

## PIGEON GENOMICS

# Natural selection shaped the rise and fall of passenger pigeon genomic diversity

Gemma G. R. Murray,<sup>1\*</sup> André E. R. Soares,<sup>1\*†</sup> Ben J. Novak,<sup>1,2</sup> Nathan K. Schaefer,<sup>3</sup> James A. Cahill,<sup>1</sup> Allan J. Baker,<sup>4†</sup> John R. Demboski,<sup>5</sup> Andrew Doll,<sup>5</sup> Rute R. Da Fonseca,<sup>6</sup> Tara L. Fulton,<sup>1,7</sup> M. Thomas P. Gilbert,<sup>6,8</sup> Peter D. Heintzman,<sup>1,9</sup> Brandon Letts,<sup>10</sup> George McIntosh,<sup>11</sup> Brendan L. O'Connell,<sup>3</sup> Mark Peck,<sup>5</sup> Marie-Lorraine Pipes,<sup>12</sup> Edward S. Rice,<sup>3</sup> Kathryn M. Santos,<sup>11</sup> A. Gregory Sohrweide,<sup>13</sup> Samuel H. Vohr,<sup>3</sup> Russell B. Corbett-Detig,<sup>3,14</sup> Richard E. Green,<sup>3,14</sup> Beth Shapiro<sup>1,14§</sup>

The extinct passenger pigeon was once the most abundant bird in North America, and possibly the world. Although theory predicts that large populations will be more genetically diverse, passenger pigeon genetic diversity was surprisingly low. To investigate this disconnect, we analyzed 41 mitochondrial and 4 nuclear genomes from passenger pigeons and 2 genomes from band-tailed pigeons, which are passenger pigeons' closest living relatives. Passenger pigeons' large population size appears to have allowed for faster adaptive evolution and removal of harmful mutations, driving a huge loss in their neutral genetic diversity. These results demonstrate the effect that selection can have on a vertebrate genome and contradict results that suggested that population instability contributed to this species's surprisingly rapid extinction.

**T**he passenger pigeon (*Ectopistes migratorius*) numbered between 3 billion and 5 billion individuals before its 19th-century decline and eventual extinction (1). Passenger pigeons were highly mobile and bred in large social colonies, and their population lacked clear geographic structure (2). Few vertebrates have populations this large and cohesive, and the neutral model of molecular evolution predicts that effective population size ( $N_e$ ) and genetic diversity will increase in proportion to population size (3). Preliminary analyses of passenger pigeon genomes have, however, revealed surprisingly low genetic diversity (4). This finding has been interpreted within the framework of the neutral theory of molecular evolution as the result of a history of large

demographic fluctuations (4). However, in large populations, natural selection may be particularly important in shaping genetic diversity: Population genetic theory predicts that selection will be more effective in large populations (3), and selection on one locus can cause a loss of diversity at other loci, particularly those that are closely linked (5–8). It has been suggested that this could explain why the genetic diversity of a species is poorly predicted by its population size (9–11).

We investigated the impact of natural selection on passenger pigeon genomes through comparative genomic analyses of both passenger pigeons and band-tailed pigeons (*Patagioenas fasciata*). Although ecologically and physiologically similar to passenger pigeons, band-tailed pigeons have a present-day population size three orders of magnitude smaller than that of their close relative, the passenger pigeon (2, 12, 13).

We applied a Bayesian skyline model of ancestral population dynamics to the mitochondrial genomes of 41 passenger pigeons from across their former breeding range (Fig. 1A and table S1) (14). This returned a most recent effective population size ( $N_e$ ) of 13 million [95% highest posterior density (HPD) interval: 2 million to 58 million] and a similar, stable  $N_e$  for the previous 20,000 years (Fig. 1B). Although this  $N_e$  is much lower than the (census) population size ( $N_c$ ), it is greater than previous estimates from analyses of nuclear genomes (4) and is likely to be conservative (14).

We compared nucleotide diversity ( $\pi$ ) in the passenger pigeon nuclear genome to  $\pi$  in the band-tailed pigeon nuclear genome. We analyzed four high-coverage passenger pigeon genome assemblies (two newly sequenced and two from published raw data; table S2) and two high-coverage band-tailed pigeon genome assemblies.  $\pi$  was

greater in passenger pigeons (average  $\pi = 0.008$ ) than in band-tailed pigeons (average  $\pi = 0.004$ ), but this difference is less than expected given their population sizes [it suggests that  $N_e/N_c$  was 0.0002 for passenger pigeons compared to 0.2 for band-tailed pigeons; (14)]. We estimated  $\pi$  for nonoverlapping 5-Mb windows across the genome and found that these species exhibit a correlated regional variation in  $\pi$ , but with greater variation in passenger pigeons (Fig. 2A and figs. S1 to S4).

To explore this variation, we mapped our scaffolds to the chicken genome assembly (14), which approximates chromosomal structure because karyotype and synteny are strongly conserved across birds (15). We found that low genetic diversity regions of the passenger pigeon genome are generally in the centers of macrochromosomes, whereas the edges of macrochromosomes and microchromosomes have higher diversity (Fig. 2B). Although this pattern is largely absent from the band-tailed pigeon genome, it is unlikely to be an artefact of ancient DNA damage as our assemblies had high coverage depth (table S2), we used conservative cut-offs for calling variants, and we recovered similar patterns after excluding variants more likely to have resulted from damage (fig. S5) (14). Instead, the pattern mirrors the recombination landscape of the bird genome, where recombination rates are lower in the centers of macrochromosomes, relative both to their edges and the microchromosomes (14, 15).

We next investigated the impact of selection on the evolution of protein-coding regions of the genome in both species. We calculated the rate of adaptive substitution relative to the rate of neutral substitution ( $\omega_a$ ) and the ratio of nonsynonymous to synonymous polymorphism ( $pN/pS$ ) for 5-Mb windows across the genome. A higher  $\omega_a$  suggests stronger or more efficient positive selection, and a lower  $pN/pS$  suggests stronger or more efficient selective constraint.  $\omega_a$  was higher (Mann-Whitney U test,  $P = 1.3 \times 10^{-5}$ ) and  $pN/pS$  lower ( $P = 8.2 \times 10^{-12}$ ) in passenger pigeons than band-tailed pigeons (Fig. 3 and fig. S6). We also found that  $\omega_a$  was higher ( $P = 2.2 \times 10^{-8}$ ) and  $pN/pS$  lower ( $P = 4.1 \times 10^{-6}$ ) in high-diversity regions of the passenger pigeon genome compared to low-diversity regions (Fig. 3 and fig. S6). In addition, codon usage bias, which is thought to reflect selection for translational optimization (16), was greater in passenger pigeons than in band-tailed pigeons, and greater in high-diversity regions (figs. S19 and S20).

We also estimated the difference between the proportions of substitutions and polymorphisms that are nonsynonymous (the direction of selection, DoS) for individual genes, where a positive DoS indicates adaptive evolution. DoS was more often positive in passenger pigeons than in band-tailed pigeons and, in passenger pigeons, DoS was correlated with diversity (fig. S7). McDonald-Kreitman tests (17) identified 32 genes with evidence of adaptive evolution in passenger pigeons (table S3). Among them are genes associated with immune defense (e.g., *CPD*), seasonal consumption of high-sugar foods in passerine birds

<sup>1</sup>Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA 95064, USA. <sup>2</sup>Revive & Restore, Sausalito, CA 94965, USA. <sup>3</sup>Department of Biomolecular Engineering, University of California Santa Cruz, Santa Cruz, CA 95064, USA. <sup>4</sup>Department of Natural History, Royal Ontario Museum, Toronto, ON M5S 2C6, Canada. <sup>5</sup>Department of Zoology, Denver Museum of Nature and Science, Denver, CO 80205, USA. <sup>6</sup>Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5-7, 1350 Copenhagen, Denmark. <sup>7</sup>Environment and Climate Change Canada, 9250-49th Street, Edmonton, AB T6B 1K5, Canada. <sup>8</sup>NTNU University Museum, 7491 Trondheim, Norway. <sup>9</sup>Tromsø University Museum, UiT–The Arctic University of Norway, 9037 Tromsø, Norway. <sup>10</sup>Department of Biology, The Pennsylvania State University, University Park, PA 16802, USA. <sup>11</sup>Collections Department, Rochester Museum and Science Center, Rochester, NY 14607, USA. <sup>12</sup>Marie-Lorraine Pipes, Zooarchaeologist Consultant, Victor, NY 14564, USA. <sup>13</sup>A. Gregory Sohrweide D.D.S., Baldwinsville, NY 13027, USA. <sup>14</sup>University of California Santa Cruz Genomics Institute, 1156 High Street, Santa Cruz, CA 95064, USA.

\*These authors contributed equally to this work. †Present address: Laboratório Nacional de Computação Científica, Petrópolis, RJ, Brazil. ‡Deceased

§Corresponding author. Email: beth.shapiro@gmail.com

(*SI*), and stress modulation (*FAAH*). Selection on these gene functions is consistent with the distinctive diet of passenger pigeons, their sociability, and their population size and density (2), which could have led to an increased burden of transmissible pathogens (18) and increased stress (19).

Differences in the impact of selection between passenger pigeons and band-tailed pigeons could derive from differences in recombination rate, mutation rate, and the distribution of fitness effects. However, the close relationship between these species makes substantial differences in most of these factors unlikely, and the most parsimonious explanation is their different population sizes. Theory predicts that larger populations will experience a greater impact of natural selection, both because they generate more mutations per generation and because selection is more effective in overcoming random drift when  $N_e$  is large (3).

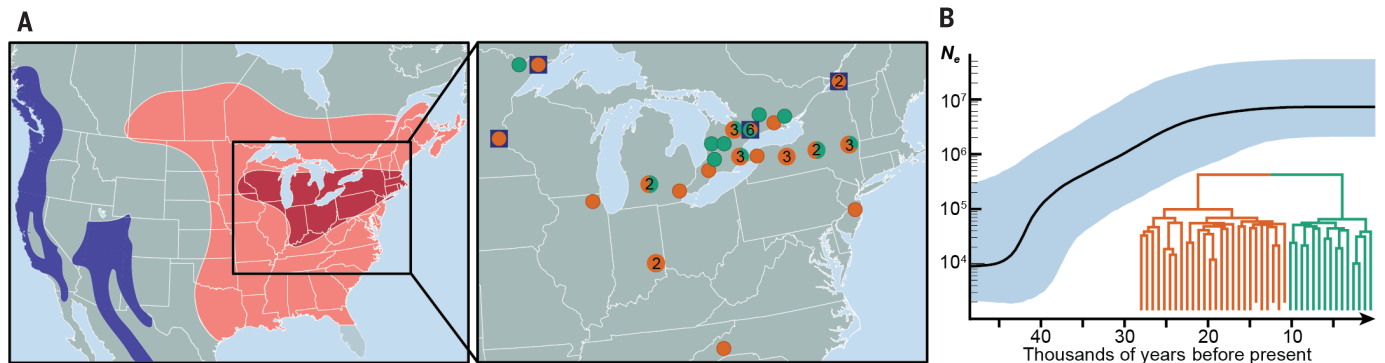
A greater impact of selection on nonsynonymous sites could also increase the impact of selection on neutral sites due to linkage. In linked regions, selection on one site can lead to reduced diversity at neutral sites and a reduced efficiency of selection at other selected sites (3, 20). The impact of this will be greater where recombination rates are low because linked regions will be

larger. Therefore, the large population size of the passenger pigeon, assuming a typical avian recombination landscape, may have resulted in an overall increased neutral genetic diversity and efficiency of selection, but genetic diversity and efficiency of selection in genomic regions with lower recombination rates were reduced due to linkage with selected variants. This explains the pattern of diversity across the passenger pigeon genome (Fig. 2), including the low diversity in the mitochondrial genome (Fig. 1B) (14). It is also supported by other avian studies, which report a correlation between recombination rate and both diversity (21, 22) and the efficiency of selection (23, 24). However, it has been argued that the observed correlation between recombination and the efficiency of selection could be an artefact of GC-biased gene conversion (gBGC) (25).

Regions of the genome with higher recombination rates are expected to accumulate GC substitutions faster as a result of gBGC. gBGC promotes the fixation of A/T to G/C mutations and the loss of G/C to A/T mutations by preferentially replacing A/T bases with G/C bases when recombination occurs at a heterozygous locus (26). gBGC is predicted to have a greater influence when  $N_e$  is large (26). We observe a

higher GC content in high-recombination regions of both pigeon species' genomes (fig. S8), indicating a long-term influence of gBGC. We also observe a higher rate of A/T to G/C substitution and a lower rate of G/C to A/T substitution in passenger pigeons than in band-tailed pigeons, indicating a greater influence of gBGC in passenger pigeons (Fig. 4, A and B).

The purging of deleterious G/C mutations or fixing of beneficial A/T mutations could create the appearance of a greater efficiency of selection in passenger pigeons (25). This is apparent in our observation that in regions of the passenger pigeon genome with high recombination rates (and high diversity), there is both a higher rate of nonsynonymous substitution relative to synonymous substitution ( $dN/dS$ ) for substitutions opposed by gBGC and a lower  $dN/dS$  for substitutions promoted by gBGC (Fig. 4, C and D, and fig. S9). We also find that gBGC influences  $\omega_a$  and  $pN/pS$  (figs. S10 and S11). To test whether our inference of more efficient selection in passenger pigeons is an artefact of gBGC, we estimated  $\omega_a$  and  $pN/pS$  separately for G/C to G/C and A/T to A/T mutations, which are unaffected by gBGC. For these mutations, we again observed higher  $\omega_a$  and lower  $pN/pS$  in passenger pigeons than in band-tailed pigeons (figs. S10 and S11),

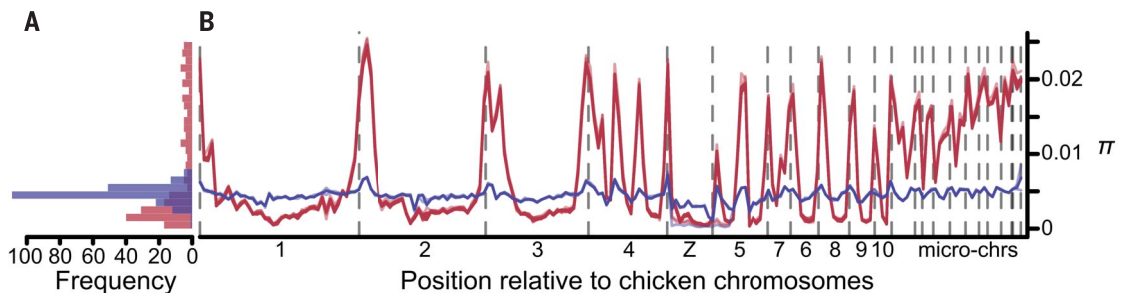


**Fig. 1. Passenger pigeon range, sample origins, and  $N_e$  estimate from mitochondrial genomes.** (A) Range of passenger pigeons at time of European contact (dark red: breeding range; light red: full range) (1) and current range of band-tailed pigeons (purple) (12), with the inset showing the location of origin of the 41 passenger pigeon samples analyzed here. Locations of the four samples

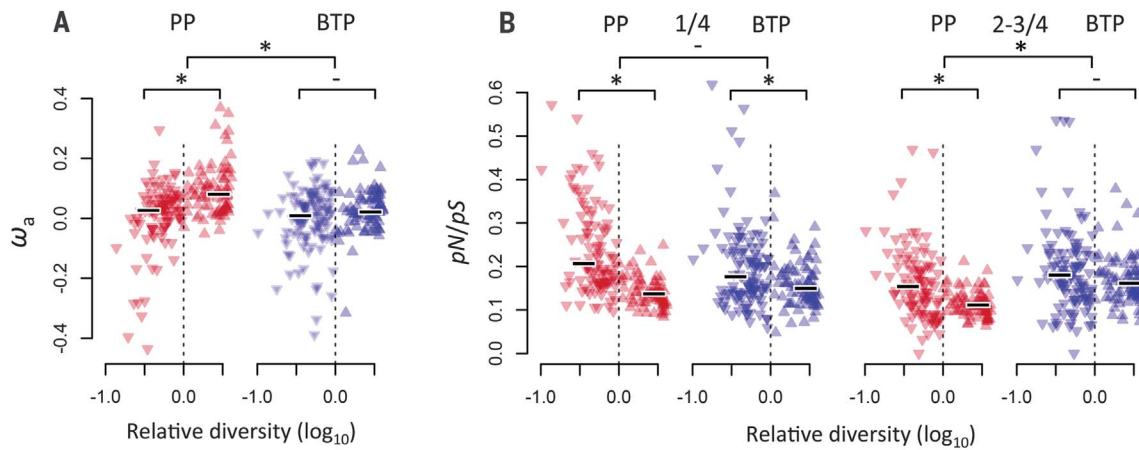
from which nuclear genomes were generated are indicated with a blue box. (B) Inferred  $N_e$  (blue shading indicates the 95% HPD interval) and mitochondrial phylogeny from a Bayesian coalescent analysis. Colors in the inset to (A) match the phylogeny in (B). The structure of the phylogeny does not correlate with geography, which is consistent with an absence of geographic population structure.

**Fig. 2.  $\pi$  across passenger pigeon and band-tailed pigeon genomes.**

(A) A histogram describing mean  $\pi$  for 5-Mb windows across the passenger pigeon (red) and band-tailed pigeon (blue) genomes. (B) Genomic distribution of individual pairwise estimates of mean  $\pi$  in 5-Mb windows across the two species' genomes. Each between- and within-individual pairwise comparison is plotted as red (28 passenger pigeon comparisons) or blue (6 band-tailed pigeon



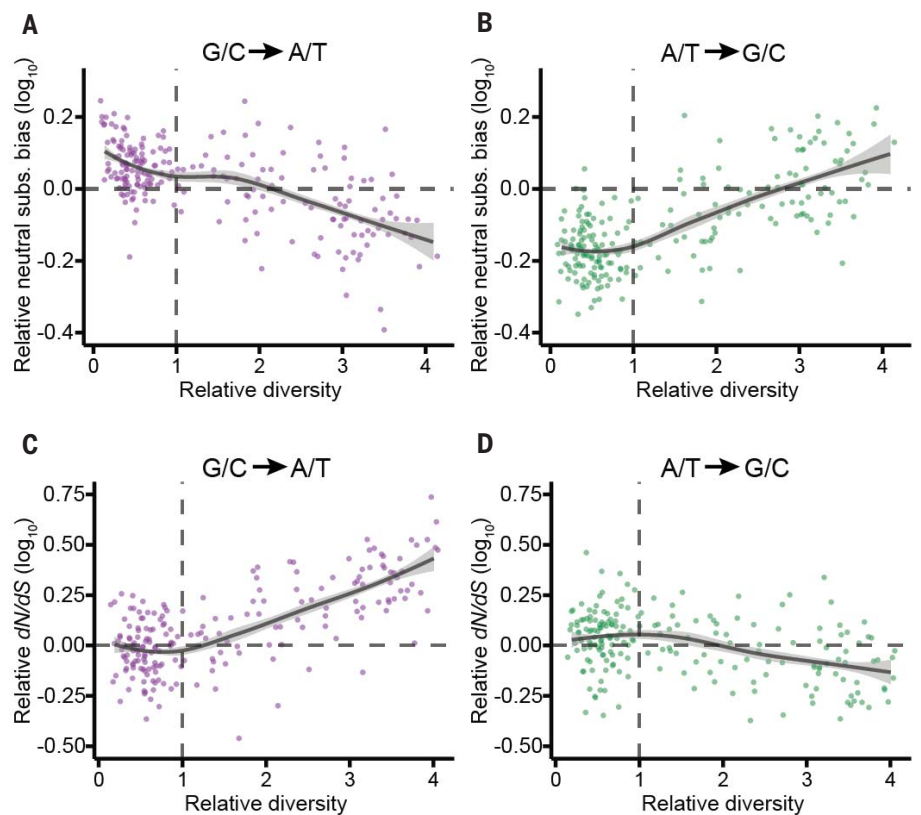
comparisons) lines. Chromosome boundaries are indicated as vertical dashed lines. Chromosomes are ordered by their size in the chicken genome.



**Fig. 3. Estimates of  $\omega_a$  and  $pN/pS$ .** Estimates are averages for 5-Mb windows and are plotted against the window's genetic diversity in passenger pigeons relative to band-tailed pigeons (on a  $\log_{10}$ -scale). Comparisons are drawn between (A)  $\omega_a$  and (B)  $pN/pS$  in passenger pigeons (PP; red) and band-tailed pigeons (BTP; blue) and between low-diversity ( $\pi_{PP} < \pi_{BTP}$ ; point-down triangles) and high-diversity ( $\pi_{PP} > \pi_{BTP}$ ; point-up triangles) windows [median values are shown as horizontal lines;

an asterisk (\*) indicates  $P \leq 1 \times 10^{-4}$  and a dash (-)  $P \geq 0.1$  in a Mann-Whitney U test]. (B)  $pN/pS$  estimates for derived mutations present in one out of four and two or three out of four individuals. A higher  $pN/pS$  for lower-frequency mutations could reflect the slow purging of weakly deleterious mutations. Estimates are based on analyses of two individuals from each species (see fig. S6 for estimates using all passenger pigeon samples).

**Fig. 4. Patterns of substitution for nucleotide base changes that are opposed (A and C) and promoted (B and D) by gBGC.** (A) The rate of G/C to A/T substitution relative to G/C to G/C substitution in passenger pigeons, divided by the same parameter in band-tailed pigeons. (B) The rate of A/T to G/C substitution relative to A/T to A/T substitution in the passenger pigeon lineage, divided by the same parameter in band-tailed pigeons. (C)  $dN/dS$  for G/C to A/T mutations in passenger pigeons, divided by the same parameter in band-tailed pigeons. (D)  $dN/dS$  for A/T to G/C mutations in passenger pigeons, divided by the same parameter in band-tailed pigeons. All estimates are for 5-Mb windows across the genome and are plotted on a  $\log_{10}$ -scale against diversity in passenger pigeons relative to band-tailed pigeons. Trend lines were estimated using the "stat\_smooth" function in *ggplot2* (method = "loess") in R. Shading reflects 95% confidence limits around the trend lines.



confirming that passenger pigeons experience more efficient selection. However, when comparing high- and low-diversity regions of the passenger pigeon genome, we only observe a difference in  $pN/pS$ . This indicates that differences in  $\omega_a$

across the passenger pigeon genome may have been driven by gBGC.

Passenger pigeons' low genetic diversity has previously been explained as the result of drastic population fluctuations driven by resource avail-

ability based on pairwise sequentially Markovian coalescent (PSMC) analyses of the nuclear genome (4, 14). By contrast, our analyses reveal both population stability preceding the species' extinction and a surprisingly pervasive influence

of natural selection. Moreover, the extent of the influence of selection across the passenger pigeon genome indicates that analyses such as PSMC are unlikely to reliably inform us of demographic history (14). Our results therefore do not support the hypothesis that natural demographic fluctuations contributed to the passenger pigeon's extinction, and instead suggest that passenger pigeons may have evolved traits that were adaptive when their population was large but that made it more difficult for them to survive after their population was diminished by the commercial harvest (2). More broadly, our results suggest that even species with large and stable population sizes can be at risk of extinction after a sudden environmental change.

#### REFERENCES AND NOTES

1. A. W. Schorger, *The Passenger Pigeon: Its Natural History And Extinction* (Literary Licensing, LLC, 1955).
2. E. H. Bucher, in *Current Ornithology*, D. M. Power, Ed. (Springer US, 1992), *Current Ornithology*, pp. 1–36.
3. B. Charlesworth, *Nat. Rev. Genet.* **10**, 195–205 (2009).
4. C.-M. Hung et al., *Proc. Natl. Acad. Sci. U.S.A.* **111**, 10636–10641 (2014).
5. J. Maynard Smith, J. Haigh, *Genet. Res.* **23**, 23–35 (1974).
6. D. J. Begun, C. F. Aquadro, *Nature* **356**, 519–520 (1992).
7. J. H. Gillespie, *Genetics* **155**, 909–919 (2000).
8. B. Charlesworth, *Genetics* **190**, 5–22 (2012).
9. R. C. Lewontin, *The Genetic Basis of Evolutionary Change* (Columbia University Press, 1974).
10. R. B. Corbett-Detig, D. L. Hartl, T. B. Sackton, *PLOS Biol.* **13**, e1002112 (2015).
11. E. M. Leffler et al., *PLOS Biol.* **10**, e1001388 (2012).
12. T. A. Sanders, Band-tailed pigeon population status, 2014. U.S. Department of the Interior, Fish and Wildlife Service, Division of Migratory Bird Management, Washington, D.C. (2014).
13. K. P. Johnson, D. H. Clayton, J. P. Dumbacher, R. C. Fleischer, *Mol. Phylogenet. Evol.* **57**, 455–458 (2010).
14. Materials and methods are available as supplementary materials.
15. H. Ellegren, *Trends Ecol. Evol.* **25**, 283–291 (2010).
16. R. Hershberg, D. A. Petrov, *Annu. Rev. Genet.* **42**, 287–299 (2008).
17. J. H. McDonald, M. Kreitman, *Nature* **351**, 652–654 (1991).
18. B. T. Grenfell, A. P. Dobson, *Ecology of Infectious Diseases in Natural Populations* (Cambridge University Press, 1995).
19. S. Creel, B. Dantzer, W. Goymann, D. R. Rubenstein, *Funct. Ecol.* **27**, 66–80 (2013).
20. W. G. Hill, A. Robertson, *Genet. Res.* **8**, 269–294 (1966).
21. B. M. Van Doren et al., *Mol. Ecol.* **26**, 3982–3997 (2017).
22. L. Dutoit et al., *Proc. Biol. Sci.* **284**, 20162756 (2017).
23. K. Nam et al., *Genome Biol.* **11**, R68 (2010).
24. T. I. Gossmann, A. W. Santure, B. C. Sheldon, J. Slate, K. Zeng, *Genome Biol. Evol.* **6**, 2061–2075 (2014).
25. P. Bolívar, C. F. Mugal, A. Nater, H. Ellegren, *Mol. Biol. Evol.* **33**, 216–227 (2016).
26. L. Duret, A. Eyre-Walker, N. Galtier, *Gene* **385**, 71–74 (2006).

#### ACKNOWLEDGMENTS

We thank L. Shiue, S. Weber, J. Kapp, M. Stiller, T. Kuhn, S. Wagner, and R. Shaw for assistance generating data. We thank J. Novembre for advice on analyzing codon usage bias. Research was supported by the Packard Foundation, the Gordon and Betty Moore Foundation, and

Revive & Restore. A.E.R.S. was funded by Ciência sem Fronteiras fellowship–CAPES, Brazil. Sequencing was supported by the Dean's Office, the Vincent J. Coates Genomics Sequencing Laboratory at University of California Berkeley (Berkeley sequencing supported by NIH S10 Instrumentation Grants S10RR029668 and S10RR027303), and the Danish National Sequencing Centre in Copenhagen (sequencing supported by Lundbeck Foundation grant R52-5062). R.E.G. is cofounder of Dovetail Genomics, LLC. K.M.S. is Senior Director of Collections and Exhibitions at the Rochester Museum and Science Center. The sequence data generated in this study are archived in the relevant National Center for Biotechnology Information databases: The band-tailed pigeon assembly and RNA-seq reads used for its annotation can be found in Bioproject PRJNA308039 and reads from passenger pigeon samples in PRJNA381231 (accession numbers are provided in table S1). The authors declare no competing interests. B.S. conceived and designed the study with critical input from G.G.R.M., A.E.R.S., R.E.G., and R.B.C.-D.; B.S., T.L.F., and B.J.N. led sample collection; A.J.B., A.D., J.R.D., A.G.S., K.M.S., M.T.P.G., and M.P. provided samples; A.E.R.S., T.L.F., B.L., B.J.N., and R.R.DaF. performed DNA extraction and library preparation experiments; A.E.R.S. and P.D.H. performed mitochondrial genome assembly and analyses; A.E.R.S., N.K.S., E.S.R., J.A.C., S.H.V., and P.D.H. performed nuclear genome assembly and analyses; G.G.R.M. designed and performed selection analyses; B.S., G.G.R.M., A.E.R.S., and R.E.G. wrote the paper; and all authors contributed to editing the manuscript.

#### SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/358/6365/951/suppl/DC1  
Supplementary Text  
Figs. S1 to S31  
Tables S1 to S10  
References (27–130)

15 June 2017; accepted 28 September 2017  
10.1126/science.aao0960

## Natural selection shaped the rise and fall of passenger pigeon genomic diversity

Gemma G. R. Murray, André E. R. Soares, Ben J. Novak, Nathan K. Schaefer, James A. Cahill, Allan J. Baker, John R. Demboski, Andrew Doll, Rute R. Da Fonseca, Tara L. Fulton, M. Thomas P. Gilbert, Peter D. Heintzman, Brandon Letts, George McIntosh, Brendan L. O'Connell, Mark Peck, Marie-Lorraine Pipes, Edward S. Rice, Kathryn M. Santos, A. Gregory Sahrweide, Samuel H. Vohr, Russell B. Corbett-Detig, Richard E. Green and Beth Shapiro

*Science* **358** (6365), 951-954.  
DOI: 10.1126/science.aao0960

### Genetics of the passenger pigeon

The now-extinct passenger pigeon used to be one of the most numerous vertebrates on Earth. Murray *et al.* examined the genomes of four passenger pigeon samples from different locales within its range. They describe the interplay between passenger pigeon population size, genome structure and recombination, and natural selection. They conclude that a reduction in genetic diversity provided few avenues for the bird to respond to human pressures, which ultimately drove it to extinction.

*Science*, this issue p. 951

#### ARTICLE TOOLS

<http://science.sciencemag.org/content/358/6365/951>

#### SUPPLEMENTARY MATERIALS

<http://science.sciencemag.org/content/suppl/2017/11/17/358.6365.951.DC1>

#### REFERENCES

This article cites 120 articles, 30 of which you can access for free  
<http://science.sciencemag.org/content/358/6365/951#BIBL>

#### PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)