored by a distance-based phylogenetic network analysis (**Figure 3B**), which showed clear

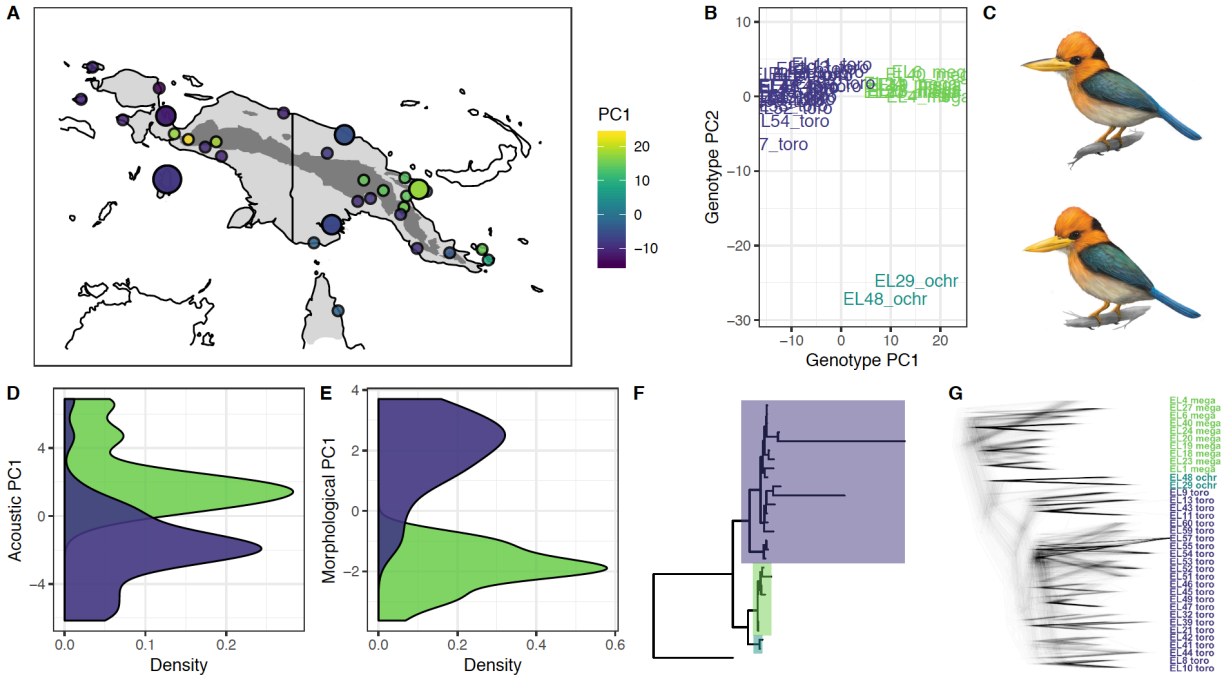

Figure 2: **Genomic and phenotypic data provide evidence of assortative mating.** A) Sampling localities for Syma kingfishers across New Guinea and Australia, color-coded by genotype PC1 and scaled by number of individuals. B) Principal component analysis of genotypes, color-coded by the best fit K-means clustering result (K=3). C) Illustration of *S. megarhyncha* (top) and *S. torotoro* (bottom), by Kevin Epperly. D) The first principal component of bioacoustic parameters measured from vocalizations is bimodally distributed by species. E) The first principal component of morphological data is bimodally distributed by species. F) A maximum likelihood phylogeny from near-complete mitochondrial genomes supports reciprocal monophyly of *S. torotoro* and a clade with *S. megarhyncha* and *S. t. ochracea.* Highlighted clades have full bootstrap support. G) A consensus neighbor-joining tree overlaid with trees bootstrapped from 10,000 loci is concordant with relationships revealed by mitochondrial data.

distinctions among the three lineages shaped by large reticulations, a pattern consistent with introgression over long time periods (69). We formally evaluated alternate demographic hypotheses by fitting the empirical joint site frequency spectrum (JSFS) to the its expected distribution using `moments` v. 1.0.0, which uses ordinary differential equations to model the evolution of allele frequencies (70). We tested models that shared growth in the most recent of two time periods following an initial divergence but differed in which of the two time periods migration was allowed in, if any (**Figure S2**). Across all parameter optimizations, a model of isolation with migration (IM; e.g. parapatric speciation with incomplete reproductive isolation) had the highest log-likelihood (-538.72), followed by a model of initial isolation with migration (IIM; parapatric speciation leading to strong reproductive isolation), with a maximum log-likelihood of -552.46, and then a model of secondary contact, with a maximum log-likelihood of -565.69 (**Figure 3C**). An allopatric speciation ("strict isolation"; SI) model was poorly supported (maximum log-likelihood -723.79). Both the wide range of likelihoods across runs and large standard deviations associated with parameter values for all models (**Table S1**) indicated difficulty reaching a global optima and / or poor fits for a more complex scenario involving selection. Parameter estimates varied widely across models, but broadly agreed in indicating greater ancestral than contemporary population sizes, larger populations in lowland species *S. torotoro*, and greater gene flow from *S. torotoro* to *S. megarhyncha* than the reverse (approximately 2-17 migrants per generation versus 0-2, respectively). Historical effective population size trajectories inferred in `SMC++` (71) were nearly identical in both species (**Figure 3D**), likely due to the confounding effects of gene flow, but agreed with results from moments in indicating large ancestral effective population sizes and a pronounced bottleneck at the last glacial maximum assuming a two-year generation time (72).

Natural selection and genomic divergence: Analysis of WGS data suggests *S. torotoro* and *S. megarhyncha* diverge primarily in small regions of the genome shaped by positive selection. Genome scans of divergence in 50 kb sliding windows were broadly consistent with expectations of speciation with gene flow driven by disruptive natural selection, revealing globally low to moderate genetic divergence (mean F_{ST}=0.0947) defined by 21 autosomal F_{ST} peaks (**Figure 4A**; **Figure S3A**). The Z chromosome showed low interspecific divergence, suggesting a limited role in speciation (73), though this may represent an artifact of high intraspecific diversity due to misincorporation of reads from the W chromosome in unsexed females. Correlations between summary statistics suggested a role for both structural reductions in recombination and positive selection in shaping F_{ST} peaks. F_{ST} was negatively but weakly correlated with D_{XY} (P<1x10^{-15}, R^2=0.1271), $\pi_{torotoro}$ (P<1x10^{-15}, R^2=0.2427), and $\pi_{megarhyncha}$ (P<1x10^{-15}, R^2=0.3785) (**Figure S3B**). D_{XY} was positively and strongly correlated with $\pi_{torotoro}$ (P<1x10^{-15}, R^2=0.9214) and $\pi_{megarhyncha}$ (P<1x10^{-15}, R^2=0.8495) (**Figure S3B**). F_{ST} was approximately exponentially distributed, while D_{XY}, $\pi_{torotoro}$, and $\pi_{megarhyncha}$ showed higher median values (**Figure S3C**). However, a method for detecting selective sweeps based on multiple signatures (74) indicated a strong role for positive selection in generating F_{ST} outliers. First, the value of the composite selective sweep summary statistic μ was significantly greater within peaks than without (Wilcox-Mann–Whitney U test; P<2x10^{-15}) (**Figure 4B**). Second, the total number of candidate selective sweeps within F_{ST} outlier regions was enriched relative to a random subset of nonoutlier windows, given a false positive rate of 0.05 (**Figure 4C**). Lastly, empirical values for $\pi_{megarhyncha}$ on chromosome 5 were significantly lower than expected under purifying selection alone (Wilcox-Mann–Whitney U test; P<2x10^{-15}), based on values generated by forward-time simulations (75) (**Figure 4D**).

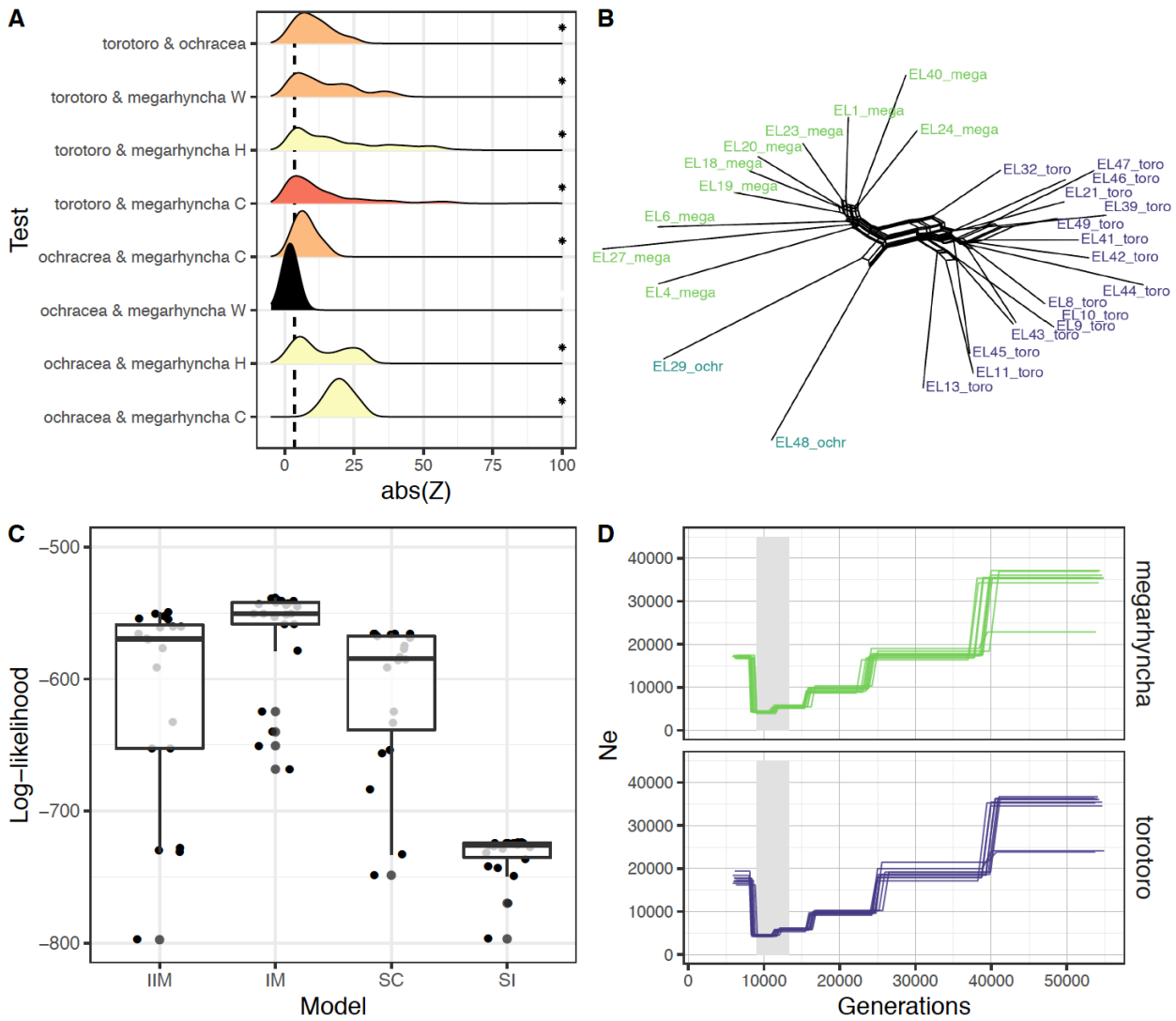

Figure 3: **Evidence for long-term introgression between species.** A) Median values for *Z*-scores from *D*-statistics ("ABBA-BABA tests") in pairwise comparisons between *S. torotoro* and allopatric *S. megarhyncha* and *S. t. ochracea* populations. Comparisons are color-coded by the proportion of individual combinations in each test that were significant, ranging from 0 to 1.0 and from black to off-white. B) A phylogenetic network supports previously inferred species limits but reveals extensive reticulation. C) Demographic inference from the joint site frequency spectrum supports a model of isolation with migration, but with broad uncertainty across replicate parameter optimizations. D) Trajectories of ancestral population sizes for indicate a bottleneck at the last glacial maximum, and are likely biased by gene flow.

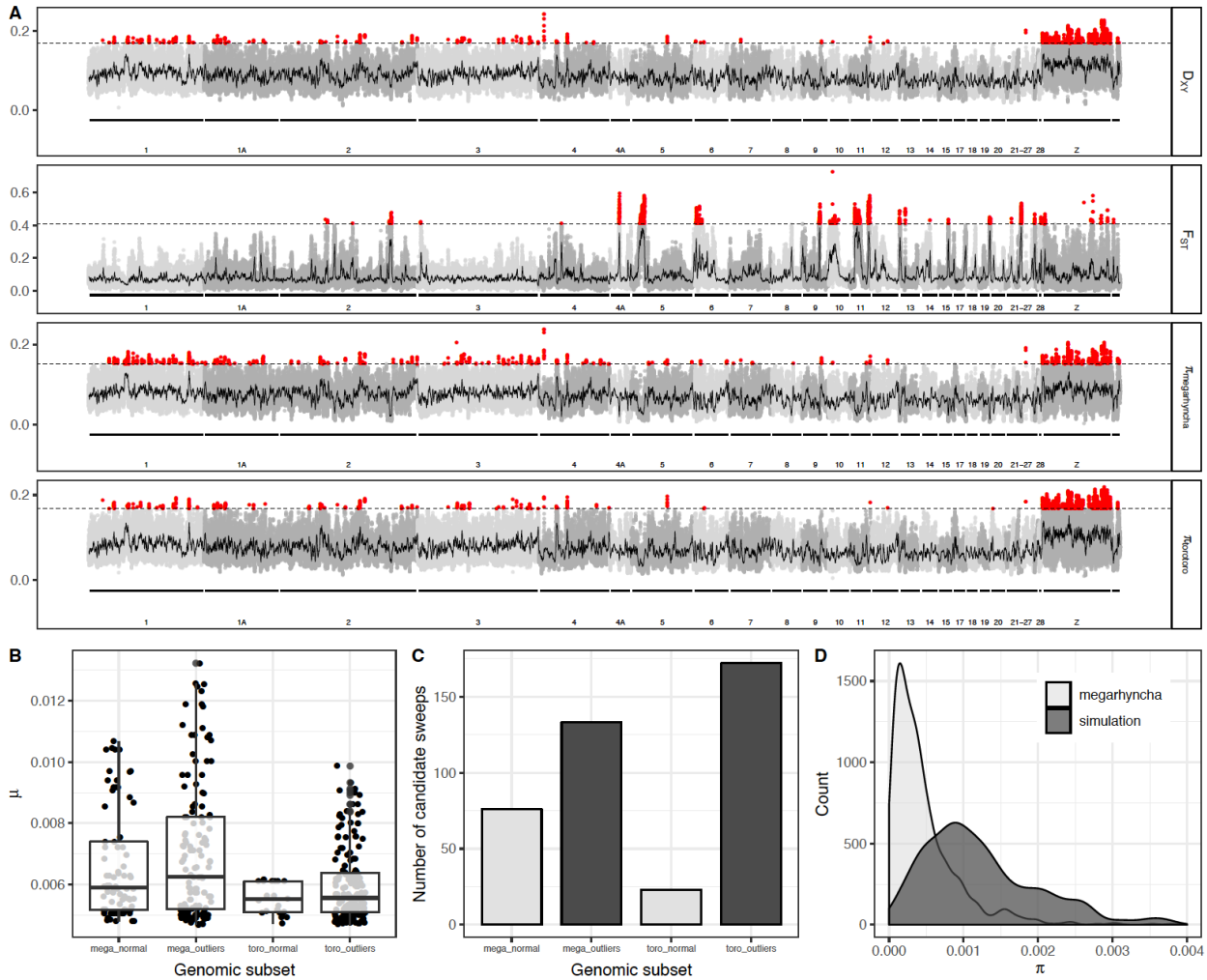

Figure 4: **Heterogeneous genomic divergence shaped by positive selection.** A) Relative divergence (F_{ST}), absolute divergence (D_{XY}), and intraspecific diversity ($\pi_{torotoro}$, and $\pi_{megarhyncha}$) calculated in 50 kb sliding windows across the genome. B) Significantly higher values of selective sweep summary statistic μ in F_{ST} peaks relative to a random sample of windows. C) F_{ST} peaks are enriched for the total number of candidate sweeps relative to non-outlier windows. D) Values of π on *S. megarhyncha* chromosome 5 are reduced below expectations from purifying selection alone.

Discussion:

Speciation across environmental gradients is commonly invoked in explanations of high tropical species richness and the latitudinal biodiversity gradient (54, 76–78). Tropical elevational gradients have high levels of temporally stable thermal stratification (40) that has been linked to selection for narrow thermal physiologies (41), reduced upslope or downslope dispersal (39), and high beta diversity (42). As a result, they have frequently been studied as a stage for this process (22, 43, 50, 53, 54, 79). Yet previous comparative phylogenetic studies have overwhelmingly supported models of divergence in allopatry followed by secondary contact and range displacement, particularly in vertebrates (53, 54, 79). Our finding of well-defined species limits (**Figure 2**) in the face of extensive historical and contemporary introgression (**Figure 3**) suggests that selection across elevational gradients is sufficient to maintain species boundaries with little evidence of strong postzygotic isolation. This result provides rare validation of an intuitive and widely cited but poorly buttressed theory: that selection across environmental gradients can serve as a motor for tropical diversification, either as the primary driver of speciation (e.g. in ecological speciation models) or as a force for reinforcement and disruptive selection after the build-up of partial reproductive isolation in allopatry.

We believe *Syma* provides a rare example of speciation with gene flow driven by niche expansion and disruptive selection across an elevational gradient (16), rather than allopatric speciation by another mechanism. Multiple lines of evidence are difficult to reconcile with an extensive period of isolation followed by secondary contact and niche displacement. Divergence in morphometric traits, the most visible phenotypic difference between *S. torotoro* and *S. megarhyncha* (**Figures 2D, 2E**), is largely thought to reflect ecological adaptation in allopatric populations (80, 81). While this may arise as a result of character displacement from competition in populations experiencing secondary contact (82), it presumes sufficient reproductive isolation has developed to permit coexistence without genetic homogenization and the fusion of lineages. In *Syma*, *D*-statistics in replicated pairwise comparisons across its full range (**Figure 3A**), demographic inference (**Figure 3C**), and a phylogenetic network analysis (**Figure 2B**) refute this with evidence for extensive gene flow between species. Further, several lines of evidence suggest gene flow has occurred since nearly their initial divergence. Extensive reticulation in a phylogenetic network paired with an apparent absence of individuals with recent hybrid ancestry in our dataset is consistent with low rates of hybridization over long time periods (69), as expected if interspecies mating occurs in the limited area of contact between parapatric ranges (**Figure 3B**). Similarly, support for an isolation-with-migration model (**Figure 3C**) over a model of secondary contact indicates historical as well as contemporary gene flow. Finally, significant introgression between mainland *S. torotoro* and insular endemic *S. t. ochracea* can be most easily explained by the sorting of introgressed loci following its split with *S. megarhyncha*, which would indicate gene flow between mainland lineages.

However, we emphasize that for most cases, parapatric speciation (and its vaguer, spatially agnostic variant "speciation with gene flow") are overly simplified models of what is likely a highly dynamic process over evolutionary time scales. Nor are genomic data ever likely to completely eliminate the possibility of a period of allopatry in the history of diverging lineages (83, 84). A more likely scenario with similar signatures might consist of numerous interdigitated periods of gene flow and isolation, driven by fluctuating local population sizes or by Pleistocene glacial cycling and its compression of elevational zones throughout New Guinea (85). Indeed, recent surveys on Mt. Wilhelm and Mt. Karimui in Papua New Guinea found a large gaps in the elevational ranges of *S. torotoro* and *S. megarhyncha* (45, 86), in contrast to earlier studies (59) suggesting that populations are effectively

allopatric at short horizontal distances. We suspect periodic gene flow is difficult to distinguish using data from site frequency spectra alone, but may be detected using other summary statistics, a process to be elaborated on elsewhere.

Though we did not identify the targets of selection, our data suggest disruptive positive selection has led to high divergence in small regions of the genome against a background of long-term introgression. These genomic "islands" have become a hallmark of the high throughput sequencing era and has been shown in wide range of nonmodel organisms (10, 87). Early interpretations of the pattern as a clear signature of speciation with gene flow–where F_{ST} peaks correlate to genes for traits under disruptive selection while low F_{ST} regions remain susceptible to introgression–have been complicated by evidence numerous processes can produce similar distributions, including selective sweeps in allopatry or recombination rate variation (61, 87–92). Using multiple signatures of selective sweeps paired with simulations to establish a false positive threshold, we found F_{ST} peaks were enriched for candidate sweeps in both species (**Figure 4B**, **Figure 4C**), with values of π on *S. megarhyncha* chromosome 5 (the site of one major peak) reduced below expectations under purifying selection alone. Evidence from multiple summary statistics across the same genomic windows can also help tease apart underlying mechanisms in some cases (88): for instance, low absolute differentiation (D_{XY}) in genomic islands can indicate F_{ST} values are inflated by low intraspecific diversity due to reduced recombination, casting doubt on models of divergence with gene flow (61). Here, our data defy easy description, highlighting the limits of this approach. A handful of major F_{ST} peaks stand out against a background of low F_{ST} windows, with a global average of 0.0947 (**Figure 4A**; **Figure S3A**). However, D_{XY} shows only a weak correlation with F_{ST} values (**Figure S3B**): low in some F_{ST} peaks and high in others, consistent with the likely hypothesis that recombination rate, selection against hybrid ancestry, disruptive selection, and gene flow have interacted to form a complex mosaic in the genomes of *S. torotoro* and *S. megarhyncha.*

As adaptive differentiation in morphology appears to have occurred during the initial stages of lineage divergence (either in the presence of gene flow or accompanied by it shortly afterwards), *Syma* is plausibly a case of ecological speciation *sensu* Nosil and Schluter (8, 10, 93). We propose this may have been facilitated by a so-called "magic trait", or trait under environmentally-mediated selection that also influences reproductive isolation (94). Magic trait models are among the easiest ways to achieve speciation with gene flow (10), and perhaps especially apply to birds, where intrinsic incompatibilities are rare and prezygotic isolation is thought to play a major role in speciation (95, 96). While we lack strong evidence for a target of disruptive selection or a specific mechanism to link it to reproductive isolation, we speculate selection on body size may also affect mate choice. Morphometric data from *Syma* are consistent with Bergmann's Rule and its prediction of larger body size in the higher elevation cooler climate species, *S. megarhyncha* (but see 97). As one of only three kingfisher species found above 2000 m in New Guinea and the only regional kingfisher that is a high elevation specialist, *Syma megarhyncha* is likely to have a thermal physiology under stronger selective constraint for warm environments, requiring greater adaptive divergence in body size to colonize montane environments. An increase in body size might directly lower frequency calls as a biproduct of morphological divergence, and assortative mating might act on either trait individually or in tandem (98-100).

What does *Syma* reveal about the origin of elevational series of congeners and by extension their significant contribution to tropical montane biodiversity (45)? First, our results suggest disruptive selection across elevational gradients can effectively maintain species limits in the absence of strong postzygotic reproductive isolation, a finding with broad relevance to speciation in tropical mountains regardless of the geographic mode of divergence or levels of gene flow. In a scenario of

secondary contact without adaptation to divergent elevational niches, *Syma* lineages would likely have fused. This suggests allopatric divergence in isolation may not be sufficient to drive many speciation events in tropical mountains. Second, aspects of *Syma's* natural history may serve as a guide to where to expect speciation with gene flow in birds: species that are vocal nonlearners, have a conserved climatic niche, and have a small realized range and total area of suitable habitat. While New Guinea is large and geologically complex, it is dwarfed by other tropical montane regions, which potentially have increased odds for speciation in allopatry (58, 101). Yet even if speciation across elevational gradients remain a rare exception among birds, the conditions that make it possible in *Syma* are more common in other taxa. Indeed, preliminary studies suggest it may be a much more common mechanism in amphibians, insects, and plants (50, 52, 53, 56). We highlight the importance of natural history studies of poorly known tropical organisms in establishing candidates for further investigation with genomic and experimental approaches.

Ernst Mayr's emphasis on geographic isolation profoundly shaped the study of speciation and diversification, establishing the primacy of divergence in allopatry and highlighting the significance of coexistence in sympatry to species concepts and speciation theory (24, 25). Though famously a skeptic of sympatric speciation and its relatives (23), he retained an appreciation of ecological factors in population divergence (25)–contrasting perspectives shaped by his foundational experience as an ornithologist, natural historian and systematist in New Guinea. Yet contemporary speciation research integrating genomic data with traditional analyses of phenotype and distributional data has only recently been applied to the New Guinea birds Mayr knew so well (52), and continues to be difficult to implement due to massive logistical hurdles. Much as the study of ancient DNA has revolutionized our understanding of human prehistory (88), widespread use of whole genome sequences from historic museum specimens has the potential to reshape our understanding of the speciation process in understudied tropical regions with few contemporary data and pressing conservation challenges. As this occurs, we expect much of Mayr's seminal work to gain new significance, as case studies like *Syma* highlight the continued relevance of the spatial organization of populations while enriching previously simple narratives of the role of selection and gene flow in the origin of species.

Materials and Methods:

Reproducibility and data availability. All code written for this study, all processed data matrices, and supplemental material can be found at https://github.com/elinck/syma_speciation/. Large data matrices are available via Dryad (pending), and raw sequence data is available via the NCBI Sequence Read Archive (pending).

Study system. Yellow-billed Kingfisher *Syma torotoro* and Mountain Kingfisher *Syma megarhyncha* (Aves: Alcedinidae) are the sole members of their genus. Tree kingfishers (subfamily Halcyoninae), they are endemic to New Guinea, its satellite islands, and the Cape York Peninsula of Australia. *S. torotoro* is found in tropical lowland forest and savannah from sea level to ~500 m elevation, or less commonly to ~1100 m (58). *S. megarhyncha* is found from ~600 m to 2700 m or higher (58). Though cited as a classic example of congeneric elevational replacements occurring in parapatry (59), their elevational ranges have also been reported to either overlap (102) or be separated by a substantial gap (45, 86). Both species are omnivorous territorial interior forest residents and differ only in *S. megarhyncha's* larger body size, deeper call, and the extent of black on the top of its bill in one subspecies (58). Insular *S. torotoro* subspecies *S. t. ochracea* differs substantially from its conspecifics and *S. megaryncha* in call, and is intermediate in size.

Morphological and bioacoustic data. We measured bill length, bill depth, tarsus, wing chord, and tail length from 72 museum specimens of *Syma torotoro* (n=40) and *S. megarhyncha* (n=32) at the American Museum of Natural History, representing all described subspecies (**Table S1**). Using these data, we performed principal component analyses in R (103) with normalized variables, and used PC1 to build mixture models using the `R` package `mclust` v. 5.4.1, which we evaluated with a maximum likelihood classification approach (104). We downloaded all available vocalizations from both species from xeno-canto and Cornell's Macaulay Library. We filtered these data for quality and quantified a suite of bioacoustic variables using the `warbleR` package v. 1.1.14 in `R` (65), and ran PCA with normalized variables on the output. We then fit these data to alternate species delimitation models using the same approach as with our morphological data.

Sampling, library preparation, and DNA sequencing. We extracted DNA from fresh tissues (n=6) and toepad samples from historical museum specimens (n=34) from 30 individuals of *S. torotoro* (n=30) and 10 individuals of *S. megarhyncha* (n=10). These samples represented all described subspecies and the full extent of both species' ranges in New Guinea and Australia (**Table S2**). We extracted DNA from fresh tissues using a Qiagen DNAeasy kit and the manufacturer's recommended protocol. For historical toepad samples (collected 1877-1977), we extracted DNA using either a using a phenol–chloroform and centrifugal dialysis method (105) (for reduced representation sequencing) or a standard salt extraction protocol (for whole genome resequencing). On a subset of samples (n=20), we performed reduced representation genome sequencing using a hybridization capture with RAD probes (hyRAD) approach, described in detail elsewhere (62). We sent the remaining samples to the UC Berkeley's Vincent J. Coates Genomic Sequencing Laboratory, where laboratory staff prepared genomic libraries for low coverage whole genome sequencing (WGS) using Illumina TruSeq Nano kits and a modified protocol that skipped sonication and enzymatically repaired fragments with RNAse. They then pooled (n=20) and sequenced these samples with 150 base pair paired-end reads on a single lane of an Illumina HiSeq 4000.

Sequence assembly and variant calling: We processed demultiplexed reads with a custom bioinformatic pipeline optimized for handling degraded DNA data and available at https://github.com/elinck/syma_speciation/. Briefly, we trimmed raw reads for adapters and low quality bases using `bbduk` from `BBTools` version 38.06 suite of bioinformatics tools. We aligned these reads to an unpublished draft genome of Woodland Kingfisher *Halcyon senegalensis* from the Bird 10K Genome Project using bbmap with a k-mer value of 12, a maximum indel length of 200 bp, and a minimum sequence identity of 0.65. We used `PicardTools` v. 2.17.8 and `GATK` v. 3.6.0 (106) to append read groups and perform local realignment on .bam files. We then used `mapDamage` 2.0.9 to account for postmortem damage to DNA from historical museum specimens by rescaling quality scores (107). We performed multisample variant calling using the `UnifiedGenotyper` tool in `GATK` v. 3.6.0, and filtered our variant calls for missing data, coverage, and quality with `VCFtools` 0.1.16 (108). To complement our nuclear DNA sequences, we assembled near-complete mitochondrial genomes from all samples using a baiting and iterative mapping approach implemented `MITObim` v 1.9.1. (109), which we for 10 cycles using a complete mtDNA genome from close relative *Todiramphus sanctus* as a reference (110).

Population structure inference: We evaluated population genetic structure within and between species using both nonparametric and model-based clustering approaches. We performed principal component analysis of genotypes (PCA) and identified putative genetic clusters for K=2 through K=4 using `adegenet` v. 2.1.1 and a 95% complete dataset with 66,917 SNPs from all samples passing quality filters and a minimum minor allele frequency of 0.05. We plotted axes PC1 and PC2 using `ggplot2` v. 3.1.0. After first determining small sample sizes made unsupervised in-

ference inappropriate for our data and resulted in extremely high cross validation error (64), we implemented model-based clustering using `ADMIXTURE` v. 1.3.0 (63) and population priors from the K=3 K-means clustering result.

Phylogenetic inference: We constructed a consensus neighbor joining tree in the `R` package `ape` v. 5.1 (111), implementing Sanderson's nonparametric rate smoothing with λ=1 using a 22,226 SNP dataset of all individuals passing quality filters with no more than 5% missing data per site, a minimum minor allele frequency of 0.05, and a minimum depth of coverage of 3x. To evaluate and visualize uncertainty, we generated 500 additional neighbor joining trees by resampling 10,000 SNPs from the matrix with replacement, and plotted them on the consensus tree using the `densiTree()` function in `ape`. We used the same underlying distance matrix to compute a phylogenetic network using the Neighbor-Net approach (112) implemented in `SplitsTree` (113) with default parameters. We inferred a maximum likelihood tree with an alignment of mtDNA genomes from the majority of samples using `RAxML` through the CIPRES portal (114) with default parameters.

Demographic inference: We calculated the JSFS using a SNP dataset with a minimum per site depth of 3x and a minimum quality of Q=30, thinned to 1 site for every 50 kbp to reduce the influence of linkage disequilibrium. We then defined four nested models which differed by level and timing of gene flow (Figure S1): isolation with migration (IM), isolation with initial migration (IIM), secondary contact (SC), and strict isolation (SI). All models featured a split of an ancestral population into two daughter lineages of arbitrary size, followed by a second time period allowing population growth. After initially optimizing parameters using the `optimize_log` method, we ran 25 additional optimizations, allowing a 1-fold random perturbation of parameter values each time. We selected the model with the maximum log-likelihood and estimated parameter uncertainty using 100 bootstrapped frequency spectra and the Godambe Information Matrix. We converted parameter values to real units using a genome-wide mutation rate of 2.3 x 10^{-9} (115), an effective sequence length scaled to reflect our LD-thinned SNP dataset, and a generation time estimate of two years, and generated parameter uncertainty estimates using a Godambe Information Matrix. To evaluate changes in historical effective population size through time in greater detail, we used `SMC++` (71), which accounts for the distribution of variation across the genome. We masked continuous stretches of homozygosity greater than 30 kb, included all contigs longer than 1 x 100 kbp, assumed a generation time of 2 years, and again assumed mutation rate of 2.3 x 10^{-9}. We visualized uncertainty by refitting models to 10 bootstrap replicates per population, resampling both samples and contigs with replacement.

Introgression: We explicitly tested for introgression between lineages using a four taxon test implemented in `ANGSD` with the `doAbbababa` function, described in detail elsewhere (68, 116). We calculated D for each combination of individuals conforming to the topology (mega,ochr),toro), and assessed significance by block jackknife over 1 mbp windows. We summarized results by collapsing individual tests representing the same geographic comparisons and calculating median Z scores and the proportion of individual tests that were significant given a cutoff of Z=3.59, which is equivalent to P=0.005 after performing a Bonferroni correction for multiple comparisons (n=31). These comparisons included *S. torotoro*, insular *S. t. ochracea* and three *S. megarhyncha* populations: *S. megarhyncha* in the mountains of the Huon Peninsula, *S. megarhyncha* in the eastern Central Ranges, and *S. megarhyncha* in the far western Central Ranges.

Genome scans and tests for positive selection: We assessed levels of divergence across the genome by calculating Wright's F_{ST}, D_{XY}, and π for each species in 50 kbp windows with a 10 kbp step size using scripts written by Simon Martin (https://github.com/simonhmartin/genomics_general), using a VCF file filtered for a minimum Q score of 30 and a minimum depth of 3x with a total

of 1,858,764 SNPs and 102,606,487 total sites. To assign chromosome identity to scaffolds and windows, we aligned the *H. senegalensis* draft genome to the chromosome-level genome assembly of *Taenopygia guttata* (117) using `NUCmer` in `MUMmer` v. 3.1 (118), allowing a maximum gap of 1 kbp and using a minimum sequence identity threshold of 10 kbp per contig. We evaluated correlations among summary statistics using simple linear models implemented in `R` (103). To formally test for the signature of positive selection in the genomes of both species, we used the program `RAiSD` (74). We calculated Wright's F_{ST}, D_{XY}, and π for each species in nonoverlapping 50 kbp windows, and used these statistics to identify a set of scaffolds hosting F_{ST} outlier windows at the P=0.005 threshold. We then created an identically sized set non-F_{ST} outlier windows using random sampling without replacement, and ran `RAiSD` on species-specific .vcf files including only these scaffolds. To establish a value of composite selective sweep summary statistic μ to use as a threshold for a false positive rate of 0.05, we used the program `SFS_CODE` (75) to simulate data similar to our own but under the influence of purifying selection, outputting a .vcf file to analyze with `RAiSD` using the `-k 0.05` option.

Acknowledgements:

We thank generations of New Guinean field assistants–traditional landowners and stewards of the island's birdlife–without whose help this paper would have been impossible to write. For tissue loans, we thank L. Josef at ANWC, K. Zykowski at YPM, R. Moyle and M. Robbins at KUMNH, P. Sweet at AMNH, and R. Prys-Jones at NHMUK-Tring. For help arranging fieldwork, we thank A. Mack, G. Kapui, J. Robbins, N. Gowep, B. Beehler, L. Dabek, N. Whitmore, F. Dem, S. Tulai, and V. Novotny. We thank Alex Wiebe for assistance measuring specimens, and thank CJ Battey for many discussions, help with analyses, and inspiration for data visualization. This work was supported by NSF Doctoral Dissertation Improvement Grant #1701224 and an NDSEG Fellowship to E.B.L., and by NSF DEB #0108247 to J.P.D.

References:

1. Turesson G (1925) The Plant Species in Relation to Habitat and Climate. Hereditas 6(2):147–236.
2. McCormack John E, Smith Thomas B (2008) Niche expansion leads to small-scale adaptive divergence along an elevation gradient in a medium-sized passerine bird. Proc R Soc B Biol Sci 275(1647):2155–2164.
3. Cheviron ZA, Brumfield RT (2009) Migration-Selection Balance and Local Adaptation of Mitochondrial Haplotypes in Rufous-Collared Sparrows (zonotrichia Capensis) Along an Elevational Gradient. Evolution 63(6):1593–1605.
4. Coyne JA, Orr HA (2004) Speciation (Sinauer).
5. Mallet J (2005) Speciation in the 21st century. Heredity 95:105–109.
6. Fitzpatrick BM, Fordyce JA, Gavrilets S (2008) What, if anything, is sympatric speciation? J Evol Biol 21(6):1452–1459.
7. Mallet J (2008) Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. Philos Trans R Soc B Biol Sci 363(1506):2971–2986.
8. Rundle HD, Nosil P (2005) Ecological speciation. Ecol Lett 8(3):336–352.
9. Via S (2009) Natural selection in action during speciation. Proc Natl Acad Sci 106(Supplement 1):9939–9946.

10. Nosil P (2012) Ecological Speciation (Oxford University Press, Oxford, New York).
11. Smith JM (1966) Sympatric Speciation. Am Nat 100(916):637–650.
12. Endler JA (1977) Geographic Variation, Speciation, and Clines (Princeton University Press).
13. Kirkpatrick M, Ravigné V (2002) Speciation by Natural and Sexual Selection: Models and Experiments. Am Nat 159(S3):S22–S35.
14. Doebeli M, Dieckmann U (2003) Speciation along environmental gradients. Nature 421(6920):259–264.
15. Doebeli M, Dieckmann U, Metz JAJ, Tautz D (2005) What We Have Also Learned: Adaptive Spectiation Is Theoretically Plausible. Evolution 59(3):691–695.
16. Hua Xia (2016) The impact of seasonality on niche breadth, distribution range and species richness: a theoretical exploration of Janzen's hypothesis. Proc R Soc B Biol Sci 283(1835):20160349.
17. Butlin Roger K, Galindo Juan, Grahame John W (2008) Sympatric, parapatric or allopatric: the most important way to classify speciation? Philos Trans R Soc B Biol Sci 363(1506):2997–3007.
18. Schliewen UK, Tautz D, Pääbo S (1994) Sympatric speciation suggested by monophyly of crater lake cichlids. Nature 368(6472):629–632.
19. Savolainen V, et al. (2006) Sympatric speciation in palms on an oceanic island. Nature 441(7090):210–213.
20. Ryan PG, Bloomer P, Moloney CL, Grant TJ, Delport W (2007) Ecological Speciation in South Atlantic Island Finches. Science 315(5817):1420–1423.
21. Nosil P (2008) Speciation with gene flow could be common. Mol Ecol 17(9):2103–2106.
22. Mayr E (1942) Systematics and the Origin of Species, from the Viewpoint of a Zoologist (Harvard University Press).
23. Mayr E (1963) Animal species and evolution (Harvard University Press).
24. Coyne JA (1994) ERNST MAYR AND THE ORIGIN OF SPECIES. Evolution 48(1):19–30.
25. Nosil P (2008) Ernst Mayr and the integration of geographic and ecological factors in speciation. Biol J Linn Soc 95(1):26–46.
26. Dobzhansky T (1937) Genetics and the Origin of Species (Columbia University Press).
27. Provine WB (2004) Ernst Mayr: Genetics and Speciation. Genetics 167(3):1041–1046.
28. Losos JB, Glor RE (2003) Phylogenetic comparative methods and the geography of speciation. Trends Ecol Evol 18(5):220–227.
29. Penalba JV, Joseph L, Moritz C (2018) Current geography masks dynamic history of gene flow during speciation in northern Australian birds. Mol Ecol 28(3):630-643.
30. Brelsford A, Milá B, Irwin DE (2011) Hybrid origin of Audubon's warbler. Mol Ecol 20(11):2380–2389.
31. Kumar V, et al. (2017) The evolutionary history of bears is characterized by gene flow across species. Sci Rep 7:46487.
32. Schumer M, et al. (2018) Natural selection interacts with recombination to shape the evolution of hybrid genomes. Science:eaar3684.
33. Edelman NB, et al. (2018) Genomic architecture and introgression shape a butterfly radiation. bioRxiv:466292.
34. Wilson EO (1959) Some Ecological Characteristics of Ants in New Guinea Rain Forests. Ecology 40(3):437–447.
35. Diamond JM (1973) Distributional Ecology of New Guinea Birds: Recent ecological and biogeographical theories can be tested on the bird communities of New Guinea. Science 179(4075):759–769.
36. Diamond JM (1972) Biogeographic Kinetics: Estimation of Relaxation Times for Avifaunas

of Southwest Pacific Islands. Proc Natl Acad Sci 69(11):3199–3203.
37. Stryjewski KF, Sorenson MD (2017) Mosaic genome evolution in a recent and rapid avian radiation. Nat Ecol Evol 1(12):1912.
38. Smith BT, et al. (2014) The drivers of tropical speciation. Nature 515(7527):406–409.
39. Polato NR, et al. (2018) Narrow thermal tolerance and low dispersal drive higher speciation in tropical mountains. Proc Natl Acad Sci:201809326.
40. Janzen DH (1967) Why Mountain Passes are Higher in the Tropics. Am Nat 101(919):233–249.
41. Sheldon KS, Huey RB, Kaspari M, Sanders NJ (2018) Fifty Years of Mountain Passes: A Perspective on Dan Janzen's Classic Article. Am Nat 191(5):553–565.
42. Jankowski JE, Ciecka AL, Meyer NY, Rabenold KN (2009) Beta diversity along environmental gradients: implications of habitat specialization in tropical montane landscapes. J Anim Ecol 78(2):315–327.
43. Cadena Carlos Daniel, et al. (2012) Latitude, elevational climatic zonation and speciation in New World vertebrates. Proc R Soc B Biol Sci 279(1726):194–201.
44. Terborgh J, Weske JS (1975) The Role of Competition in the Distribution of Andean Birds. Ecology 56(3):562–576.
45. Freeman BG, Freeman AMC (2014) The avifauna of Mt. Karimui, Chimbu Province, Papua New Guinea, including evidence for long-term population dynamics in undisturbed tropical forest. Bull Br Ornithol Club 134(1):30–51.
46. Wiersma P, Muñoz-Garcia A, Walker A, Williams JB (2007) Tropical birds have a slow pace of life. Proc Natl Acad Sci 104(22):9340–9345.
47. Harvey MG, Aleixo A, Ribas CC, Brumfield RT (2017) Habitat Association Predicts Genetic Diversity and Population Divergence in Amazonian Birds. Am Nat 190(5):631–648.
48. Smith BT, Seeholzer GF, Harvey MG, Cuervo AM, Brumfield RT (2017) A latitudinal phylogeographic diversity gradient in birds. PLOS Biol 15(4):e2001073.
49. DuBay SG, Witt CC (2014) Differential high-altitude adaptation and restricted gene flow across a mid-elevation hybrid zone in Andean tit-tyrant flycatchers. Mol Ecol 23(14):3551–3565.
50. Funk WC, et al. (2016) Elevational speciation in action? Restricted gene flow associated with adaptive divergence across an altitudinal gradient. J Evol Biol 29(2):241–252.
51. Gadek CR, et al. (2018) Why are tropical mountain passes "low" for some species? Genetic and stable-isotope tests for differentiation, migration and expansion in elevational generalist songbirds. J Anim Ecol 87(3):741–753.
52. Elias M, et al. (2009) Out of the Andes: patterns of diversification in clearwing butterflies. Mol Ecol 18(8):1716–1729.
53. Arteaga A, et al. (2016) Comparative Phylogeography Reveals Cryptic Diversity and Repeated Patterns of Cladogenesis for Amphibians and Reptiles in Northwestern Ecuador. PLOS ONE 11(4):e0151746.
54. Caro LM, Caycedo-Rosales PC, Bowie RCK, Slabbekoorn H, Cadena CD (2013) Ecological speciation along an elevational gradient in a tropical passerine bird? J Evol Biol 26(2):357–374.
55. Cadena CD, Céspedes LN (In press) Origin of elevational replacements in a clade of nearly flightless birds – most diversity in tropical mountains accumulates via secondary contact following allopatric speciation.
56. Chapman MA, Hiscock SJ, Filatov DA (2013) Genomic Divergence during Speciation Driven by Adaptation to Altitude. Mol Biol Evol 30(12):2553–2567.
57. Osborne OG, Batstone TE, Hiscock SJ, Filatov DA (2013) Rapid Speciation with Gene Flow

Following the Formation of Mt. Etna. Genome Biol Evol 5(9):1704–1715.
58. Pratt TK, et al. (2015) Birds of New Guinea: Second Edition (Princeton University Press). REV-Revised, 2 Available at: https://www.jstor.org/stable/j.ctt7zvd34 [Accessed December 8, 2018].
59. Diamond JM (1972) Avifauna of the Eastern Highlands of New Guinea (Nuttall Ornithological Club).
60. Rand AL, Guinea (1933-1934) AE to N, History (1920-1941) WSSE of the AM of N (1936) Altitudinal variation in New Guinea birds. American Museum novitates ; no. 890. Available at: http://digitallibrary.amnh.org/handle/2246/4107 [Accessed December 8, 2018].
61. Cruickshank TE, Hahn MW (2014) Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. Mol Ecol 23(13):3133–3157.
62. Linck EB, Hanna ZR, Sellas A, Dumbacher JP (2017) Evaluating hybridization capture with RAD probes as a tool for museum genomics with historical bird specimens. Ecol Evol 7(13):4755–4767.
63. Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. Genome Res. doi:10.1101/gr.094052.109.
64. Puechmaille SJ (2016) The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. Mol Ecol Resour 16(3):608–627.
65. Araya-Salas M, Smith-Vidaurre G (2017) warbleR: an r package to streamline analysis of animal acoustic signals. Methods Ecol Evol 8(2):184–191.
66. Stamakis A (2014) RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. Bioinformatics 30(9):1312-1313
67. Kopelman NM, Stone L, Gascuel O, Rosenberg NA (2013) The behavior of admixed populations in neighbor-joining inference of population trees. Pac Symp Biocomput:273–284.
68. Green RE, et al. (2010) A Draft Sequence of the Neandertal Genome. Science 328(5979):710–722.
69. Martin SH, Davey JW, Salazar C, Jiggins CD (2019) Recombination rate variation shapes barriers to introgression across butterfly genomes. PLOS Biol 17(2):e2006288.
70. Jouganous J, Long W, Ragsdale AP, Gravel S (2017) Inferring the Joint Demographic History of Multiple Populations: Beyond the Diffusion Approximation. Genetics 206(3):1549–1567
71. Terhorst J, Kamm JA, Song YS (2017) Robust and scalable inference of population history from hundreds of unphased whole-genomes. Nat Genet 49(2):303–309.
72. Clark PU, et al. (2009) The Last Glacial Maximum. Science 325(5941):710–714.
73. Irwin DE (2018) Sex chromosomes and speciation in birds and other ZW systems. Mol Ecol 27(19):3831–3851.
74. Alachiotis N, Pavlidis P (2018) RAiSD detects positive selection based on multiple signatures of a selective sweep and SNP vectors. Commun Biol 1(1):79.
75. Hernandez RD (2008) A flexible forward simulator for populations subject to selection and demography. Bioinforma Oxf Engl 24(23):2786–2787.
76. Schneider CJ, Smith TB, Larison B, Moritz C (1999) A test of alternative models of diversification in tropical rainforests: Ecological gradients vs. rainforest refugia. Proc Natl Acad Sci 96(24):13869–13873.
77. Portik DM, et al. (2017) Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection. Mol Ecol 26(19):5245–5263.
78. Beheregaray LB, Cooke GM, Chao NL, Landguth EL (2015) Ecological speciation in the tropics: insights from comparative genetic studies in Amazonia. Front Genet 5. doi:10.3389/fgene.2014.00477.

79. Moyle RG, et al. (2017) A genome-wide assessment of stages of elevational parapatry in Bornean passerine birds reveals no introgression: implications for processes and patterns of speciation. PeerJ 5. doi:10.7717/peerj.3335.
80. Ricklefs RE, Bermingham E (2007) The Causes of Evolutionary Radiations in Archipelagoes: Passerine Birds in the Lesser Antilles. Am Nat 169(3):285–297.
81. Winger BM, Bates JM (2015) The tempo of trait divergence in geographic isolation: Avian speciation across the Marañon Valley of Peru. Evolution 69(3):772–787.
82. Diamond J, Pimm SL, Gilpin ME, LeCroy M (1989) Rapid Evolution of Character Displacement in Myzomelid Honeyeaters. Am Nat 134(5):675–708.
83. Strasburg JL, Rieseberg LH (2011) Interpreting the estimated timing of migration events between hybridizing species. Mol Ecol 20(11):2353–2366.
84. Martin SH, et al. (2013) Genome-wide evidence for speciation with gene flow in Heliconius butterflies. Genome Res:gr.159426.113.
85. Hope G (2014) The Sensitivity of the High Mountain Ecosystems of New Guinea to Climatic Change and Anthropogenic Impact. Arct Antarct Alp Res 46(4):777–786.
86. Sam K, Koane B, Novotny V (2014) New avian records along the elevational gradient of Mt. Wilhelm, Papua New Guinea. Bull Br Ornithol Club 134(2):116–133.
87. Bay RA, Ruegg K (2017) Genomic islands of divergence or opportunities for introgression? Proc R Soc B 284(1850):20162414.
88. Irwin DE, et al. (2018) A comparison of genomic islands of differentiation across three young avian species pairs. Mol Ecol 27(23):4839–4855.
89. Turner TL, Hahn MW, Nuzhdin SV (2005) Genomic Islands of Speciation in Anopheles gambiae. PLOS Biol 3(9):e285.
90. Nosil P, Feder JL (2012) Genomic divergence during speciation: causes and consequences. Philos Trans R Soc B Biol Sci 367(1587):332–342.
91. Via S (2012) Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow. Philos Trans R Soc B Biol Sci 367(1587):451–460.
92. Burri R (2017) Interpreting differentiation landscapes in the light of long-term linked selection. Evol Lett 1(3):118–131.
93. Schluter D (2001) Ecology and the origin of species. Trends Ecol Evol 16(7):372–380.
94. Servedio MR, Doorn GSV, Kopp M, Frame AM, Nosil P (2011) Magic traits in speciation: 'magic' but not rare? Trends Ecol Evol 26(8):389–397.
95. Edwards SV, et al. (2005) Speciation in birds: Genes, geography, and sexual selection. Proc Natl Acad Sci 102(suppl 1):6550–6557.
96. Price T (2008) Speciation in Birds (Roberts and Company).
97. Freeman BG (2017) Little evidence for Bergmann's rule body size clines in passerines along tropical elevational gradients. J Biogeogr 44(3):502–510.
98. Derryberry EP, et al. (2018) Ecological drivers of song evolution in birds: Disentangling the effects of habitat and morphology. Ecol Evol 8(3):1890–1905.
99. Slabbekoorn H, Smith TB (2002) Habitat-Dependent Song Divergence in the Little Greenbul: An Analysis of Environmental Selection Pressures on Acoustic Signals. Evolution 56(9):1849–1858.
100. Zhen Y, et al. (2017) Genomic divergence across ecological gradients in the Central African rainforest songbird (Andropadus virens). Mol Ecol 26(19):4966–4977.
101. Schulenberg TS (2007) Birds of Peru (Princeton University Press).
102. Gregory P (2017) Birds of New Guinea (Lynx Edicions, Barcelona).
103. R Core Team (2018) R: A language and environment for statistical computing. (R Foundation for Statistical Computing, Vienna, Austria) Available at: https://www.R-project.org/.

104. Scrucca L, Fop M, Murphy TB, Raftery AE (2016) mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models. R J 8(1):289–317.
105. Dumbacher JP, Fleischer RC (2001) Phylogenetic evidence for colour pattern convergence in toxic pitohuis: Müllerian mimicry in birds? Proc R Soc Lond B Biol Sci 268(1480):1971–1976.
106. McKenna A, et al. (2010) The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20(9):1297–1303.
107. Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L (2013) mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinforma Oxf Engl 29(13):1682–1684.
108. Danecek P, et al. (2011) The variant call format and VCFtools. Bioinformatics 27(15):2156–2158.
109. Hahn C, Bachmann L, Chevreux B (2013) Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. Nucleic Acids Res 41(13):e129.
110. Andersen MJ, et al. (2015) Rapid diversification and secondary sympatry in Australo-Pacific kingfishers (Aves: Alcedinidae: Todiramphus). R Soc Open Sci 2(2). doi:10.1098/rsos.140375.
111. Paradis E, Claude J, Strimmer K (2004) APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics 20(2):289–290.
112. Bryant D, Moulton V (2004) Neighbor-Net: An Agglomerative Method for the Construction of Phylogenetic Networks. Mol Biol Evol 21(2):255–265.
113. Huson DH, Bryant D (2006) Application of Phylogenetic Networks in Evolutionary Studies. Mol Biol Evol 23(2):254–267.
114. Miller M, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for Inference of Large Phylogenetic Trees. Proc Gatew Comput Environ Workshop GCE:1–8.
115. Smeds L, Qvarnström A, Ellegren H (2016) Direct estimate of the rate of germline mutation in a bird. Genome Res 26(9):1211–1218.
116. Korneliussen TS, Albrechtsen A, Nielsen R (2014) ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics 15(1):356.
117. Warren WC, et al. (2010) The genome of a songbird. Nature 464(7289):757–762.
118. Delcher AL, et al. (1999) Alignment of whole genomes. Nucleic Acids Res 27(11):2369–2376.