PHENOTYPIC EVOLUTION

Convergent regulatory evolution and loss of flight in paleognathous birds

Timothy B. Sackton[1,2]*, Phil Grayson[2,3], Alison Cloutier[2,3], Zhirui Hu[4], Jun S. Liu[4], Nicole E. Wheeler[5,6], Paul P. Gardner[5,7], Julia A. Clarke[8], Allan J. Baker[9,10], Michele Clamp[1], Scott V. Edwards[2,3]*

A core question in evolutionary biology is whether convergent phenotypic evolution is driven by convergent molecular changes in proteins or regulatory regions. We combined phylogenomic, developmental, and epigenomic analysis of 11 new genomes of paleognathous birds, including an extinct moa, to show that convergent evolution of regulatory regions, more so than protein-coding genes, is prevalent among developmental pathways associated with independent losses of flight. A Bayesian analysis of 284,001 conserved noncoding elements, 60,665 of which are corroborated as enhancers by open chromatin states during development, identified 2355 independent accelerations along lineages of flightless paleognaths, with functional consequences for driving gene expression in the developing forelimb. Our results suggest that the genomic landscape associated with morphological convergence in ratites has a substantial shared regulatory component.

C onvergent evolution—the independent evolution of similar phenotypes in divergent taxa—produces some of the most striking examples of adaptation, but the molecular architecture of convergent traits is not well understood (1–4). In some cases, similar or identical mutations in independent lineages appear to be associated with convergent phenotypes (5, 6); in other cases, convergent phenotypes appear to arise by diverse molecular paths (7). Studies of convergence are increasingly focused on detecting genome-wide patterns (8–13), but the relative contributions of regulatory regions and protein-coding genes to convergent phenotypes remain undetermined.

A particularly notable example of a convergent trait involves loss of powered flight, which has occurred many times independently in the course of avian evolution. One of the most iconic groups of flightless birds is the ratites, consisting of extant ostriches, kiwi, rheas, cassowaries, and emus, as well as the extinct moa and elephant birds. All ratites show morphological similarities including forelimb reduction (ranging from moderate in ostriches and rheas to a complete absence in moa), reduced pectoral muscle mass associated with the absence of the sternal keel, and feather modifications, as well as generally larger body size (14–16). Historically, ratites were thought to be a monophyletic sister clade to the volant tinamous (17); however, despite remaining uncertainties in paleognath relationships, recent molecular phylogenetic evidence strongly supports ratite paraphyly, implying as many as six independent losses of flight within this group based on phylogenetic relationships and biogeographic scenarios (16, 18–22).

To study the genomic correlates of flight loss in ratites, we assembled and annotated 11 new paleognath genomes (fig. S1 and tables S1 to S3) (23), including eight flightless ratites [greater rhea, lesser rhea, cassowary, emu, Okarito kiwi, great spotted kiwi, little spotted kiwi, and the extinct little bush moa (19, 24)] and three tinamous (thicket tinamou, elegant crested tinamou, and Chilean tinamou), and analyzed them together with the published ostrich, white-throated tinamou (25), and North Island brown kiwi (26) genomes, along with 30 existing neognath and outgroup genomes.

We compiled a total data set of 41,184,181 base pairs of aligned DNA from 20,850 noncoding loci, including introns, ultraconserved elements (27), and conserved non-exonic elements (CNEEs) (table S4) (23, 28) to test the hypothesis of ratite paraphyly and to resolve the placement of rheas, which remains an outstanding question in paleognath phylogenetics (18, 22, 29). Consistent with recent molecular phylogenies (18–22, 29), we recovered a basal divergence between the ostrich and the remaining paleognaths, including the tinamous, which are therefore nested within a paraphyletic ratite clade (Fig. 1A). However, contrary to recent concatenation analyses (22), our coalescent analysis using MP-EST [Maximum Pseudo-likelihood for Estimating Species Trees (30)] consistently places the rheas as sister to the kiwi + emu + cassowary clade (figs. S2 and S3). Our species tree is further corroborated by 4274 informative CR1 retroelement insertions, including 18 absent in ostriches but shared among all non-ostrich paleognaths, including tinamous (31). Discrepancies with previous analyses are explained by the existence of an empirical anomaly zone (32) in paleognaths, in which incomplete lineage sorting across short internal branches leading to the ancestor of rheas, kiwi, and emu + cassowary results in a set of most common gene trees that differ from the species tree (Fig. 1B and figs. S4 and S5), compromising the concatenation approach. Our analysis and biogeographic considerations (20–22) imply three to six losses of flight in the history of this group (Fig. 1A). The alternative—a single loss of flight at the base of the paleognaths, followed by a regain of flight in tinamous—appears implausible given evidence for repeated losses of flight across birds and the lack of any evidence for regains of flight after loss (fig. S6).

Fig. 1. Genome-wide data place rheas as sister to a kiwi/cassowary-emu clade.
(**A**) MP-EST species tree topology, with 100% bootstrap support throughout (not shown). Branch lengths in units of substitutions per site were estimated with ExaML using a fully partitioned alignment of 20,850 noncoding loci constrained to the MP-EST topology. Ratite paraphyly is consistent with either a single loss of flight in the paleognath ancestor followed by regain in tinamous (green arrows) or a minimum of three independent losses of flight in the ratites (dark blue arrows), with at least five losses suggested by a proposed sister group relationship between elephant birds and kiwi (light blue arrows; elephant bird branch not shown). (**B**) Distribution of the 25 most common gene tree topologies, showing that the species tree topology (in red and marked with an asterisk) is not the most common gene tree topology.
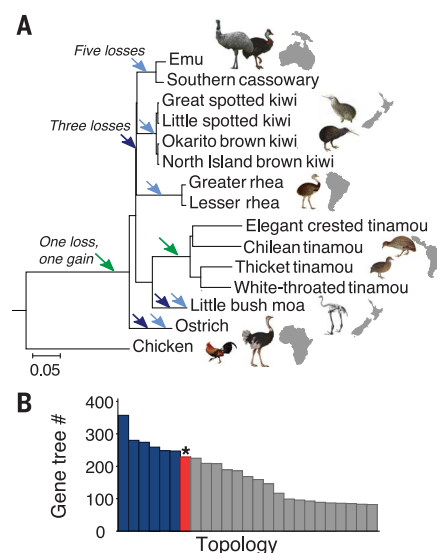
[1]Informatics Group, Harvard University, Cambridge, MA, USA. [2]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA. [3]Museum of Comparative Zoology, Harvard University, Cambridge, MA, USA. [4]Department of Statistics, Harvard University, Cambridge, MA, USA. [5]School of Biological Sciences, University of Canterbury, Christchurch, New Zealand. [6]Centre for Genomic Pathogen Surveillance, Wellcome Sanger Institute, Wellcome Genome Campus, Cambridge, UK. [7]Department of Biochemistry, University of Otago, Dunedin, New Zealand. [8]Jackson School of Geosciences, University of Texas, Austin, TX, USA. [9]Department of Natural History, Royal Ontario Museum, Toronto, Ontario, Canada. [10]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada.
*Corresponding author. Email: tsackton@g.harvard.edu (T.B.S.); sedwards@fas.harvard.edu (S.V.E.)

We analyzed 7654 high-quality gene alignments across 44 species of birds, including nine flightless ratites, four volant tinamous, and 21 neognaths, to test for convergent evolution in protein-coding genes among ratite lineages (*23*). We found no evidence for an excess of convergent amino acid substitutions among ratites (table S5) and minimal evidence for ratite-specific convergent shifts in rates of amino acid divergence (table S6 and fig. S7). However, we did find 238 genes exhibiting positive selection on a subset of codons specifically in ratites, including 63 that are selected in multiple species, and an additional 294 genes with positive selection predominately in ratites (>50% of selected branches are ratites).

The number of convergently selected genes in ratites is greater than expected under the assumption of independent selection in each lineage (permutation *P* < 0.001). However, the degree of convergence we observe declines substantially if we make the conservative assumption that only three independent losses of flight occurred (permutation *P* = 0.0475; fig. S8); we find no cases of genes positively selected in three independent losses of flight under the conservative loss model (in which flight is lost independently only in moa, ostrich, and the rhea/emu/kiwi/cassowary clade; table S7). We also find no evidence that the convergently selected genes are enriched for any Gene Ontology (GO) or Kyoto Encyclopedia of Genes and Genomes (KEGG) terms (*Q* value > 0.20 for all categories after

multiple test correction). The larger set of genes with selection specifically or predominantly in ratites (regardless of degree of convergence) is enriched for terms involved in protein metabolic processes (data S1), but not transcription factors or proteins with functions in development. This finding suggests that convergent protein-coding evolution in these species may be associated with other phenotypes, such as changes

in feathers, body size, or other traits, but not morphological evolution of forelimbs and related skeletal phenotypes such as the loss of a keel.

We applied profile hidden Markov models to detect sequence changes that are underrepresented in homologs of a protein (*33*), similar to that used to identify putative function-altering mutations in protein-coding genes associated with loss of flight in the Galapagos cormorant (*34*). This
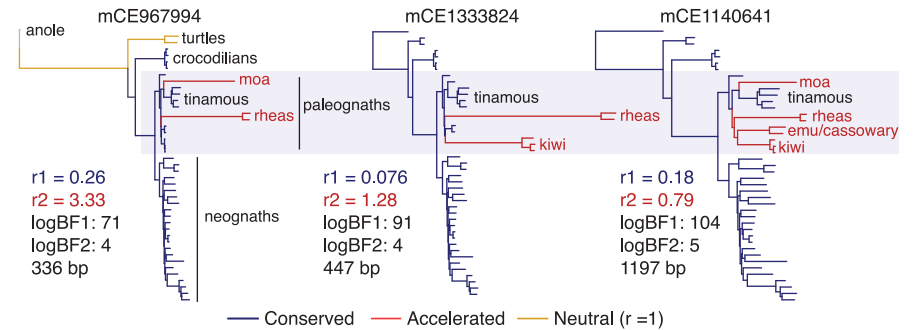

Fig. 2. A Bayesian model identifies convergently accelerated CNEEs in ratites. Shown here are trees for three examples of convergent ratite accelerated elements (specific CNEE IDs are listed above each tree). Major groups of birds and non-avian outgroups are indicated on the leftmost tree. Branch lengths are proportional to substitution rate relative to the neutral rate for each tree (*r* = 1); branches are colored according to the maximum a posteriori conservation state. Bayes factors (BFs) 1 and 2 [see supplementary materials and (*44*)], the conserved (r_1) and accelerated (r_2) rates, and length of each element in base pairs are indicated. Gray branches indicate that the element is missing or ambiguous for that taxon.
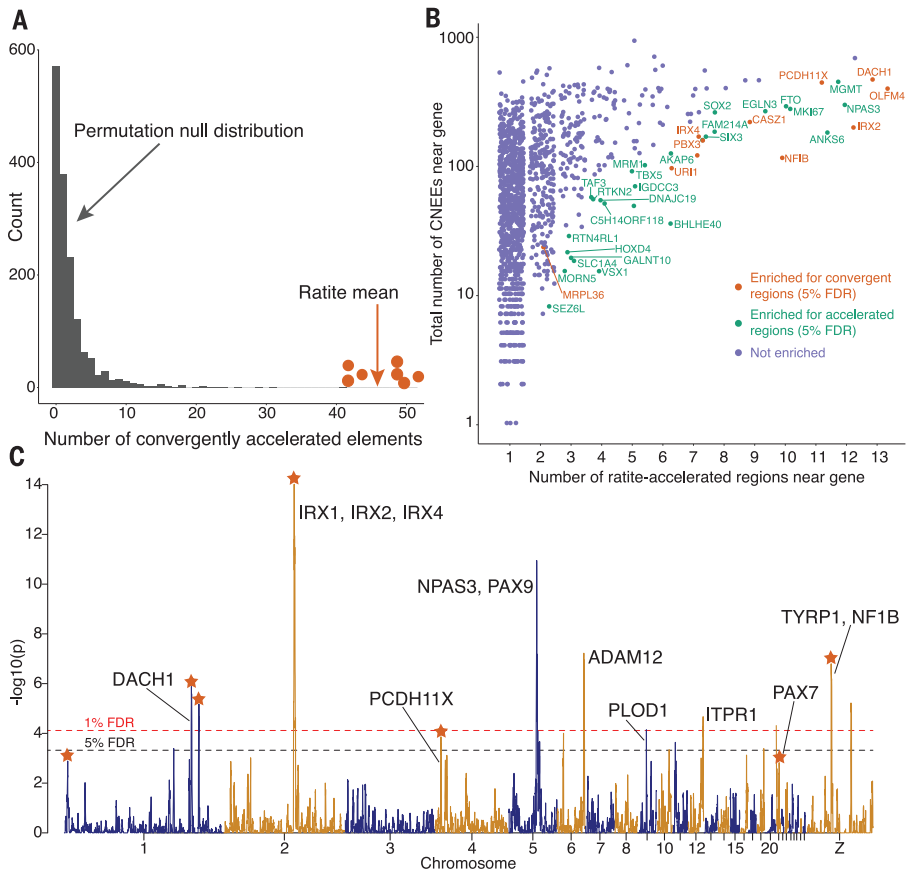

Fig. 3. Ratites exhibit unusually high numbers of convergently accelerated CNEEs. (**A**) Histogram of the null distribution of the observed number of convergently accelerated elements in random trios of neognaths (gray), with individual ratite trios (red points) and ratite mean (red arrow) indicated; ratites have significant excess of convergence (permutation *P* = 0.0063). (**B**) Genes with evidence for excess of nearby accelerated and convergently accelerated elements. Colored and labeled points are genes with excess convergent or accelerated elements based on a permutation test (5% FDR). Only genes with at least one accelerated element are plotted. (**C**) Evidence for spatial clustering of accelerated elements across paleognath genomes. The *y* axis shows the $-\log_{10} P$ value for a test of excess convergent ratite accelerated elements in 1-Mb sliding windows (100-kb slide), computed using a binomial test, where the probability of success is calculated as the total accelerated elements divided by the total number of elements, and the number of samples is the number of CNEEs in each window. Genes in each window that may be of particular interest are noted; windows with an excess of convergently accelerated elements are indicated with a purple star. The reference coordinates are the chicken genome (galGal4).

approach yielded a single gene, *Neurofibromin 1* (*NF1*), showing evidence for convergent functional evolution in ratites [at a 5% false discovery rate (FDR)], but we did not detect an excess of convergent function-altering protein-coding substitutions relative to genome-wide distributions (fig. S9). Little is known about *NF1* function in chickens, but in mice it has been shown to be involved in skeletal (*35*) and skeletal muscle development (*36*).

Regulatory regions may be subject to less pleiotropic constraint than protein-coding genes (*5*, *37*), and they may be more likely to underlie convergent phenotypes than protein-coding genes if common pathways are involved. To identify candidate regulatory regions in paleognath genomes, we used phylogenetic conservation (*38*) to identify 284,001 CNEEs (≥50 base pairs), which are thought to have regulatory roles in birds (*39*) and other taxa (*40–43*). A Bayesian method (*44*) for detecting changes in conservation of these elements across the phylogeny identified 2355 elements specifically accelerated in ratites, 256 to 630 of which have experienced multiple independent acceleration events depending on modeling assumptions (Fig. 2, fig. S10, and table S8); these results are consistent across input datasets and parameter values (figs. S11 and S12).

To determine whether ratites experience more convergent acceleration in CNEEs than is expected by chance, we randomly sampled trios of non-sister neognaths and counted how many CNEEs were convergently accelerated in all three lineages, using the posterior probabilities from an unrestricted Bayesian model (*23*). Among random trios of neognaths, we observed a mean of 1.99 convergently accelerated CNEEs (across 1552 permutations). In contrast, we observed a mean of 45.9 convergently accelerated CNEEs among seven ratite trios (under the conservative model of three independent losses of flight in moa, ostrich, and the ancestor of rhea/kiwi/cassowary/emu).

The mean number of convergently accelerated CNEEs we observed in ratites is thus unusually high compared to an empirical null distribution (permutation *P* = 0.0064, Fig. 3A). Although incomplete lineage sorting can increase apparent substitution rates (*45*), we found no evidence that ratite-accelerated elements are associated with increased gene tree discordance (fig. S13).

We looked for evidence of functional convergence by testing for spatial and pathway enrichment among both the 2355 ratite-specific accelerated CNEEs and the subset that are convergently accelerated in multiple lineages. Convergently accelerated CNEEs were enriched near genes associated with sequence-specific DNA binding, limb morphogenesis, Wnt signaling, and regulation of epithelial cell proliferation GO terms (fig. S14 and data S2). Both ratite-accelerated and convergently accelerated CNEEs were overrepresented near a variety of genes relevant to morphology and development (Fig. 3B), were spatially clustered along the genome (Fig. 3C), and included at least one element (mCE1916069) overlapping a known limb enhancer in mammals. No single gene or accelerated element is expected to account for the full extent of morphological convergence among ratites; however, accelerated elements were enriched near key limb developmental genes including *TBX5*, *DACH1*, *PAX9*, and genes from the *IRX* family (Fig. 3, B and C). *TBX5* is essential for forelimb development in tetrapods (*46*) and pectoral fin development in fish (*47*) and has been previously implicated in morphological divergence in emus (*15*). *IRX5* is critical for both size and anteroposterior patterning in the limbs (*48*). *DACH1* has been shown to repress bone morphogenetic protein (BMP), promote proximodistal patterning, and maintain the apical ectodermal ridge during limb development (*49*). Finally, *PAX9*-deficient mice have craniofacial and limb abnormalities as well as complete loss of teeth (*50*). Some genes enriched for ratite-accelerated elements are also reported near clusters of lineage-accelerated elements in mammals, including *DACH1*, *NPAS3*, and *NF1B* in humans (*51*), *HoxD* in bats (*52*), and *NPAS3* in primates (*53*). Deviations from model expectations, such as base compositional shifts (which some accelerated elements exhibit and could be driven by gene conversion) and the presence of simple sequence repeats (*54*), appear unlikely to explain this overlap (fig. S15). This concordance would suggest that certain key developmental genes are reused repeatedly during morphological evolution across amniotes, or that a high density of nearby regulators predisposes certain genomic regions to repeated evolution.

Ratite accelerated elements are candidates for regulatory regions with functional relevance to the range of convergent phenotypes associated with loss of flight in ratites, although we expect this suite of traits to be highly polygenic. To functionally characterize the association between sequence conservation and regulatory activity in birds, we used ATAC-seq (assay for transposase-accessible chromatin sequencing) to identify genomic regions of open chromatin for eight tissues
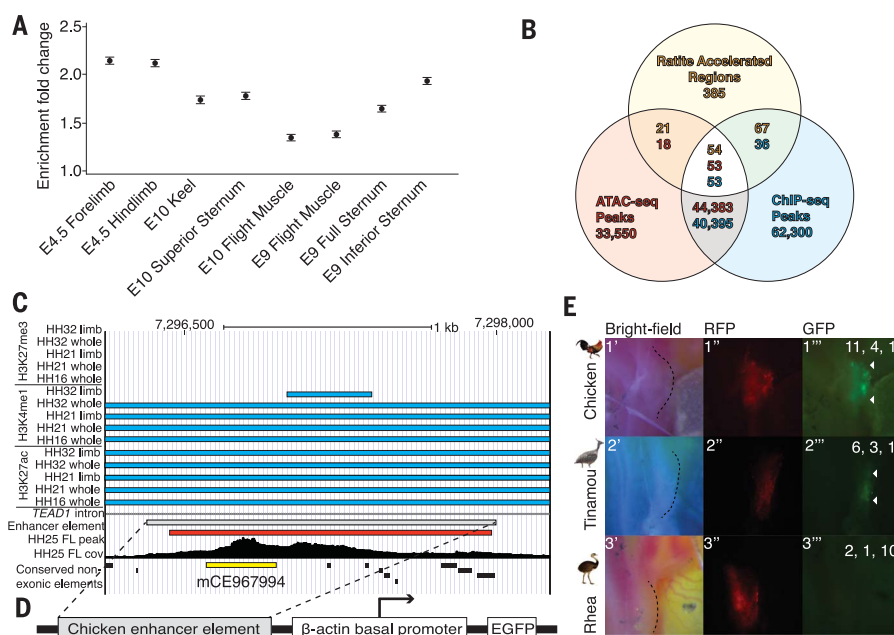

Fig. 4. Ratite-specific convergent acceleration is associated with altered activity in a candidate enhancer. (**A**) Conserved non-exonic elements are enriched under ATAC-seq peaks across all eight chicken tissues. Enrichment fold change (mean and 95% confidence interval) was calculated against chicken genomic background using a permutation-based test for interval overlaps implemented in GAT (10,000 samples). (**B**) Candidate forelimb ATAC-seq peaks for enhancer screen were identified by overlap with convergently accelerated elements and ChIP-seq peaks. (**C**) Genomic region associated with strong enhancer activity in early chicken forelimbs. (**D**) Schematic of β-actin/GFP vector used in enhancer screen. (**E**) Results of enhancer screen of candidate ATAC-seq peak in chicken forelimb bud. Pictures are representative images of 10 to 16 replicates showing a developing forelimb region 24 hours after electroporation (Hamburger and Hamilton stages 19–20). Dashed line on the bright-field image (left column) indicates the forelimb region. Red fluorescent protein expression (middle column) indicates the area electroporated. GFP expression (green, right column) is driven by enhancer activity of the candidate region. Numbers of replicates showing strong, partial, and no GFP expression are indicated in the upper right corner of each image. Top row: Chicken ATAC-seq peak drives consistent GFP expression throughout the developing forelimb (strong GFP 11/16, partial GFP 4/16, no GFP 1/16). Second row: The homologous genomic region in the elegant crested tinamou also drives consistent GFP expression throughout the developing chicken forelimb (strong GFP 6/10, partial GFP 3/10, no GFP 1/10). Third row: The homologous genomic region in the greater rhea fails to drive consistent GFP expression throughout the developing chicken forelimb (strong GFP 2/13, partial GFP 1/13, no GFP 10/13).

during the course of chick development (*55*, *56*). Consistent with recent work (*39*), CNEEs are enriched under ATAC-seq peaks (Fig. 4A), which supports the regulatory function of CNEEs in this and other datasets. To test whether ratite-specific sequence acceleration results in functional differentiation of enhancer activity in vivo, we screened a set of candidates for enhancer activity in developing chicken forelimbs. We focused on putative forelimb enhancers identified by examining the intersection of consistent forelimb ATAC-seq peaks, convergent ratite-accelerated elements, and previously published embryonic chicken ChIP-seq (chromatin immunoprecipitation sequencing) peaks, resulting in 54 candidate enhancers (Fig. 4B). Using an electroporated β-actin/green fluorescent protein (GFP) enhancer construct assay, we identified a promising chicken region from among these candidates, consisting of the ATAC-seq peak containing the convergent ratite-accelerated element mCE967994, which produced consistent strong enhancer activity in early chicken forelimbs (Fig. 4, C to E). This element is located in the intron of the *TEAD1* gene, which has been implicated in cell proliferation, cell survival, mesoderm patterning, and the epithelial-mesenchymal transition (*57–59*) and contains motifs associated with TEAD1 binding, which suggests that it may play an autoregulatory role. We tested for enhancer activity of the homologous genomic region in the volant elegant crested tinamou, in which mCE967994 is identified as conserved in our Bayesian analysis, and the flightless greater rhea, in which mCE967994 is accelerated. We found that the tinamou version of this enhancer consistently drove GFP expression (*N* = 9 of 10 embryos), whereas the rhea version did not (Fig. 4E) (*N* = 10 of 13 embryos). Thus, accelerated sequence evolution of this element in rheas appears to be associated with functional divergence of regulatory activity. Although we cannot rule out the possibility that compensatory trans-regulatory evolution in rheas maintains function of this and other accelerated elements in their natural context, this assay demonstrates that our unbiased comparative genomic screen readily identifies cis-regulatory elements with functionally divergent cis-regulatory activity in vivo.

Evolutionary biology has long been focused on attempting to understand the relative roles of regulatory and protein-coding change in phenotypic evolution, as well as genomic mechanisms underlying convergent phenotypes (*60*). Our unbiased statistical and functional screens, which emphasized genomic changes occurring in parallel on multiple lineages of ratites, suggest that convergent morphological evolution and loss of flight in ratites is associated more strongly with regulatory evolution in noncoding DNA than with evolutionary changes in protein-coding genes. Our findings offer a contrast to previous work that emphasized protein-coding correlates of flightlessness in birds. Although our results do not rule out a role for lineage-specific genomic

drivers of flightlessness, they provide a template for future genome-wide studies of loss of flight and other convergent phenotypes across the Tree of Life.

REFERENCES AND NOTES

1. J. B. Losos, *Evolution* **65**, 1827–1840 (2011).
2. A. Martin, V. Orgogozo, *Evolution* **67**, 1235–1250 (2013).
3. D. L. Stern, *Nat. Rev. Genet.* **14**, 751–764 (2013).
4. J. F. Storz, *Nat. Rev. Genet.* **17**, 239–250 (2016).
5. Y. F. Chan et al., *Science* **327**, 302–305 (2010).
6. Y. Zhen, M. L. Aardema, E. M. Medina, M. Schumer, P. Andolfatto, *Science* **337**, 1634–1637 (2012).
7. K. L. Cooper et al., *Nature* **511**, 41–45 (2014).
8. S. Yeaman et al., *Science* **353**, 1431–1433 (2016).
9. R. Partha et al., *eLife* **6**, e25884 (2017).
10. S. Xu et al., *Mol. Biol. Evol.* **34**, 1008–1015 (2017).
11. M. J. Berger, A. M. Wenger, H. Guturu, G. Bejerano, *Nucleic Acids Res.* **46**, 9299–9308 (2018).
12. W. K. Meyer et al., *Science* **361**, 591–594 (2018).
13. V. Sharma et al., *Nat. Commun.* **9**, 1215 (2018).
14. P. Houde, *Nature* **324**, 563–565 (1986).
15. S. R. B. Bickley, M. P. O. Logan, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 17917–17922 (2014).
16. S. J. Nesbitt, J. A. Clarke, *Bull. Am. Mus. Nat. Hist.* **406**, 1–91 (2016).
17. J. Cracraft, *Ibis* **116**, 494–521 (1974).
18. J. Harshman et al., *Proc. Natl. Acad. Sci. U.S.A.* **105**, 13462–13467 (2008).
19. A. J. Baker, O. Haddrath, J. D. McPherson, A. Cloutier, *Mol. Biol. Evol.* **31**, 1686–1696 (2014).
20. K. J. Mitchell et al., *Science* **344**, 898–900 (2014).
21. A. Grealy et al., *Mol. Phylogenet. Evol.* **109**, 151–163 (2017).
22. T. Yonezawa et al., *Curr. Biol.* **27**, 68–77 (2017).
23. See supplementary materials.
24. A. Cloutier et al., bioRxiv 262816 [preprint]. 11 February 2018.
25. E. D. Jarvis et al., *Science* **346**, 1320–1331 (2014).
26. D. Le Duc et al., *Genome Biol.* **16**, 147 (2015).
27. J. E. McCormack et al., *Genome Res.* **22**, 746–754 (2012).
28. S. V. Edwards, A. Cloutier, A. J. Baker, *Syst. Biol.* **66**, 1028–1044 (2017).
29. M. J. Phillips, G. C. Gibb, E. A. Crimp, D. Penny, *Syst. Biol.* **59**, 90–107 (2010).
30. L. Liu, L. Yu, S. V. Edwards, *BMC Evol. Biol.* **10**, 302 (2010).
31. A. Cloutier et al., bioRxiv 262949 [preprint]. 9 February 2018.
32. J. H. Degnan, N. A. Rosenberg, *Public Libr. Sci. Genet.* **2**, 762–768 (2006).
33. N. E. Wheeler, L. Barquist, R. A. Kingsley, P. P. Gardner, *Bioinformatics* **32**, 3566–3574 (2016).
34. A. Burga et al., *Science* **356**, eaal3345 (2017).
35. M. Kolanczyk et al., *Hum. Mol. Genet.* **16**, 874–886 (2007).
36. N. Kossler et al., *Hum. Mol. Genet.* **20**, 2697–2709 (2011).
37. S. B. Carroll, B. Prud'homme, N. Gompel, *Sci. Am.* **298**, 60–67 (May 2008).
38. A. Siepel et al., *Genome Res.* **15**, 1034–1050 (2005).
39. R. Seki et al., *Nat. Commun.* **8**, 14229 (2017).
40. L. A. Pennacchio et al., *Nature* **444**, 499–502 (2006).
41. A. Visel et al., *Nat. Genet.* **40**, 158–160 (2008).
42. J. A. Capra, G. D. Erwin, G. McKinsey, J. L. R. Rubenstein, K. S. Pollard, *Philos. Trans. R. Soc. B* **368**, 20130025 (2013).
43. D. E. Dickel et al., *Cell* **172**, 491–499.e15 (2018).
44. Z. Hu, T. B. Sackton, S. V. Edwards, J. S. Liu, *Mol. Biol. Evol.* msz049 (2019).
45. F. K. Mendes, M. W. Hahn, *Syst. Biol.* **65**, 711–721 (2016).
46. C. Rallis et al., *Development* **130**, 2741–2751 (2003).
47. D. G. Ahn, M. J. Kourakis, L. A. Rohde, L. M. Silver, R. K. Ho, *Nature* **417**, 754–758 (2002).
48. D. Li et al., *Dev. Cell* **29**, 233–240 (2014).
49. Y. Kida, Y. Maeda, T. Shiraishi, T. Suzuki, T. Ogura, *Development* **131**, 4179–4187 (2004).
50. H. Peters, A. Neubüser, K. Kratochwil, R. Balling, *Genes Dev.* **12**, 2735–2747 (1998).
51. K. S. Pollard et al., *PLOS Genet.* **2**, e168 (2006).
52. B. M. Booker et al., *PLOS Genet.* **12**, e1005738 (2016).
53. D. Kostka, A. K. Holloway, K. S. Pollard, *Mol. Biol. Evol.* **35**, 2034–2045 (2018).
54. K. T. Xie et al., *Science* **363**, 81–84 (2019).
55. J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang, W. J. Greenleaf, *Nat. Methods* **10**, 1213–1218 (2013).
56. A. R. Gehrke et al., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 803–808 (2015).
57. H. Zhang et al., *J. Biol. Chem.* **284**, 13355–13362 (2009).
58. B. Zhao et al., *Genes Dev.* **22**, 1962–1971 (2008).
59. A. Sawada et al., *Mol. Cell. Biol.* **28**, 3177–3189 (2008).
60. S. Lamichhaney et al., bioRxiv 574756 [preprint]. 12 March 2019.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/364/6435/74/suppl/DC1
Figs. S1 to S15
Tables S1 to S8
References (*61–113*)
Data S1 to S5

Science

Convergent regulatory evolution and loss of flight in paleognathous birds

Timothy B. Sackton, Phil Grayson, Alison Cloutier, Zhirui Hu, Jun S. Liu, Nicole E. Wheeler, Paul P. Gardner, Julia A. Clarke, Allan J. Baker, Michele Clamp and Scott V. Edwards

All roads lead to regulation

Species from widely divergent taxa can experience similar changes in traits. What underlying genetic drivers cause these parallel changes remains an open question. Sackton *et al.* looked across groups of birds that have repeatedly lost flight, the ratites and tinamous, and found that there is convergence in the regulatory regions associated with genes related to flight, but not within the protein coding regions. Changes within these regulatory regions influenced limb development and may represent quick paths toward convergent change across taxa.
Science, this issue p. 74

ARTICLE TOOLS	http://science.sciencemag.org/content/364/6435/74
SUPPLEMENTARY MATERIALS	http://science.sciencemag.org/content/suppl/2019/04/03/364.6435.74.DC1
REFERENCES	This article cites 111 articles, 26 of which you can access for free http://science.sciencemag.org/content/364/6435/74#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service