THE UNIVERSITY OF CHICAGO

THE ECOLOGY, EVOLUTION, AND ENVIRONMENTAL HISTORY OF BIRDS IN NATURAL AND HUMAN IMPACTED ENVIRONMENTS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

COMMITTEE ON EVOLUTIONARY BIOLOGY

BY

SHANE GREGORY DUBAY

CHICAGO, ILLINOIS

AUGUST 2018

Contents

List of Figures

List of Tables

Acknowledgments

This work would not have been possible without the labor, support, and commitment of many people. First, Wu Yongjie and Carl Fuldner were true collaborators and friends. The work presented in this dissertation is just as much theirs as it is mine. My co-advisors, John Bates and Trevor Price fostered independence, while providing crucial support along the way. They have both greatly impacted the scientist I have become, for which I am grateful. I am also thankful for the guidance I received from Zachary Cheviron and Graham Scott, who both welcomed me into their communities, helping me make sense of my physiological data. I also thank Marcus Kronforst, who provided critical feedback along the way, which shaped the trajectory of my last two chapters. As a dual member of the Committee on Evolutionary Biology at the University of Chicago and The Field Museum, I have been fortunate to interact with a broad network of people that fostered my success as a graduate student, including Michael Coates, Carolyn Johnson, Elizabeth Eakin, Bonnie Brown, Connie Homan, Jeff Wisniewski, and Mary Johnson at the university, and Shannon Hackett, David Willard, Ben Marks, Josh Engel, Sushma Reddy, Mary Hennen, Tom Gnoske, and Corrie Moreau at the museum.

With respect to Chapter 1, I thank contributors past and present to natural history collections. Projects like this would not be possible without the commitment of individuals to collections, specifically: David Willard, Ben Marks, Sushma Reddy, Josh Engel, Shannon Hackett, and John Bates at The Field Museum; Janet Hinshaw and Benjamin Winger at the University of Michigan Museum of Zoology; and Steve Rogers at the Carnegie Museum of Natural History. David Willard and John Bates first introduced me and Carl Fuldner to “sooty” birds in collections. I also thank Julie Marie Lemon, Marissa Lee Benedict, and the Arts, Science

+ Culture Initiative at the University of Chicago, which funded this chapter. I thank Philipp Heck and Levke Kööp for help with SEM imaging.

Much of the work for my last three chapters was conducted in China, in which countless individuals poured time and energy into making these projects successful. I must first thank Lei Fumin, Qu Yanhua, Gang Song, Zhu Xiaojia, Gao Bin, and Cheng Yalin at the Institute of Zoology, Chinese Academy of Science who helped facilitate much of my research in China and provided a second academic home during my many visits to China. Furthermore, my research would not have been possible without the help, support, and kindness of the staff of Gongga Shan National Nature Reserve and Hailuoguo Forestry Bureau, specifically Zhou Huaming, Li Yugang, Mao Shaojun, Xie Qiang, and Ni Tianzhen. In addition, my field data would not have been collected without the help of Xin Chao, Joel Smith, Andy Hart Reeve, and Dylan Meyer. My research in China was generously supported by grants from the National Geographic Society/Waitt Fund, the American Philosophical Society, the American Museum of Natural History, the American Ornithologists' Union, the Society for the Study of Evolution, the Paulson Institute, and the Animal Behavior Society.

I have to express my gratitude to my colleagues in Graduate Students United at the University of Chicago, the American Federation of Teachers, and the American Association of University Professors, who have taught me what true community can look like. I am specifically grateful to: Benjamin Blanchard, Clara del Junco, Neville Eclov, Mathilde Gerbelli-Gauthier, Claudio Gonzales, Eyjólfur Guðmundsson, Rafadi Hakim, Gourav Khullar, Dallas Krentzel, Chaz Lee, Grant Macdonald, Peter Malonis, Natalia Piland, Supriya, and Audrey Williams with GSU; Candis Castillo, Sam Jordan, Tyler Kissinger, and Emma LaBounty with AFT; and Kira

Schuman with AAUP. I look forward to the day when the University's administration treats its graduate students and its other workers with respect and dignity.

Finally, there are a multitude of individuals to thank that positively impacted me personally and professionally during graduate school: Brett Aiello, Hussein Al-Asadi, Mark Bitter, Chloe Blackshear, Benjamin Blanchard, Natasha Bloch, Joao Capurucho, Carlos Cardenas, Sophia Carryl, Katie Collins, Johanna Colon, Chris Cooney, Jacob Cooper, Stewart Edie, Chad Eliason, Kristina Fialko, Victoria Flores, Peter Flynn, Ryan Fuller, Valentina Gomez, Swiss Nick Greber, Andre Green, Dave Grossnickle, Sebastian Heilpern, Daniel Hooper, David Jablonski, Jacob Johnson, Jon Kennedy, Evan Koch, Dallas Krentzel, Simon Lax, Michael Le Chevallier, Woo Chan Lee, Peter Lido, Chloe Nash, Grant Macdonald, Roberto Marquez, Darli Massardo, Pamela Martinez, Joel Mercado-Diaz, Liz Moyer, Ben Otoo, John Park, Cathy Pfister, Natalia Piland, Arvind Pillai, Joyce Pieretti, Supriya, Suresh Rana, Rick Ree, Maurico Santos, Mariah Scott, Carlos Servan, Mo Siddiq, Heather Skeen, Joel Smith, Peter Smits, Collin Soderberg-Chase, Tim Sosa, Jenn Stanley, Tom Stewart, Gabriel Vargas, Iuri Ventura, Chentao Wei, Brooke Weigel, Alex White, and Ben Winger. I have to express my gratitude to my parents, Paul and Corinne DuBay, who provided me the opportunity, confidence, and support to choose this path. Finally, Daniela Palmer has been a true partner, teaching me how real advocacy and academic pursuit can be balanced to build a more just and impactful academy. I am grateful to have shared much of this experience with her.

I did take the puzzle piece.

Introduction

The contents of this dissertation integrate a diverse set of disciplines, spanning biological and physical sciences, intact and human modified environments, and across time scales. Each chapter covers seemingly disparate topics, from atmospheric science and environmental history to behavior, physiology, and their roots in evolutionary biology. However, each chapter is firmly embedded in natural history, pushing disciplinary boundaries with the shared goal of understanding the consequences of environmental gradients, through time or space, on the organisms around us. Together, these projects leverage natural history to ask new questions about the ecology, evolution, and environmental history of birds in natural and human impacted environments.

In Chapter 1, we analyzed a time-series of black carbon deposited on >1,300 bird specimens to reconstruct historical levels of atmospheric black carbon in the US Manufacturing Belt. Black carbon is a major contributor to anthropogenic climate change, and historical emission inventories of black carbon are thus consequential for assessing past climate sensitivity and modeling future climate scenarios (Ramanathan and Carmichael, 2008; Hansen et al., 2000; Jacobson, 2000). However, due to a lack of coordinated atmospheric monitoring prior to the 1950s, current emission inventories suffer from a high degree of uncertainty for the early industrial era (Bond et al., 2013). We document that bird specimens provide an incidental record of atmospheric pollutants from which we have recovered historical emissions trends that predate the advent of coordinated air quality monitoring by over 70 years. Among the insights revealed through our analysis, we found that prevailing emission inventories underestimate black carbon emissions through the first decades of the 20th century, suggesting that the contribution of black carbon to past climate forcing may also be underestimated.

In this chapter, we've sought to capture the integrative nature of the study. While our key findings are rooted in atmospheric science, their implications naturally extend into biology, environmental science, and policy. For biologists, our project restores the severed link between museum specimens and the landscapes they once inhabited. For environmental policy advocates, our dataset provides a quantitative benchmark to assess policies that targeted air pollution. For atmospheric scientists our results generate an independent resource to validate and refine historical emission inventories, while also filling a major gap in atmospheric sampling records.

In Chapter 2, we leveraged field experiments to empirically test the capacity in birds to use human disturbance within their environment, a behavioral phenomenon that has fascinated natural historians for centuries (Buffon, 1771–1783; Lack 1948). Interspecific foraging associations are well documented in animals, such as birds following and foraging in association with army ant swarms and non-human mammals (Rand, 1953; Dean and MacDonald, 1981; Zhang and Wang, 2000; Kuniy et al., 2003; Beiseigel, 2007; King and Cowlishaw, 2008; Willis and Oniki, 1978). In rare cases birds can directly exploit human disturbance, but these examples are often limited to species that naturally occur in edge, open, or disturbed habitats. In this chapter, we discovered a capacity in insectivorous birds to exploit human disturbance in primary subalpine forest to gain novel foraging opportunities.

In Chapter 3, we explored physiology, behavior, and life history to ask how diverse functional demands on the avian flight muscle are reconciled in birds that experience different environmental and ecological pressures. As a classic multifunctional organ, the avian flight muscle is deeply integrated into avian life history and ecology, and must balance the diverse demands of powered flight and thermoregulation (Rosser and George, 1985; Marsh and Dawson, 1989; Scott et al., 2009; Segre et al., 2015; Wright et al., 2016; Dakin et al., 2018). In this

chapter we leveraged an intersectional framework in which we asserted that flight muscle phenotypes are a manifestation of the combined and balanced selective pressures that individuals experience as a result of multi-class identity (i.e. species, age, and sex). We found that age- and sex-structured variation in flight muscle phenotypes is associated with the distinct functional demands and selective pressures imposed on different demographics, revealing an underlying functional basis for two common life history strategies in birds: (1) age- and sex-structure migration and (2) delayed plumage maturation, in which young males retain female-like plumage to avoid the costly aggression with dominant adult males.

In Chapter 4, we experimentally tested how birds mitigate the physiological stress of shifting environments in seasonal mountains. Past works has uncovered the adaptive mechanisms that facilitate living at high elevations (reviewed in Storz et al., 2010), but this research largely ignores that many animals seasonally migrate in and out of high elevation environments. This behavior presents metabolic challenges for animals as they shift seasonally between disparate environments and elevations because physiological adaptions to high elevation environments are often maladaptive at low elevation, and vice versa (Projecto-Garcia et al., 2013). Through common garden experiments, in which we manipulated temperature and partial pressure of oxygen, we found that birds with higher degrees of seasonal movements exhibit increased physiological flexibility in traits that influence whole-organism metabolism, suggesting a mechanism to mitigate the physiological stress of shifting environments. This chapter provides empirical evidence for the role of metabolic flexibility in facilitating seasonal movements in and out of high elevation environments.

1 Bird specimens track 135 years of atmospheric black carbon and environmental policy[1]

1.1 Abstract

Atmospheric black carbon has long been recognized as a public health and environmental concern. More recently, black carbon has been identified as a major, ongoing contributor to anthropogenic climate change, thus making historical emission inventories of black carbon an essential tool for assessing past climate sensitivity and modeling future climate scenarios. Current estimates of black carbon emissions for the early industrial era have high uncertainty, however, because direct environmental sampling is sparse before the mid-1950s. Using photometric reflectance data of >1,300 bird specimens drawn from natural history collections, we track relative ambient concentrations of atmospheric black carbon between 1880 and 2015 within the US Manufacturing Belt, a region historically reliant on coal and dense with industry. Our data show that black carbon levels within the region peaked during the first decade of the 20th century. Following this peak, black carbon levels were positively correlated with coal consumption through midcentury, after which they decoupled, with black carbon concentrations declining as consumption continued to rise. The precipitous drop in atmospheric black carbon at midcentury reflects policies promoting burning efficiency and fuel transitions rather than regulating emissions alone. Our findings suggest that current emission inventories based on predictive modeling underestimate levels of atmospheric black carbon for the early industrial era, suggesting that the contribution of black carbon to past climate forcing may also be

[1] A version of this chapter has been published as: DuBay SG and CC Fuldner. 2017. Bird specimens track 135 years of atmospheric black carbon and environmental policy. Proceedings of the National Academy of Sciences 114: 11321-11326

underestimated. These findings build toward a spatially dynamic emission inventory of black carbon based on direct environmental sampling.

1.2 Significance statement

Emission inventories of major climate-forcing agents like black carbon suffer high uncertainty for the early industrial era, thereby limiting their utility for extracting past climate sensitivity to atmospheric pollutants. We identify bird specimens as incidental records of atmospheric black carbon, filling a major historical sampling gap. We find that prevailing emission inventories underestimate black carbon levels in the United States through the first decades of the 20th century, suggesting that black carbon's contribution to past climate forcing may also be underestimated. This study builds toward a robust, spatially dynamic inventory of atmospheric black carbon, highlighting the value of natural history collections as a resource for addressing present-day environmental challenges.

1.3 Introduction

Black carbon, the light-absorbing component of soot, is a complex carbonaceous aerosol that results from the incomplete combustion of organic matter, such as fossil fuels (Bond et al., 2013). Starting in the mid-19th century, cities within the US Manufacturing Belt—such as Chicago, Detroit, and Pittsburgh—experienced sharp rises in atmospheric soot due to their reliance on regional supplies of highly volatile soft, bituminous coal for manufacturing, domestic heating, and railway transportation (Stradling, 1999). By the late 19th century, the palls of coal smoke hanging over industrial cities galvanized early civic reformers, who fought urban smoke pollution as an unsightly nuisance, an economic inefficiency, and a public health concern tied to

respiratory illness and increased mortality (Stradling, 1999; Stern, 1982). These early, city-level efforts to mitigate atmospheric soot laid the groundwork for the modern environmental movement in the United States. While US cities no longer experience levels of atmospheric black carbon comparable to the historic peaks of the early 20th century, particle pollution remains a pressing public health and environmental issue in the United States and globally (Ramanathan and Carmichael, 2008; Shindell et al., 2012).

Black carbon has more recently become recognized as a major contributor to anthropogenic climate change (Ramanathan and Carmichael, 2008; Hansen et al., 2000; Jacobson, 2000). As such, historical emission inventories are consequential for understanding black carbon's effect on past climate and accurately modeling future climate scenarios. Estimates of black carbon emissions, however, have high uncertainty for the early industrial era (Bond et al., 2013), limiting our ability to use past emissions data to extract climate sensitivity. In the United States, efforts to measure concentrations of atmospheric soot were limited to sporadic city-level surveys before the mid-1950s (Ives et al., 1936), when federal legislation targeting air pollution gave rise to a coordinated national network for atmospheric monitoring (Stradling, 1999; Stern, 1982). As a result, our current understanding of atmospheric black carbon levels before midcentury in the US Manufacturing Belt is limited to anecdotal evidence and piecemeal records. Building accurate emission inventories of climate-forcing agents like black carbon remains a key step toward establishing a more rigorous understanding of how atmospheric pollutants affect climate.

Recent efforts to estimate historical black carbon emissions have used predictive models that combine fuel consumption data with emission factors, a variable that rates the efficiency of burning technologies (Novakov et al., 2003; Ito and Penner, 2005; Bond et al., 2007). Emission

inventories generated by these models have been instrumental in evaluating the contribution of atmospheric black carbon to climate change (Jacobson, 2004; Hansen et al., 2005; Shindell et al., 2009), but their power is contingent on the ability of emission factors to accurately capture changes in real-world burning efficiency over time. The robustness of predictive models can be independently evaluated by direct sampling data, such as the Greenland ice-core record (McConnell et al., 2007), which captures free-tropospheric emissions of black carbon from North America and stands as one of the few inventories based on a standardized, direct sampling metric of black carbon that extends back before the 1950s. The emission trends inferred from predictive models [such as the Speciated Pollutant Emissions Wizard (SPEW) database from Bond et al. (2007) generally mirror the Greenland ice-core record, indicating a rise in atmospheric black carbon in the late 19th and early 20th centuries associated with increased coal consumption, with emissions dropping to near preindustrial levels shortly after midcentury. While these contrasting methods achieve comparable results, there are inconsistencies between them: The ice-core record indicates a peak in black carbon concentrations in the first decade of the 20th century, while predictive models place this peak two decades later. Reconciling this disparity not only strengthens our understanding of environmental history and policy, but also holds important consequences for downstream climate analyses.

Here, we develop an alternative direct-sampling method for estimating historical trends in atmospheric black carbon by analyzing black carbon deposition on bird specimens collected within the US Manufacturing Belt over the past 135 y. In contrast to the Greenland ice-core record, our dataset recovers historical trends in atmospheric black carbon that are geographically localized. Our method therefore bypasses assumptions about the origin of atmospheric pollutants that are necessary to interpret the ice-core samples. As a direct sampling metric, our dataset also

bypasses the need to make assumptions about burning efficiency and technology shifts on which predictive models rely, providing an independent means for evaluating such models. By providing a more accurate, localized picture of historical trends in atmospheric black carbon, the results of the study yield a diverse set of implications that advance our understanding of human impacts on the physical and natural world, from assessing the impacts of black carbon on the environment to evaluating historical policies designed to clean up the air in some of the world's smokiest cities.

1.4 Results and discussion

1.4.1 Black carbon deposition on bird specimens

Birds accumulate black carbon and other particulate matter on their feathers from the surrounding environment. References to plumage discoloration appear in the scientific literature as early as the 1930s, which prompted a debate over whether darkened feathers had resulted from superficial soiling or changes to internal pigmentation. The latter would indicate industrial melanism, an evolutionary phenomenon in which darker phenotypes thrive within soot-filled environments (Hardy, 1937; Harrison, 1963). Using scanning electron microscopy (SEM), we confirmed that darkened breast feathers from specimens collected in the early 20th century are indeed covered in black carbon particles (Figures 1 and 2). Given that most bird species undergo at least one annual molt to replace body feathers, black carbon deposition on specimens functions as an environmental sample from the year in which each bird was collected.

We measured reflectance values from a time series of bird specimens that were collected within the US Manufacturing Belt between 1880 and 2015. We defined the US Manufacturing

Figure 1 | Comparison of two Field Sparrows (*S. pusilla pusilla*), one from 1906 and one from 1996. (*Lower*) SEM micrographs of belly feathers plucked from the specimens in *Upper*. SEM micrographs of the Field Sparrow from 1906 show black carbon aggregates composed of small sphericals [a detailed description of black carbon morphology with microscopy images can be found in Bond et al. (2013)]. The feather from the 1996 specimen lacks black carbon deposition. Both specimens were collected during spring months in the vicinity of Chicago. SEM images were made with a Tescan LYRA3 field emission microscope with secondary electron (SE) detection and an acceleration voltage (HV) of 3.0 kV. Feather samples were carbon-coated before imaging.

Belt as the states of Pennsylvania, Ohio, Indiana, Michigan, Illinois, and Wisconsin because these states were historically dense with industry and traditionally relied upon soft, bituminous coal, which emits greater quantities of particulate matter than the harder, anthracitic coal more prevalent in the Eastern United States (Hartshorne, 1936; Bond et al., 2004) (Figure 3 shows localities of specimens used in study). We measured reflectance from the breast and belly

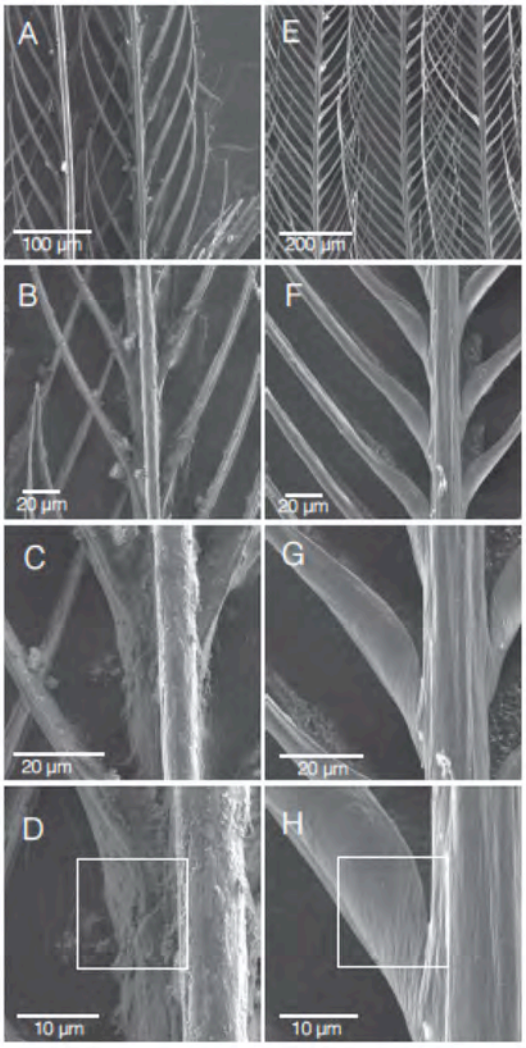

Figure 2 | Additional SEM micrographs, taken at different magnifications, from the Field Sparrows (*S. pusilla pusilla*) in figure 1. *A–D* are from the soiled 1906 specimen. *E–H* are from the clean 1996 specimen. The micrographs for each specimen are progressively higher in magnification. The white boxes in *D* and *H*outline the areas shown in Figure 1.

feathers of each specimen. Reflectance is a photometric that describes the proportion of light reflected from a surface and is expressed as a percentage value ranging from 0% (pure black) to 100% (pure white). Black carbon, defined as the primary light-absorbing component of soot, has low reflectance properties, allowing us to quantify the relative "sootiness" of each specimen as a function of reflectance. These data establish a relative estimate of ambient concentrations of

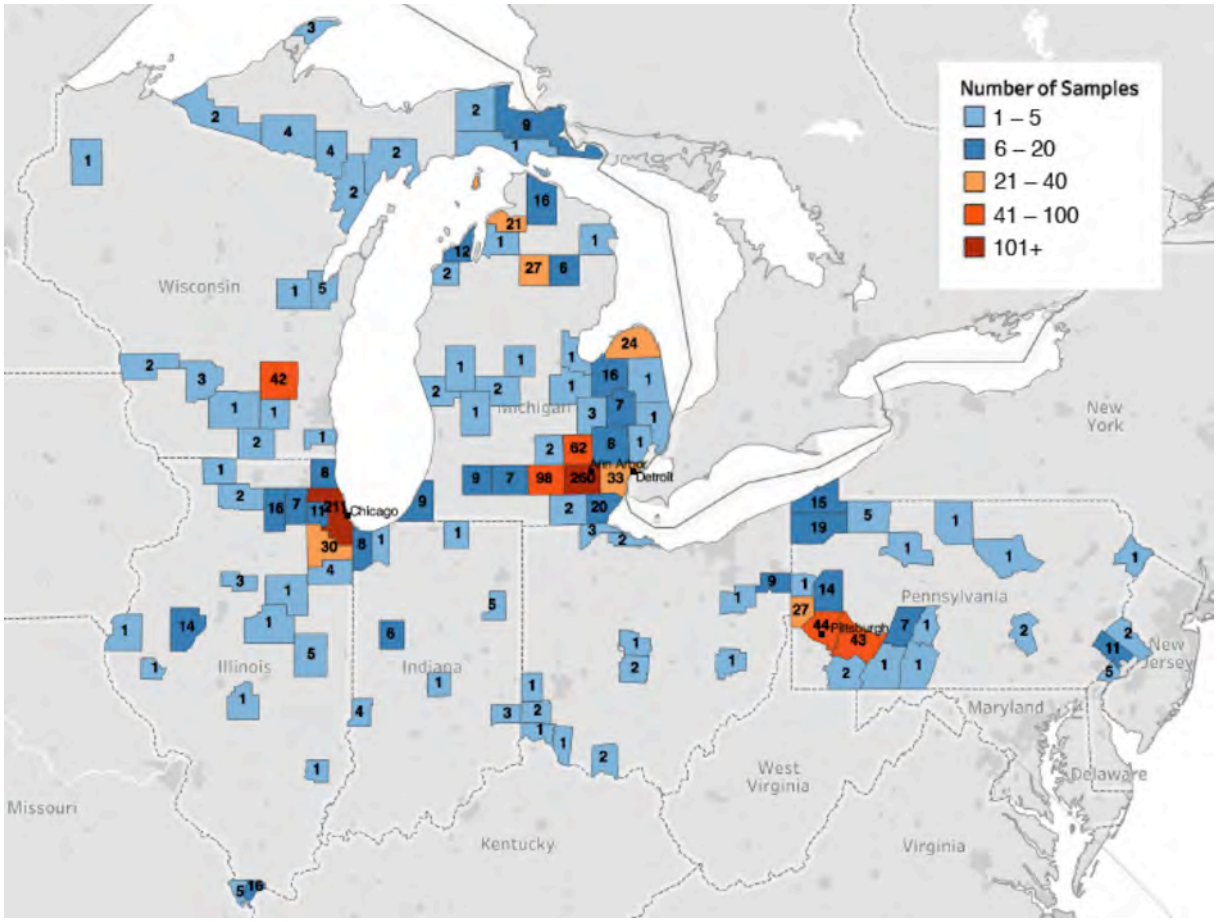

Figure 3 | Map showing the collection localities for 1,345 of 1,347 specimens used in this study. The remaining two specimens lack county locality data. Counties are shaded based on the density of sampling within the county. The number of specimens from each county is printed within each county.

black carbon based on a standardized, direct sampling metric that extends back >70 y before continuous, multicity air-monitoring networks were in place (Stradling, 1999; Stern, 1982).

We imaged 1,347 birds from five species collected within our geographic range, comprising >95% of available specimens with adult plumage at three major natural history museums. A small number of specimens were excluded because of preparation concerns or missing collection data. Starting from a list of species that breed in the US Manufacturing Belt, we selected species that have naturally light, uniform breast and belly coloration to maximize the signal strength while allowing for comparisons across species (Figures 1 and 4). These species

included the Field Sparrow (*Spizella pusilla pusilla*), Grasshopper Sparrow (*Ammodromus savannarum pratensis*), Eastern Towhee (*Pipilo erythrophthalmus erythrophthalmus*), Horned Lark (*Eremophila alpestris pratensis*), and Red-headed Woodpecker (*Melanerpes erythrocephalus*) (Dataset S1 associated with DuBay and Fuldner [2017] provides a list of vouchered specimens). Given that differences in plumage are often used to define subspecies classifications, we restricted sampling to a single subspecies per species. Each specimen was digitally photographed under standard lighting conditions, and reflectance measurements were taken from a uniform patch on the ventral side.

Each species in our sample undergoes an annual molt beginning in late summer that can last through the fall (Pyle, 1997), which replaces soiled plumage. The absence of discoloration on freshly molted fall birds further confirmed that soot accumulation, and not industrial melanism, was primarily responsible for darkened plumage. Since molt patterns vary by individual, specimens sampled during these annual molting periods included a mix of fresh and soiled birds (Figures 5 and 6), interfering with an accurate signal. The annual molt for each species thus had to be accounted for to recover an accurate trend in black carbon levels. We initially determined molting periods based on Pyle (1997), which we corroborated by examining monthly variances within our dataset (Figure 5). For all species, specimens collected in the months of September–November were excluded from the final dataset; for two species, specimens from August were also removed. From the 1,347 specimens initially sampled, 250 fell within the defined molting periods and were excluded from final analyses, leaving 1,097 usable samples within the nonmolting monthly ranges. Historically, black carbon levels are highest during winter months (Ives, 1936; McConnell et al., 2007), such that the period of heaviest accumulation on feathers occurs directly after the molt. This seasonal coincidence likely

obscures a monthly pattern of accumulation, but further work is needed to fully understand how

Figure 4 | Comparisons of old and young specimens for the four species pairs not shown in Figure 1. (*A*) Grasshopper Sparrows (*A. savannarum pratensis*) from 1907 (*Upper*) and 1996 (*Lower*). (*B*) Horned Larks (*E. alpestris pratensis*) from 1904 (*Upper*) and 1966 (*Lower*). (*C*) Eastern Towhees (*P. erythrophthalmus erythrophthalmus*) from 1906 (*Upper*) and 2012 (*Lower*). (*D*) Red-headed Woodpeckers (*M. erythrocephalus*) from 1901 (*Upper*) and 1982 (*Lower*).

birds seasonally accumulate and retain soot. Since no month-to-month trends were apparent within the designated nonmolting months, specimens were placed into groups organized by year (Figure 5).

To link reflectance data to black carbon levels for a single year, it had to be established that black carbon accumulation occurred before collection. Multiple lines of evidence indicated that the black carbon accumulated on bird specimens originated from the environment while the birds were alive and not from posthumous soiling or discoloration that occurred while being stored in a collection:

(1) Since posthumous soiling would accrete continuously, if soiling had occurred over time in storage, it would not have been possible to observe seasonal differences, and any monthly trends that result from the annual molting cycle would have been erased or vastly diminished, particularly in older specimens. We found that consistent numbers of birds collected during the fall were much cleaner in a given year, indicating freshly molted individuals (Figures 5 and 6). These patterns were observable even among birds that had been in the same collections as soiled birds, stored together since the time of collection.

(2) We conducted a visual survey of bird specimens collected outside the US Manufacturing Belt from other parts of the United States or from less industrialized countries during our 135-y sampling period. If posthumous soiling had occurred within our sample, we would have expected specimens collected in these nonindustrialized regions to have exhibited comparable levels of soiling to those in our sample, which we did not find. A visual example of this evidence can be seen in Figure 7, which shows five Horned Larks collected in Illinois and five Horned Larks collected along the western coast of North America. All 10 birds were collected during nonmolting months between 1903 and 1922, a period in which consistently high levels of black carbon deposition were found on bird specimens collected within the US Manufacturing Belt.

(3) If specimens in our sample accumulated black carbon from sitting in museum collections, we would have expected specimens to have soiled ventral sides and cleaner dorsal sides because they generally rest in drawers with their breast and belly facing up.

The dorsal side of the specimens would thus have been protected from soot precipitate.

We found, however, that both sides of specimens exhibited soiling (Figure 8).

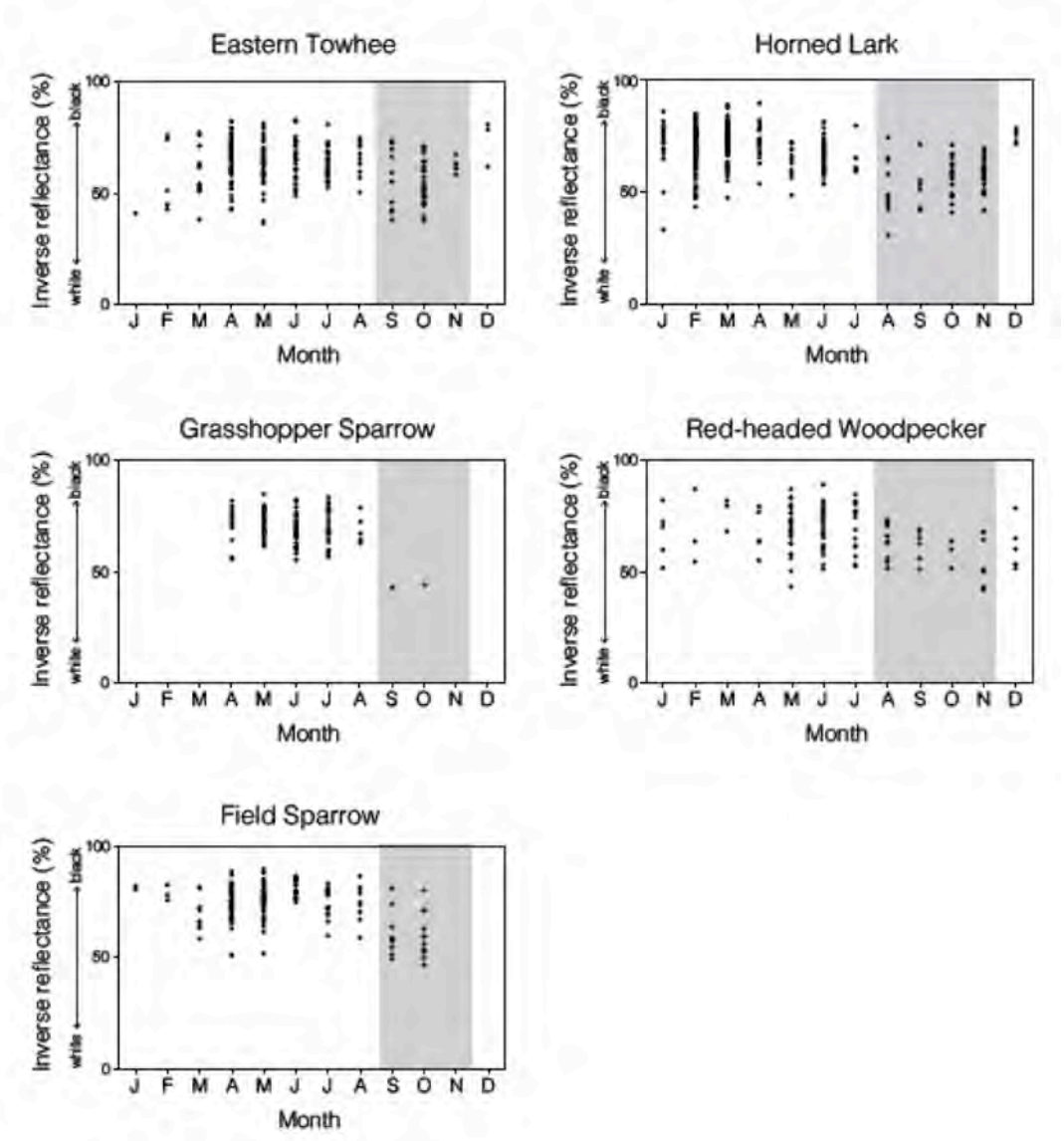

Figure 5 | Monthly trends in black carbon deposition for each species before 1950. Inverse reflectance is reported rather than reflectance to express drops in black carbon emissions, which register as increased reflectance values. The shaded areas are the months excluded from final analyses for each species, which are applied to all years. Sampling is sparse for Grasshopper Sparrows and Field Sparrows in the US Manufacturing Belt during fall and winter months because these species predominately migrate out of the region.

To integrate data across species, we normalized inverse raw reflectance values by calculating z scores within each species set. The z score for each specimen is defined as: (inverse raw reflectance value for an individual – inverse mean reflectance value for the species)/(SD of inverse reflectance for the species). Inverse reflectance was used rather than reflectance to better visualize drops in black carbon deposition, which register as an increase in reflectance and a decrease in inverse reflectance. With the normalized dataset, we estimated a trend in black carbon deposition through time using a generalized additive model (GAM) in the mgcv R package (Wood, 2011).

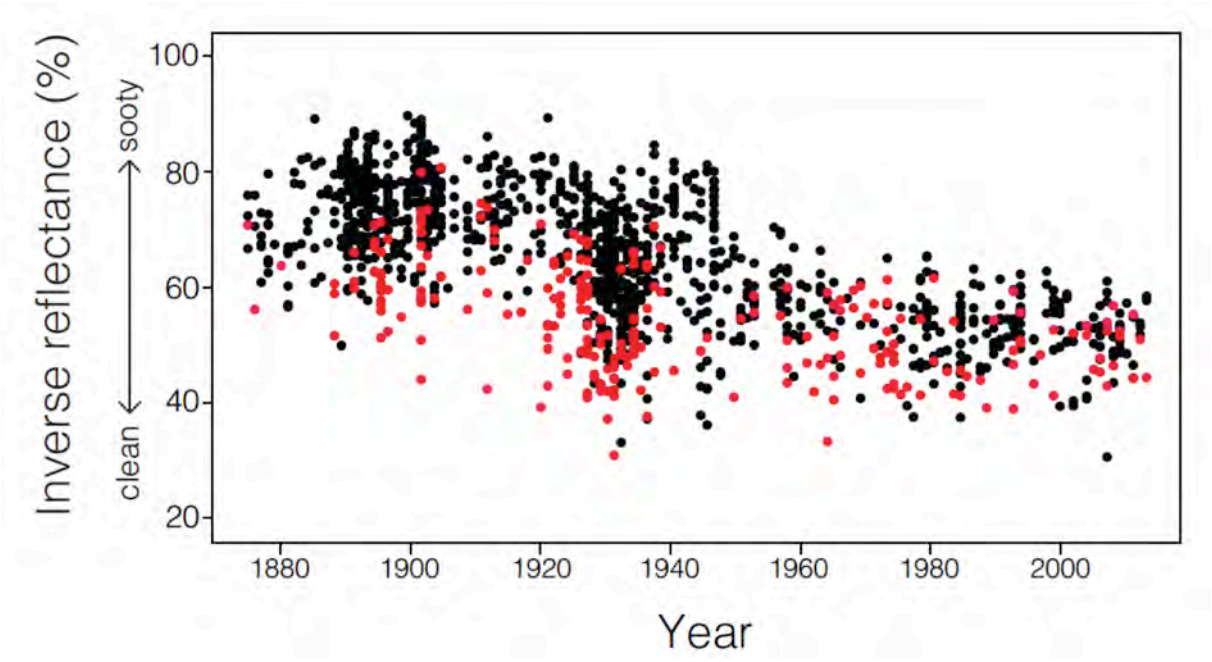

Figure 6 | Black carbon deposition for all 1,347 individuals sampled for this study, showing that specimens from molting months (red points) are substantially cleaner than specimens from winter–summer (black points). Black points are individuals included in the final dataset (n = 1,097), and red points are individuals from molting months that were excluded in final analyses (n = 250) (Figure 5). Inverse reflectance is reported rather than reflectance to express drops in black carbon emissions, which register as increased reflectance values. Before 1950, individuals from molting months are noticeably cleaner than individuals from the rest of the year, warranting the exclusion of specimens from these months for all years.

Figure 7 | Ten Horned Larks (*E. alpestris pratensis*) at The Field Museum, showing that specimens collected in nonindustrial regions do not exhibit comparable levels of soiling to birds collected within the US Manufacturing Belt. The five specimens in *Left* were collected in Illinois, inside the US Manufacturing Belt. The five specimens in *Right* were collected along the western coast of North America, outside of the US Manufacturing Belt. All 10 specimens were collected during nonmolting months (January–April) between 1903 and 1922.

1.4.2 Historical trends in black carbon and environmental history

Our results show that black carbon deposition on bird specimens peaked during the first decade of the 20th century (Figure 9). This peak is consistent with the Greenland ice-core record, anecdotal accounts, and surveys conducted in Chicago and Pittsburgh during the second decade of the 20th century, all of which indicate modest improvements in air quality after 1910 despite a steady increase in overall coal consumption (Davidson, 1979; Bachmann, 2007) (Figure 9). Black carbon accumulation on specimens remained high through the 1920s. The first precipitous drop in deposition coincided with a temporary reduction in overall coal consumption during the

Figure 8 | Images of the dorsal side of specimens from Figures 1 and 4. These images, paired with Figures 1 and 4, show that even soiling appears over the entire bird, indicating that the soiled birds in our sample acquired black carbon from the environment while alive. (*A*) Field Sparrows (*S. pusilla pusilla*) from 1906 (*Upper*) and 1996 (*Lower*). (*B*) Grasshopper Sparrows (*A. savannarum pratensis*) from 1907 (*Upper*) and 1996 (*Lower*). (*C*) Horned Larks (*E. alpestris pratensis*) from 1904 (*Upper*) and 1966 (*Lower*). (*D*) Eastern Towhees (*P. erythrophthalmus erythrophthalmus*) from 1906 (*Upper*) and 2012 (*Lower*). (*E*) Red-headed Woodpeckers (*M. erythrocephalus*) from 1901 (*Upper*) and 1982 (*Lower*).

Great Depression, which rebounded during World War II. A second and lasting drop in deposition began in the postwar period, with coal consumption declining as other fossil fuels gained traction (Figure 9 and 10). The second drop in black carbon deposition continues to present day, despite subsequent increases in coal consumption (Figure 9). This sharp drop in

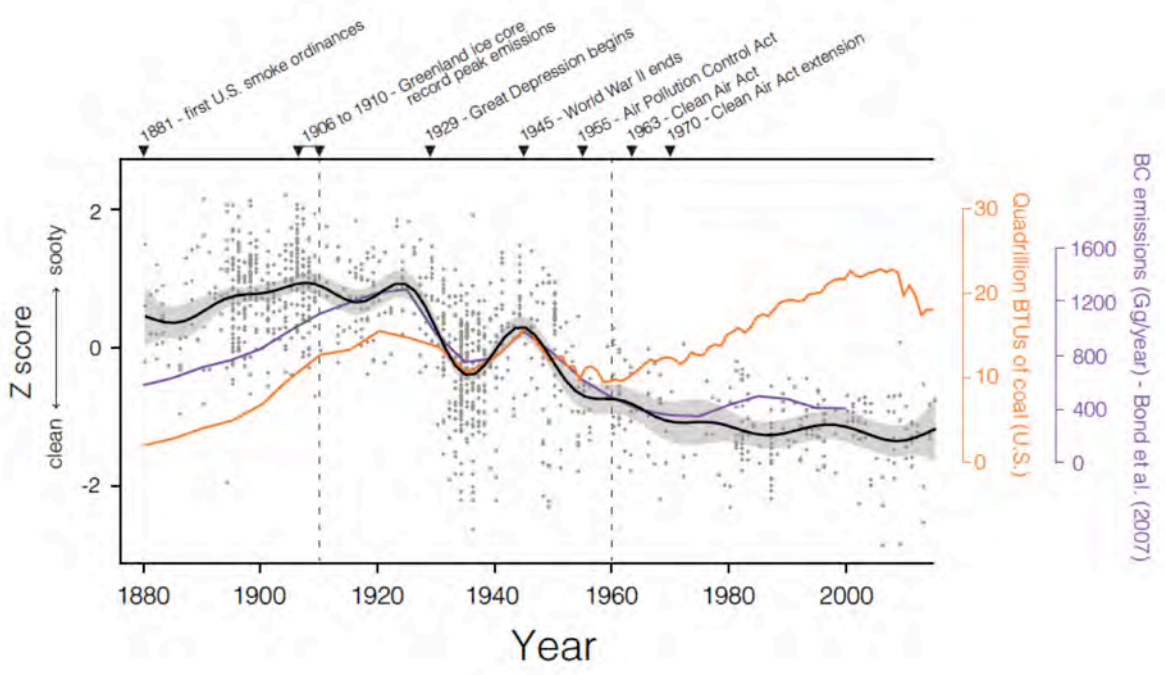

Figure 9 | Black carbon deposition on specimens of five bird species from the US Manufacturing Belt, collected between 1880 and 2015. Each point represents the z score for an individual specimen ($n = 1{,}097$) based on the inverse raw reflectance value taken from its breast and belly feathers. The black line is a GAM ($k = 20$) with 95% confidence limits (indicated by the shaded area), determined from the individual specimens (details on how k was determined can be found in 1.6 [Materials and Methods] and Figure 14. Figure 15 shows species-specific trends). The orange line is consumption for coal in the United States expressed in British thermal units (BTUs) (US Energy Information Administration). Before 1950, coal consumption data are available in 5-y intervals. After 1950, coal consumption data are yearly. The purple line shows estimates of total US black carbon (BC) emissions from Bond et al. (2007), which uses fuel consumption data and emission factor data to generate a historical emission inventory. The dashed line at 1910 denotes the progressive shift in cities within the US Manufacturing Belt from prosecuting to educating emissions violators. The dashed line at 1960 denotes the approximate moment after which black carbon emissions become decoupled from coal consumption.

atmospheric black carbon is consistent with the midcentury drop recovered in the Greenland ice-core record and predicted in Bond et al. (2007). The Greenland ice-core record also identifies forest fires as a historical source of atmospheric black carbon, but the associated emissions were comparatively low through our study period (McConnell et al., 2007), such that their influence on our trend would be minimal.

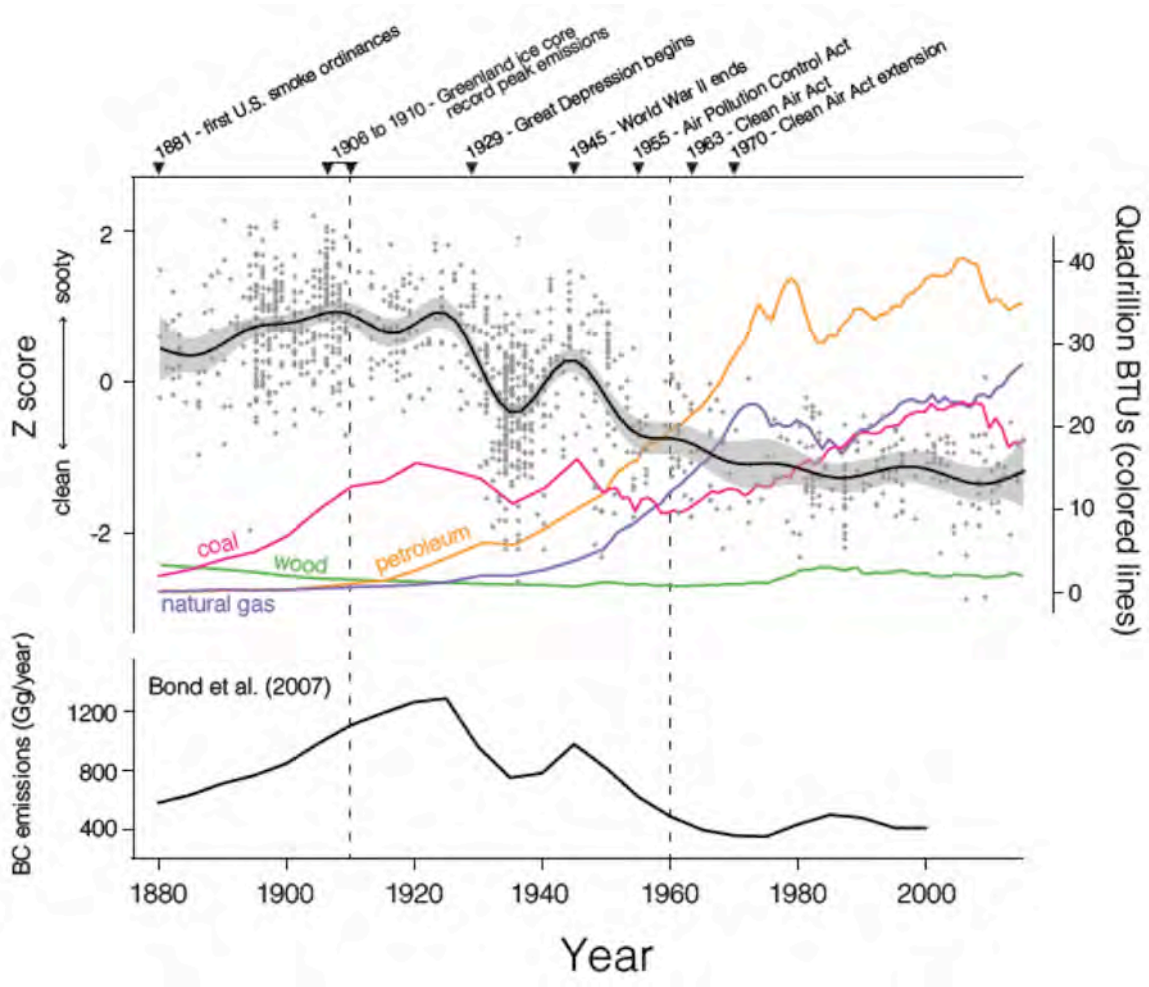

Figure 10 | Black carbon deposition on specimens (five bird species) from the US Manufacturing Belt, collected between 1880 and 2015. Each point represents the z score for an individual specimen (n = 1,097), based on the inverse raw reflectance value taken from its breast and belly feathers. The black line in *Upper* is a GAM (k = 20) with 95% confidence limits (indicated by the shaded area), determined from the individual specimens (details on how k was determined can be found in 1.6 [Materials and Methods] and Figure 14. Figure 15 shows species-specific trends). The colored lines are consumption trends for biofuels and fossil fuels expressed in British thermal units (BTUs) (US Energy Information Administration). Before 1950, fuel consumption data are available in 5-y intervals. After 1950, fuel consumption data are yearly. *Lower* shows estimates of total US black carbon (BC) emissions from Bond et al. (2007), which uses fuel consumption data and emission factor data to generate a historical emission inventory. The dashed line at 1910 denotes the progressive shift in cities within the US Manufacturing Belt from prosecuting to educating emissions violators. The dashed line at 1960 denotes the approximate moment after which black carbon emissions becomes decoupled from coal consumption.

From 1880 to 1910, black carbon deposition is not strongly correlated with coal consumption (Figures 9 and 11a). Our results show high black carbon levels with only a slight upward trend across these three decades, despite a sharp increase in coal consumption over the same period (Figure 9). While ambient concentrations of black carbon hit a historical peak in this period, the relatively constant levels of black carbon on specimens suggest that period reforms and antismoke initiatives registered a modest mitigating effect, reducing the growth in black carbon levels relative to consumption. Toward the end of the 19th century, civic reformers organized to combat urban smoke pollution (Stradling, 1999; Stern, 1982). In 1881, Chicago and Cincinnati passed the first municipal smoke ordinances in the United States. These laws focused on regulating emissions, but they exempted residential burning and proved difficult to enforce

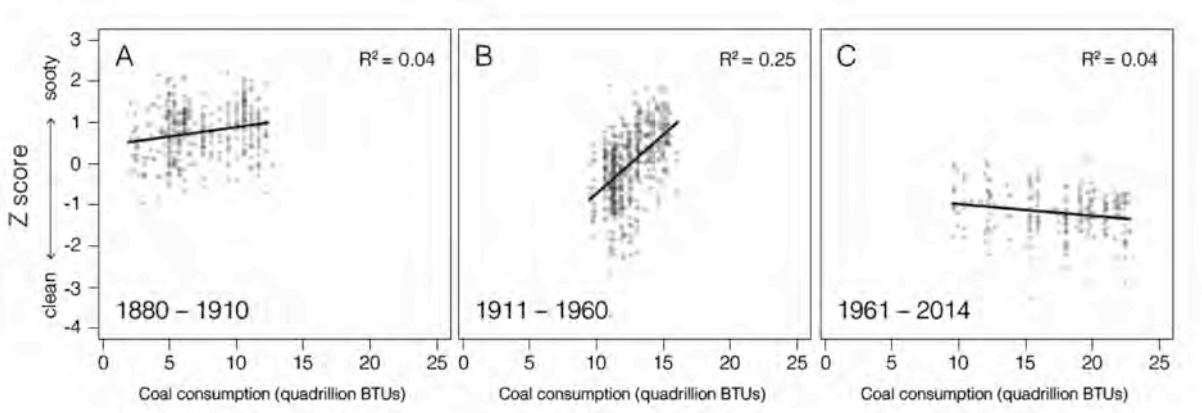

Figure 11 | Black carbon deposition on specimens plotted against US coal consumption for the three time periods defined by the dashed lines in Figure 9. (*A*) Between 1880 and 1910, black carbon deposition is not correlated with coal consumption. Black carbon deposition is high and remains relatively constant, trending upward only slightly as consumption increases sharply. (*B*) Between 1911 and 1960, black carbon deposition and coal consumption are positively correlated. (*C*) After 1960, black carbon deposition is decoupled from coal consumption. As consumption increases, black carbon deposition remains low. Before 1950, fuel consumption data are only available in 5-y intervals. We thus interpolated consumption values between points to estimate consumption for the year in which each specimen was collected before 1950. After 1950, yearly fuel consumption data are available.

(Stradling, 1999). By 1910, most cities in the US Manufacturing Belt had established municipal departments specifically devoted to smoke abatement (Stern, 1982), and reform efforts began to expand beyond litigation to encompass education and technology-based solutions (Grinder, 1980).

From 1910 to 1960, black carbon deposition was positively correlated with trends in coal consumption (Figures 9 and 11b). During this period, reform efforts focused on curbing emissions through education and by promoting technologies to burn coal more efficiently. Despite these concerted efforts to rein in soot emissions, our data confirm that the overall concentration of atmospheric black carbon remained tied to coal consumption through midcentury (Figure 9). Our results suggest that efforts to regulate emissions directly were largely ineffective at reducing overall levels of atmospheric black carbon.

During the second half of the 20th century, black carbon deposition on specimens became decoupled from coal consumption (Figures 9 and 11c). As consumption began to rise again in the postwar period, atmospheric black carbon continued to decline. This decoupling can be explained by a new approach to city-level legislation, which targeted the types of fuel consumed in both domestic and industrial sectors rather than regulating emissions directly. New regulations addressed the distribution of bituminous coal and mandated that consumers of soft coals use mechanical stokers or switch to smokeless fuels (Stradling et al., 1999). These reforms effectively eliminated bituminous coal as a fuel source from residential furnaces, which are estimated to have produced over half of black carbon emissions during the early 20th century (Novakov et al. 2003). The success of these new regulations was contingent upon providing economically viable fuel alternatives. Following a successful model implemented in St. Louis in 1940, Pittsburgh began subsidizing harder, low-volatile coal for domestic use in 1946 (Davidson,

1979). St. Louis had seen the benefits of this new approach almost immediately, experiencing an 83.5% decrease in the total hours of thick atmospheric soot during the winter of 1940–41 (Tarr and Zimring, 1997). Following WWII, US cities also began transitioning to alternative fuel sources, specifically petroleum and natural gas (Figure 10). By 1950, 66% of households in Pittsburgh were heated with natural gas, up from 17% a decade earlier (Tarr, 1981). Around the same time, electricity production in the United States shifted away from scattered, coal-powered steam boilers to centralized power plants (Platt, 1991). While these plants were more efficient, they drove the steady rise in coal consumption in the second half of the 20th century as they met the increasing demands for electricity. Together, the increased availability of fuel alternatives and the centralization of power production account for the decoupling of coal consumption from black carbon deposition on specimens. While soot mitigation in the United States took decades to achieve, the solutions proved to be relatively straightforward: Regulate the types of fuel consumed and promote affordable alternative fuel options.

With black carbon levels declining by midcentury (Figure 9), the United States entered a new era of air pollution and environmental policy. Decades of research and activism aimed at mitigating soot pollution culminated in the passage of the Air Pollution Control Act of 1955, the first federal air pollution legislation in the United States. This act did not regulate or control pollution levels, but directed money toward research into air pollution, helping to establish a coordinated, national network to monitor air quality. In 1963, the first incarnation of the Clean Air Act established federal limits on a variety of atmospheric pollutants, but by then, high levels of atmospheric black carbon had already receded (Figure 9).

1.4.3 Black carbon levels exceed predictive model estimates

Our results suggest that black carbon levels were higher at the start of the 20th century than estimates generated from predictive models (Novakov et al., 2003; Bond et al., 2007). While Bond et al. (2007) considered black carbon emissions on a national scale, our studies are largely comparable since their estimates for the United States are driven by bituminous coal, which was disproportionately consumed within the Manufacturing Belt (Stradling, 1999; Bond et al., 2004). Between 1880 and 1910, we recovered black carbon levels that were higher than the one estimated by Bond et al. (2007) (Figures 9 and 12a), a finding corroborated by the Greenland ice-

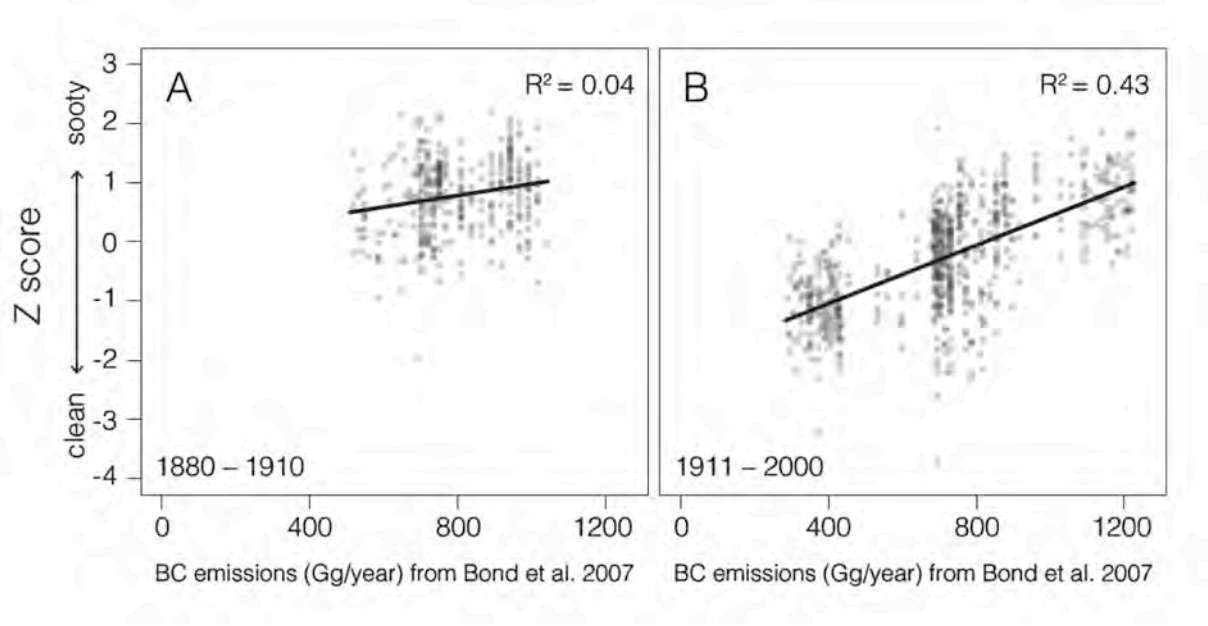

Figure 12 | Black carbon deposition on specimens plotted against black carbon (BC) emissions estimates from Bond et al. (2007) for the three time bins defined in Figure 9. The second two time bins (1911 to 1960 and 1961 to 2014) are combined to illustrate the strong correlation across both intervals. (*A*) Before 1910, we recovered relatively constant, high levels of black carbon deposition on specimens, while Bond et al. (2007) estimated a sharp rise in black carbon emissions. (*B*) After 1910, black carbon deposition is positively correlated with black carbon emissions estimates from Bond et al. (2007). Our results independently recovered similar trends in atmospheric black carbon. Bond et al. (2007) report BC emissions in 5-y intervals. We thus interpolated emissions values between points to estimate values for the year in which each specimen was collected.

core record. Our results are consistent with the peak concentrations in the ice-core record during the first decade of the 20th century, but we recovered higher relative concentrations between 1880 and 1900. This discrepancy could be explained if certain types of particles precipitated from the atmosphere before reaching Greenland.

The lower estimates recovered by Bond et al. (2007) are likely explained by the lack of reliable emissions data from the early industrial period used to parameterize predictive models. As Bond et al. (2007) acknowledge, this lack of data introduces high uncertainty into their model. Additionally, inconsistent burning practices and technology shifts across and within emitting sectors are difficult to account for in emission factors. Bond et al. (2007) assumes that the burning efficiency for a given technology remains constant, without accounting for incremental improvements in operating procedures through time. Operating efficiency became a key target of reform and education efforts after 1910, and our data suggest that these efforts did in fact register a modest mitigating effect. Emission factors in predictive models are thus limited in their ability to reflect real-world burning efficiency because of the inherent difficulties in quantifying variables like operating efficiency and asymmetric technology shifts throughout a given region or sector. Since our measurements are based on direct sampling of ambient concentrations, we are able to bypass assumptions about efficiency and technology implementation. However, translation of our measurements to emissions is not straightforward since, as with the ice-core record, the relationship between ambient concentrations and emissions depends on meteorological factors, which may have changed over time.

While our results show that current predictive models likely underestimate levels of atmospheric black carbon for the early industrial era, as consumption and emission factor data become more reliable through the 20th century, our results are positively correlated with

predictive models (Figures 9 and 12b). After 1910, we recovered a trend in atmospheric black carbon that is strikingly consistent with the trend produced by Bond et al. (2007) (Figures 9 and 12b). This result suggests that predictive models effectively recover emissions when sufficient data exist to parameterize the model. Our study thus provides support for the power of predictive modeling methods, while also indicating that black carbon emissions in the United States at the outset of the 20th century were higher than current estimates. This finding suggests that the climate-forcing effects of black carbon may also be underestimated for this period.

1.4.4 Building a usable emission inventory

A limitation of our current method is that reflectance values from specimens track relative trends in black carbon concentrations rather than recovering mass concentrations of atmospheric black carbon. Calibrating reflectance to a standard unit of mass concentration represents a next step toward building a usable, spatially dynamic emission inventory. Black carbon levels in Asian cities like Beijing and Delhi resemble those of the US Manufacturing Belt of a century ago (Bond et al., 2007), and we now have precise methods for measuring black carbon mass concentrations. By comparing contemporaneous measurements taken from these Asian cities with specimens collected from the same locations, our pre-1950 reflectance values could be calibrated to derive historical mass concentrations of black carbon from our sample. Once mass concentrations have been established, it would then become possible to estimate overall emissions for the region, although this step introduces additional challenges. Ambient concentrations and emissions are related yet distinct measures, and the translation between the two requires consideration of local topography and meteorology.

Additional insights are discoverable through a more thorough material analysis of black carbon on birds. The size and shape of black carbon particles and aggregates define their optical properties and climate-forcing effects (Bond et al., 2013; Koch et al., 2009), and thus knowing the historical size distribution of black carbon particles is critical for evaluating their climate impacts and building a usable emission inventory. Robust datasets of size distribution of black carbon particles, however, are similarly lacking before the 1950s, and these data are likewise difficult to predict with any certainty because particle size is dependent on a number of interacting variables, including the chemical composition of the fuel source, along with the technology and operating procedures used to burn the fuel. By analyzing black carbon deposits on bird specimens for their physical properties, the size distribution from the early industrial era could be directly estimated for a given year and locale. These data would be invaluable for developing more robust emission inventories of atmospheric black carbon.

1.5 Conclusions

This research highlights the unexpected ways in which museum materials can yield insights about the physical and natural world and help address present-day environmental challenges. Natural history collections are powerful resources for tracking environmental pollutants through time (Hickey and Anderson, 1968; Thompson et al., 1998) because specimens provide durable snapshots of the past environments from which they were drawn. For this study, bird specimens provided an incidental record of atmospheric black carbon from a period before standardized methods and coordinated systems for assessing air quality were in place. We focused on the US Manufacturing Belt because of its historical importance as a polluting region, but our dataset can naturally be expanded to encompass other regions with long industrial histories, such as Western

Europe. Natural history collections thus represent a unique resource for exploring past environments and environmental history.

For the purpose of this study, we used bird specimens as a direct sampling metric to assess historical concentrations of black carbon, which we used in turn to evaluate past environmental policy. Our study, however, also highlights the impact of environmental pollution on wildlife. Our samples show that black carbon particulate covered the landscape along with its living inhabitants. Black carbon accumulation on birds has potential implications for evolutionary pathways because plumage is fundamental in avian displays and signaling. Birds use their plumage to attract mates, defend territories, and/or camouflage themselves within the landscape to escape detection from predators. What happens when bright, sexually selected plumage patches are coated in soot, obscuring plumage signals that have evolved over hundreds of thousands of years? What are the consequences of black carbon deposition for visual predators when animal prey coloration is homogenized with the surrounding environment? How black carbon deposition on feathers has impacted signaling within and among species remains an open question.

1.6 Materials and methods

Reflectance has long been used as an efficient and reliable metric in atmospheric sampling (Penner and Novakov, 1996). For the purposes of this study, we were interested in deriving relative ambient concentrations from black carbon deposition on bird feathers. Since black carbon is defined by its light-absorbing properties, trends of black carbon deposition on specimens can be quantified as a function of the reduction in reflectance relative to unsoiled

specimens. We adapted photography methods from Stevens et al. (2007) and McKay (2013) to quantify the reflectance of each specimen.

1.6.1 Photographing specimens

Specimens were imaged with a mirrorless interchangeable lens camera (Sony a7R II) paired with a native 55 mm lens (Sonnar T* FE 55 mm F1.8 ZA), positioned at a fixed height of 72 cm over a self-contained light box (MK Digital Direct Photo-e-Box BIO) outfitted with 28-W continuous full-spectrum fluorescent bulbs (6,500 K, 84CRI) run through 120-V AC 60-Hz electronic ballasts. Specimens were illuminated by using top, side, and back bulbs in the light box, omitting the bottom (stage) bulbs and supplemental LED bulbs to ensure an even distribution of diffuse light from a single illuminant type source. At the beginning of each imaging session, the lighting elements were turned on and allowed to warm up for 20 min before shooting. Specimens were oriented so that the target area on the breast was positioned at the center of the camera's field of view. The light box was fully enclosed during each exposure, except for a rectangular aperture on the top, sized to fit the camera's field of view. Overhead lighting was turned off in each of the shooting locations, and windows were covered to further reduce ambient light leakage.

The images were captured in 14-bit uncompressed raw format and analyzed by using RawDigger software (Version 1.2.11), which provides access to raw data directly recorded by the digital camera's CMOS sensor. Analyzing the raw sensor data directly enabled us to bypass the linearization step described by Stevens et al. (2007), and McKay (2013), since the raw values have not been altered by nonlinear gamma encoding algorithms that are introduced when raw sensor data are converted into conventional image formats, such as JPEG or TIFF (Westland et al., 2012). Before shooting, we tested the linearity of the camera's CMOS sensor following the

procedure outlined in Stevens et al. (2007) and we found that the sensor provided a linear response over the entire dynamic range (Figure 13).

Exposure settings (shutter speed, aperture, and ISO) were optimized through a series of trials using reflectance standards. We conducted trials using four types of reflectance standards, including the XRite ColorChecker Passport (8-step), QPcard 101 (3-step), Labsphere Spectralon Diffuse Reflectance Standards (10 reference targets), and Munsell Neutral Value Scale matte finish (31-step). We found that each standard provided comparable results, but we selected the Munsell Neutral Value Scale as our primary standards because it was relatively affordable, provided the largest number of reference points, and included published reflectance percentages printed directly on the cards for easy reference. To determine exposure settings, we analyzed trial images in RawDigger with a goal of maximizing the dynamic range (defined as the distance between minimum and maximum light intensities) without introducing signal clipping on any of the color channels (R-G-B-G2), which occurs when certain clusters of pixels fall outside of the dynamic range due to overexposure (saturation). It is essential to refer to the raw data when assessing whether signal clipping has occurred, since the channel-specific histograms on many digital cameras' displays incorporate gamma-encoding algorithms that make it difficult to tell whether signal clipping has actually occurred. Exposure settings maximizing dynamic range will often indicate overexposed areas on the camera's built-in displays, when no signal clipping in the raw file has taken place.

The ISO was set to 100 to ensure a limited amount of digital noise. Based on the trials, an aperture of f/16 was chosen to minimize optical vignetting (light falloff), which is introduced at lower focal ratios, while providing a depth of field that would ensure that the target area appeared in focus for all specimens, which varied in height due to differences in natural size and

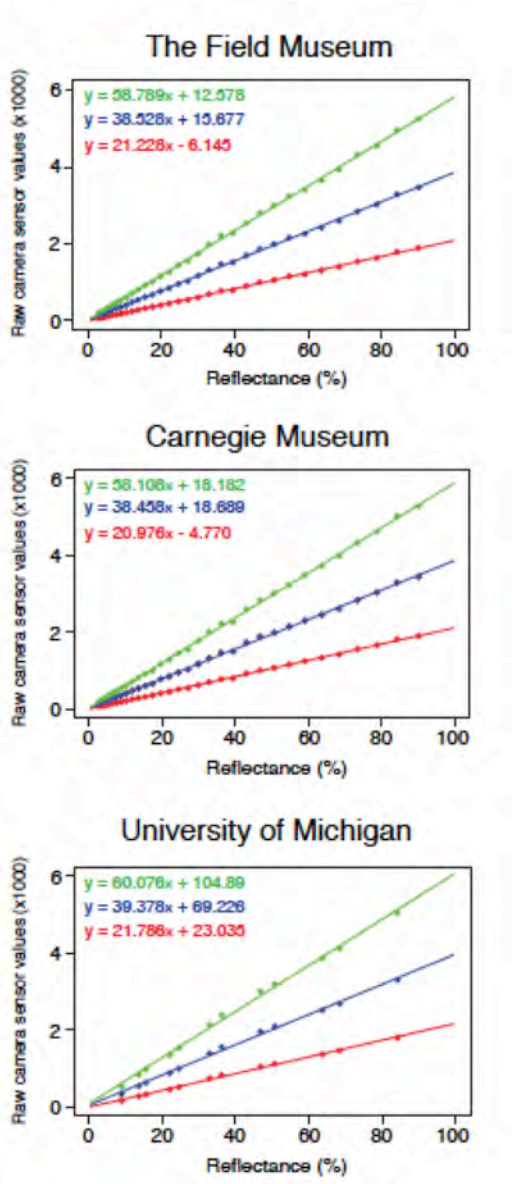

Figure 13 | Raw R, G, and B channel-specific regressions based on the Munsell Neutral Value Scale reflectance standards for each shooting location. The regression equations for each channel were used to calculate channel-specific reflectance from raw CMOS sensor data recovered in RawDigger for each specimen.

preparation of the specimens. With these parameters in place, a shutter speed of 1/25 s was selected to maximize the dynamic range.

While the use of a light box ensured relatively even and continuous illumination compared with open studio lighting arrangements, perfectly consistent illumination is difficult to

achieve in practice. Some unevenness was discovered in blank reference images, which was determined to have resulted from lens variables (optical vignetting and lens flare) and may have also been influenced by the arrangement of the bulbs in the light box. To account for these factors, the target area for each specimen was confined to a 3- × 3-inch square, which limited variance in illumination to <1%.

Under the constant lighting conditions that a light box provides, reflectance standards theoretically only need to be photographed once over the course of shooting to generate calibration regressions. In practice, however, some minor variations in overall illumination were discovered between the three locations, which may have been due to light leakage or slight variations in the voltage supply to the bulbs at each location. This variation, however, was easily accounted for by imaging the Munsell Neutral Value Scale reflectance standards at each location and calculating reflectance values for specimens with location-specific reflectance regressions. Since reflectance is expressed as a percentage, and these percentage values are relative to the standards, no additional adjustments were needed to normalize the color channels or calibrate the values across shooting locations. We photographed each card of the Munsell Neutral Value separately at The Field Museum and Carnegie Museum of Natural History, positioning each card at the center of the field of view in the same area where we measured reflectance from bird feathers. To determine reflectance regressions from these locations, we used all 31 reflectance standards (ranging from 3.1 to 90% reflectance). At the University of Michigan Museum of Zoology, we photographed the Munsell Neutral Value Scale fanned out in single photograph. For this sample, we only included 12 reflectance steps (ranging from 9 to 84.2% reflectance) that fell within the target area (Figure 13).

1.6.2 Quantifying reflectance of specimens

To determine the reflectance value for each specimen from a digital image, we used regression equations calculated from reflectance standards to convert raw sensor data to known reflectance values. We calculated R, G, and B channel-specific regressions from Munsell Neutral Value Scale reflectance standards in RawDigger (Version 1.2.11) for each of our three shooting locations: The Field Museum, Chicago; University of Michigan Museum of Zoology, Ann Arbor; and Carnegie Museum of Natural History, Pittsburgh (Figure 13). Since our camera's CMOS sensor incorporates an additional G channel (G2), we averaged both G-channel values to produce a single G-channel regression. The equations for each regression line can be found in Figure 13. We uploaded the digital photograph of each specimen into RawDigger and sampled the uniform white patch on the ventral side of each specimen. We recovered median raw R, G/G2, and B channel sensor values from a sampling area that ranged from 25 to 900 mm^2. Since feathers are a textured, heterogeneous surface, median values were used to minimize any effect of outliers. For each specimen, the sample area was determined by selecting a large continuous area without conspicuous portions of exposed skin, staining due to residual fat deposits, or other preparation and conservation issues (see Dataset S1 for sample areas). We used the collection-specific regression equations to calculate reflectance values separately for R, G/G2, and B channels for each specimen. We then averaged the three channel-specific reflectance values to obtain a composite reflectance value for each specimen.

1.6.3 Determining the smoothing function for the GAM

Smoothing parameters for GAMs can be determined in mgcv by using functions such as GCV that minimize residual deviance (goodness of fit) and degrees of freedom (Wood, 2011). With

our final dataset, the GAM estimated a smoothing function of $k = 10$ (this model is plotted in Figure 14), which recovered a smoother curve than $k = 20$ (Figure 9). Oversmoothing, however, can obscure signals in the data (Peng et al., 2006; Chuang et al., 2011), which appears to be happening with $k = 10$ based on our knowledge of likely inflection points (such as the 1929 US stock market crash) that are present in the consumption data and the Greenland ice-core record. For reference, in Figure 14, we include a variety of smoothing functions from $k = 10$ to $k = 100$. Based on the comparison of possible k values, $k = 10$ appears to apply an overly powerful smoothing operation in the GAM, forcing the first decline of black carbon to begin in the early 1920s rather than the end of the decade where we would expect it to appear based on consumption trends; $k = 13$ through $k = 35$ recovers trends that are effectively identical, which appears to recover important signals in the data that over smoothing misses; $k = 36$ and greater generate toothy trends that overrepresent random variations within the sample set. Based on the variation in the shape of different GAMs, we selected a smoothing function of $k = 20$ to produce a relatively smooth trend line that still maintained a distinctive shape that allowed for comparison against consumption data.

1.6.4 How sampling months were determined

Beginning in late summer, each species used in the study initiates an annual molt to replace worn and soiled body feathers with fresh plumage. This molting period can last through the fall months (Pyle, 1997). Natural variation in the timing of the molt produces a mix of birds with fresh and soiled plumage among specimens sampled from these months. This annual molt signal was apparent in our sample, with samples from fall months producing shifts in mean reflectance caused by the introduction of freshly molted birds, along with uncharacteristically broad ranges

in reflectance values compared with other months (Figures 5 and 6). Freshly molted individuals do not provide evidence for atmospheric conditions in a given year, warranting their removal from the final dataset. Since freshly molted birds begin to accumulate particulate matter immediately after the molting cycle is complete, rather than selectively evaluating which individuals had recently molted, all of the specimens sampled during these months were removed. We determined the months to exclude for each species based on abrupt shifts in mean reflectance between months, which are indicative of annual molting patterns. For example, in Horned Larks, reflectance values shift abruptly between July and August and then increase again between November and December, indicating that the sample of birds in the months of August–November includes a substantial number freshly molted individuals (Figure 5). Following this method, the months of August–November were excluded for Horned Larks and Red-headed Woodpeckers, and the months of September–November were excluded for Field Sparrows, Grasshopper Sparrows, and Eastern Towhees (Figure 5). We could be confident in these shifts given their seasonal timing, since overall black carbon emissions seasonally trend in the opposite direction for a given year in the Northern Hemisphere, as fuel consumption increases to meet heating needs when average temperatures drop (Ives, 1936; McConnell et al., 2007). We limited this inquiry to the years 1880–1950 because after midcentury, birds are substantially cleaner in all months, compromising our ability to detect monthly breakpoints.

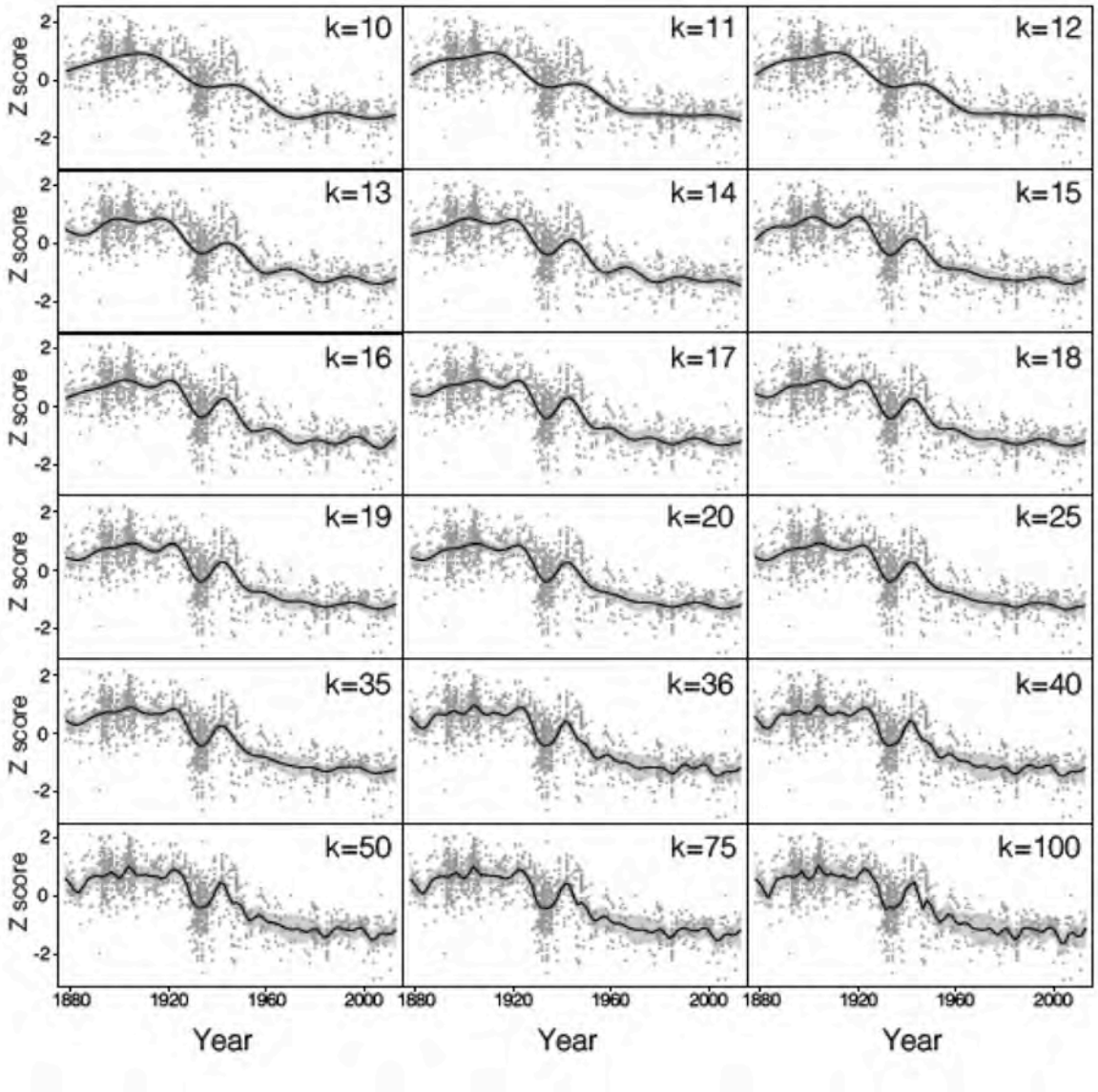

Figure 14 | GAMs with various smoothing functions applied to the normalized 1,097-specimen dataset. k = 10–12 applies an overly powerful smoothing operation in the GAM; k = 13–35 recovers trends that are effectively identical, which appear to recover important signals in the data absent from the k = 10–12 models; and k = 36 (and greater) generates a toothy trend that overrepresents random variations within the sample set.

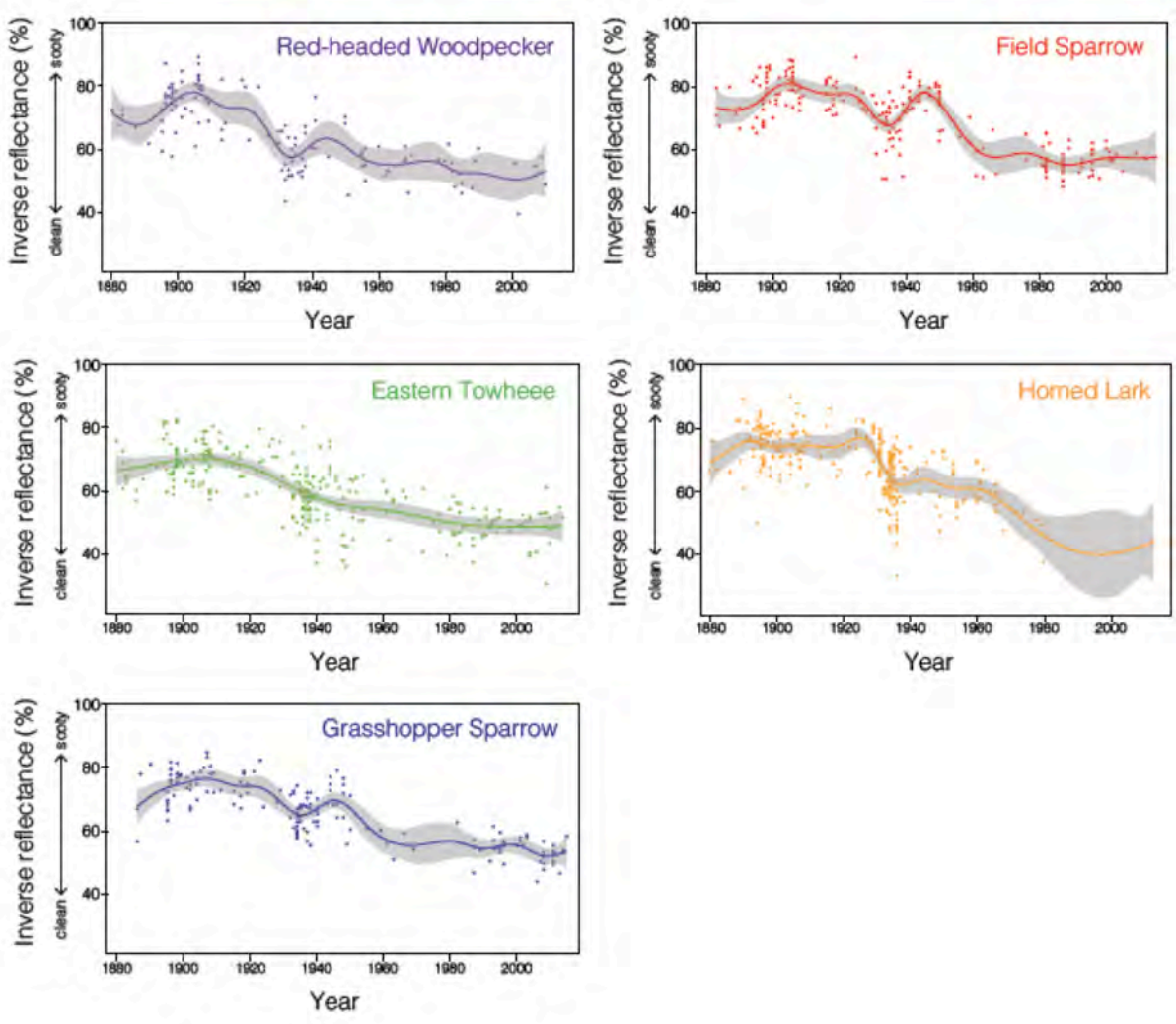

Figure 15 | Species-specific trends in black carbon deposition. Each point represents an individual specimen. The colored lines are GAMs ($k = 20$) with 95% confidence limits (shaded area) for each species [fall-month birds are excluded (Figure 5)]. Inverse reflectance is reported, rather than reflectance, to visualize drops in atmospheric black carbon.

2 Human disturbance provides foraging opportunities for birds in primary subalpine forest[2]

2.1 Abstract

Interspecific foraging associations are well-documented phenomena, characterized by one or more species exploiting the behavior of another species to decrease predation or increase foraging success. In rare cases, birds directly exploit human behavior, but examples of these interactions are limited to species that naturally occur in edge, open, or disturbed habitats. With observations and experiments we provide evidence of insectivorous birds exploiting human disturbance in primary subalpine forest in the mountains of southern China, displaying behavioral flexibility to gain novel foraging opportunities. We cut and cleared small swaths of dense bamboo growth for an unrelated study. Multiple insectivorous species were recruited to the cleared areas, foraging extensively in the disturbed earth, often within 1 m of us. These species included *Tarsiger chrysaeus, Tarsiger indicus*, *Cettia brunnifrons* or *Cettia major*, and *Heteroxenicus stellatus*. This behavior is likely a modification of pre-existing interspecific foraging associations with pheasants and large mammals in the region. These larger animals disturb the earth and lower vegetation layers upon passage and while foraging, exposing previously inaccessible invertebrate prey items on which small insectivorous birds can feed. Our findings highlight a behavioral capacity in birds to use human disturbance in an ecosystem with limited human presence.

[2] A version of this chapter has been published as: DuBay SG, Hart Reeve A, Wu YJ. 2017. Human disturbance provides foraging opportunities for birds in primary subalpine forest. Journal of Ornithology 158: 833-839.

2.2 Introduction

Almost 70 years ago, Lack (1948) noted an exceptional foraging behavior in the European Robin (*Erithacus rubecula*): robins will follow Common Pheasants (*Phasianus colchicus*), keeping close and feeding on invertebrates in the earth disturbed by the pheasants. Buffon (1771–1783) noted a similar behavior centuries before, observing that European Robins were attracted to humans, often following travelers through the forest. The robins described by Lack (1948) and Buffon (1771–1783) were tame and unafraid of game birds and large mammals, including humans. It has been suggested that these birds are, in fact, attracted to larger-bodied birds and mammals that break up the ground and disturb the leaf litter upon passage (Lack 1948), exposing invertebrate prey items hidden beneath the surface. Given their size and anatomy, robins are unable to turn over large debris and leaf litter (Lack, 1948), thus this foraging strategy provides access to previously inaccessible prey. The foraging behaviors observed by Lack (1948) and Buffon (1771–1783) have fascinated observers since their description, but explicit tests to better understand these behaviors are rarely performed. In this study we coupled behavioral observations of insectivorous birds exploiting human disturbance with experimental manipulation to extend our understanding of this adaptive foraging strategy first described by Buffon (1771–1783) in the 18th century.

Interspecific foraging associations have been documented in a diversity of avian taxa. For example, birds often follow and forage in association with Army Ant swarms and non-human primates, or other large mammals (Rand, 1953; Dean and MacDonald, 1981; Zhang and Wang, 2000; Kuniy et al., 2003; Beiseigel, 2007; King and Cowlishaw, 2008; Willis and Oniki, 1978). The follower benefits from increased foraging opportunities as the other animals flush and expose prey items. Given the repeated evolution of interspecific foraging associations (see

Heymann and Hsia, 2015), and the limited geographic scope of foraging association studies (confined mostly to the Neotropics) (King and Cowlishaw, 2008), these interactions are likely more taxonomically and geographically widespread than currently recognized.

Evidence of human/bird foraging associations are largely limited to species that naturally occur in edge, open, and/or disturbed habitats, such as animals that use urban and agricultural environments (Rand, 1953; Dugatkin, 2013). It is no surprise, or coincidence, that birds that have evolved to occupy disturbed and open habitats have flourished in human-modified environments through association with human activities, such as New World blackbirds (family Icteridae) foraging behind tractors (Beasley and Carothers, 1974) and European Robins (*Erithacus rubecula*) foraging in association with gardeners and gravediggers (Lack, 1948). Less common, however, is evidence that birds can exploit human activities in old-growth forest.

Here, we present evidence that insectivorous birds can exploit human disturbance in primary subalpine forest in the mountains of southern China. In our research we routinely cut and clear small swaths of dense bamboo in the forest to erect mist nets to catch birds. We observed insectivorous species that were attracted to the recently cleared mist-net lanes, foraging in the disturbed earth and bamboo cuttings, often within a few meters of us. Following our initial observations, we experimentally tested the hypothesis that insectivorous birds are attracted to, and utilize, habitat disturbance.

2.3 Methods

2.3.1 Study site

We conducted this study in Gongga Shan National Nature Reserve, Sichuan, China (latitude 29°34′ 21.6084″ N, longitude 101°59′ 10.6188″E). Mount Gongga, which is located within the reserve, is the easternmost peak in Asia above 7000-m elevation and is located in the central Hengduan Mountains of southern China. The Hengduan Mountains form the southeastern edge of the Tibetan Plateau and western margin of the Sichuan Basin, spanning steep elevational and environmental gradients. In these mountains, bird habitats extend from subtropical lowlands to alpine meadows. We conducted our study at subalpine elevations on the eastern slope of Mount Gongga, where the forest is dominated by a tall conifer (*Abies fabri*) canopy, rhododendron (*Rhododendron vernicosum*) mid-story, and dense bamboo (*Sinarundinaria* spp.) understory.

The mountains of southern China harbor unparalleled temperate diversity (Fjeldså et al., 2012). To date, 939 bird species are recognized from this region (Wu et al., 2016). These mountains are the center of diversity for pheasants (family Phasianidae) (del Hoyo et al., 2005), and seven species of pheasant occur in Gongga Shan National Nature Reserve (Zhou et al., 2014). These large birds disturb the ground layer as they forage. In addition, large mammals that also disturb the ground layer are diverse in the mountains of southern China. These mammals include badgers (family Mustelidae), Asiatic bears and pandas (family Ursidae), and numerous species of ungulates, like Takin and other bovids (*Budorcas taxicolor*; family Bovidae), Wild Boar (*Sus scrofa*), and deer (family Cervidae).

The Hengduan Mountains experience dramatic seasonal fluctuations in climate because of their temperate latitude. At middle and upper elevations, snow covers the ground for much of

the winter and daily winter temperatures routinely drop below freezing (Figure 16). At 3000-m elevation, minimum daily temperatures can vary by as much as 25 °C between winter and summer (Figure 16). In response to this seasonality, the majority of montane birds seasonally migrate across elevations to escape harsh winter conditions at upper elevations (del Hoyo et al., 2005). These altitudinal migrants return to breeding elevations in spring as snows thaw and temperatures warm. In Gongga Shan National Nature Reserve, we have recorded altitudinal migrants arriving to breeding elevations above 3000 m as early as March, but the majority of spring migration to these upper elevations occurs in April and early May (unpublished data of the authors). By June, most species that breed above 3000 m have already started breeding. During the transition from winter to summer (March–June), migrants can experience unexpected snowstorms and high degrees of climatic variability (Figure 16).

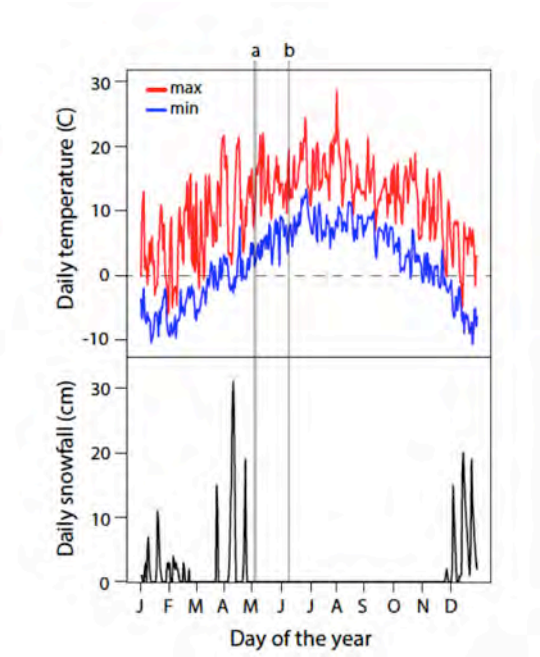

Figure 16 | Daily temperature and snowfall at 3000-m elevation on the eastern slope of Mount Gongga, 2015. *X-axis marks* indicate the first day of each month. *a* Indicates the date of the initial observations, *b*indicates the date of the controlled experiments.

2.3.2 Initial observations and controlled experiments

On 4 and 5 May 2015, while mist-netting birds for an unrelated study at 3250-m elevation on the eastern slope of Mount Gongga, we observed that insectivorous birds were attracted to our disturbance, foraging in the disturbed earth along cleared mist-net lanes. To erect mist nets, we cut and cleared dense bamboo to create lanes 1.5 m wide and 12 m long in primary subalpine forest.

On 9–12th June 2015, we returned to Gongga to collect video footage of birds foraging in disturbed net lanes (Figure 17) and experimentally test the hypothesis that birds are recruited to habitat disturbance. In the experiment we simulated standardized net-lane clearing because this disturbance is uniquely human and because we initially observed birds foraging in our net lanes, allowing us to make comparisons between our initial observations and the experiment. The experiment was as follows:

1. Two video cameras were placed facing each other in the bamboo understory, 10–12 m apart and 1 m off the ground.

2. S. G. D. and Y. W. were each positioned at the base of a camera, where they remained silent for 10 min before beginning the control treatment. This 10-min buffer period ensured that the birds observed were not recruited to the area by our noise and movements as we entered the area and set up cameras.

3. We then began the control treatment, observing and recording the presence of birds for 20 min in the undisturbed bamboo between the cameras. Birds were recorded as

"present" if they crossed the transect between the cameras within 3 m of the ground. We made visual observations, which were corroborated with video footage to ensure that all birds present were recorded.

4. Following the control treatment, S. G. D. and Y. W. cut a 1.5-m-wide lane between the cameras for 10 min, cutting the bamboo at its base and trampling the vegetation and ground layer. This method disturbed the topsoil and leaf litter layer.

5. After the disturbance, we resumed our positions at the cameras, observing and recording the presence of birds in the disturbed area between the cameras for the next 20 min. Birds were recorded for the disturbance treatment following similar protocols as the control treatment.

We conducted eight paired replicates of the above experiment.

Figure 17 | Still photograph from a video showing a female-plumaged White-browed Bush-robin (*Tarsiger indicus*) foraging in a recently cleared net lane.

The nature of the disturbance treatment prevented us from conducting the disturbance treatments before the controls at the same site. However, we preformed two additional replicates for each treatment at independent sites. That is, for two additional control replicates we performed steps 1–3 without subsequently performing steps 4 and 5, and for two additional disturbance replicates we performed steps 1, 4, and 5 to the exclusion of steps 2 and 3. If the ordering of the paired treatments increases bird recruitment in the disturbance treatment, negatively biasing the experiment, then we would expect fewer relative recruitment events for the disturbance treatment during independent replicates compared to paired replicates. All experimental replicates (totaling ten replicates for each treatment) were conducted between 3000- and 3300-m elevation, 50–100 m apart, and in areas with similar habitat structure. Based on the results outlined below, we are confident that a distance of >50 m among replicates was sufficient to avoid biases associated with pseudoreplication (i.e., detecting the same individual in multiple replicates).

2.4 Results

2.4.1 Initial observations

In May 2015, we observed four species (*Cettia brunnifrons* or *Cettia major*, *Heteroxenicus stellatus*, *Tarsiger chrysaeus*, *Tarsiger indicus*) foraging on the disturbed earth beneath cleared mist net lanes. On one occasion, two male-plumaged *T. chrysaeus* foraged within 5 m of each other for approximately 5 min, before one chased the other away. On another occasion, a female-plumaged *T. indicus* appeared and began foraging in the cleared areas as we were cutting bamboo. This bird continued to forage alongside us despite the chaos and noise of our clearing

efforts, often coming within 1 m of us. After this net lane was cleared, we observed three distinct individuals (two female-plumaged *Tarsiger indicus* and one male-plumaged *Tarsiger chrysaeus*) foraging within the 12-m net lane. On multiple occasions we observed more than one individual foraging at a given time in the cleared net lane in an otherwise dense bamboo understory. At no point did any of the birds appear to be disturbed by our presence.

2.4.2 Controlled experiment

We observed increases in bird recruitment in the disturbance treatment (Figure 18). We observed birds in six of the disturbance replicates (eight birds total) and in one of the control replicates (one bird total). The increase in birds observed in the disturbance treatment from the control treatment was statistically significant in a one-tailed Fisher's exact test ($p = 0.029$), testing the alternative hypothesis that recruitment increases with disturbance. In contrast to our initial observations, at no point during the experiment did we observe more than one bird at a time foraging in the cleared lane, nor did we observe birds during the 10-min cutting period. Interestingly, only three of the eight birds recorded in the disturbance treatment were observed foraging in the lane. The remaining five individuals landed in the cleared area or flew across the transect without foraging during the 20-min observation period. The birds observed during the disturbance treatment were: one male *Tarsiger chrysaeus*, two female-plumaged *Tarsiger chrysaeus*, two female-plumaged *Tarsiger indicus*, one *Ficedula strophiata*, and two unidentified passerines. We did not observe birds in the independent control treatments, but we observed birds in both independent disturbance replicates, suggesting that the ordering of the paired treatments did not negatively bias the experiment.

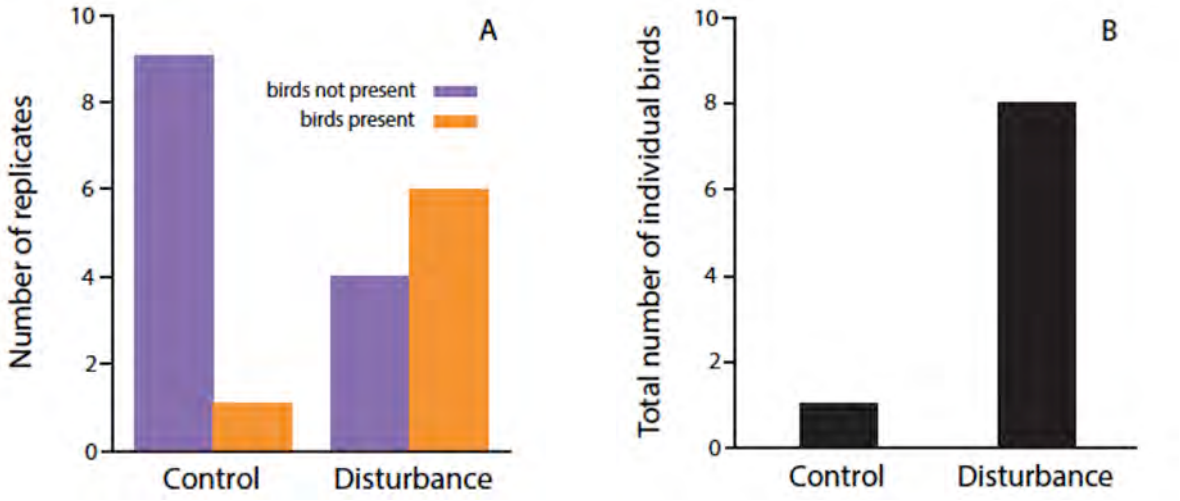

Figure 18 | Experimental results testing if birds are attracted to disturbance. (A) The number of experimental replicates in which birds were observed. (B) The total number of individual birds observed in each experimental treatment (n = 10 replicates/treatment).

We identified two possible instances of pseudoreplication: first, a female-plumaged *T. chrysaeus* was recorded in two replicates that were 75 m apart. We cannot be positive that these birds were not the same individual, but our results would remain unchanged even if they were because both of these experiments had other birds present. If we remove this individual from the dataset (or attribute it to only one of the replicates) the result remains the same because the number of birds recruited in each replicate is not taken into account in the Fisher's exact test, only whether birds were present or absent in a replicate. Second, a female-plumaged *T. indicus* was recorded during two replicates that were 135 m apart. These experiments were done in an area that we previously surveyed for *T. indicus* the month before. We know that these two replicates were in distinct *T. indicus* territories with an additional territory situated between. These two lines of evidence suggest that pseudoreplication did not bias our results or interpretations.

2.5 Discussion

The experimental results, paired with the initial observations, support the hypothesis that insectivorous birds are attracted to, and use, habitat disturbance. The foraging behavior that we observed is likely an extension of a pre-existing foraging behavior. We (as large mammals) were viewed not as a threat, but as agents of disturbance, providing insectivorous birds access to prey items. Small songbirds (order Passeriformes), such as those observed here, have co-existed with a diversity of large-bodied game birds and mammals, capable of disturbing the ground layer, for millions of years in the subtropical and temperate mountains of southern China, providing ample time and opportunity for interspecific foraging associations to evolve and persist. Given that we specifically cut and disturbed the bamboo understory, it is tempting to suggest that the recruitment behavior we observed possibly evolved in association with the Giant Panda (*Ailuropoda melanoleuca*), which is famously reliant on this food source and disturbs the bamboo understory as it forages (Sheldon, 1937; Johnson et al., 1988). It is possible that the observed behavior evolved in association with the Giant Panda during the past millennia when the species was widespread in the mountains and lowlands of central and southern China (Loucks et al., 2001).

The recruitment of birds to our disturbance may be an extension of foraging associations with particular large-bodied species, general associations with the collective suite of larger animals that occur in the region, or a more general attraction to any active forest disturbance, including landslides, tree falls, or edges of fires. It is most plausible that this foraging strategy is not species specific, but has rather evolved in association with general disturbance and high diversity of pheasants and large mammals in the region. If this foraging behavior evolved in association with a suite of large-bodied animals and/or general disturbance, this might explain

the ease with which the birds exploited our disturbance in the forest. Thus, the foraging behavior that we observed may naturally extend from a pre-existing adaptive foraging strategy rather than represent a stark modification of behavior.

Of the birds that we observed foraging in the cleared net lanes, all the species are insectivores and predominately forage for insects and larvae on the ground and in low vegetation (del Hoyo et al., 2005). For example, species like Gould's Shortwing (*Heteroxenicus stellatus*) are largely terrestrial. The species that we observed foraging the most extensively in the disturbed areas (the shortwing and the two *Tarsiger* bush-robin species) belong to the same family (Muscicapidae) as the European Robin. It is unknown whether these species independently evolved this foraging strategy or if this behavior is ancestral in muscicapids. Variation in the strength of human associations among European Robin populations suggests that this adaptive foraging strategy may be easily lost and gained (Lack, 1948). A detailed study of the geographic and taxonomic extent of this adaptive foraging strategy would increase our understanding of the evolutionary and ecological context of this behavior.

At middle and upper elevations in the mountains of southern China the ground is covered in snow for large portions of the year. In response, the majority of insectivorous birds migrate to lower elevations for the winter to escape cold temperatures and a lack of resources at upper elevations (del Hoyo et al., 2005). When these birds return to breeding elevations in March and April the ground can still be covered in snow, with the possibility of additional snow accumulation (Figure 16). It is possible that disturbance from pheasants and large mammals may contribute to the survival of insectivorous birds at upper elevations during periods of snow cover when invertebrate prey items are otherwise trapped beneath the snow layer. Lack (1948) noted that European Robins were more likely to be associated with humans in winter. He hypothesized

that the strength of foraging associations fluctuates among seasons in response to resource availability. Seasonal variation in resource availability in the mountains of southern China may explain the difference we observed in the strength of response to disturbance between our initial observations and the experiment. The recruitment response to disturbance appeared weaker in mid-July (during the experiment) than it did 1.5 months earlier when we made our initial observations. For example, in early May we observed multiple individuals foraging in net lanes at the same time, and birds foraged while bamboo was being cut, neither of which occurred during the experiments in mid-June. Additionally, only three of the eight individual birds that we observed in the disturbance treatment foraged extensively in the cleared areas. Our observations are consistent with Lack's (1948) seasonal observations in the European Robin, suggesting that seasonal resource fluctuations may regulate the strength of adaptive foraging associations through the year.

Additional explanations for the apparent difference in the response strength between our initial observations and the experiment include decreased density and mobility of birds in mid-June around 3000-m elevation. In early May (when our initial observations were made) the community is more dynamic and birds are more likely to cross paths with larger animals; birds are establishing breeding territories, females are finding mates, and a portion of the community at 3000-m elevation is migrating upslope. These dynamics are exemplified in our initial observation of two adult male *T. chrysaeus* foraging together in a cleared net lane, an unlikely scenario once breeding territories are set. By mid-June, breeding territories are established, spring migration is complete, and birds are nesting, decreasing the density of birds foraging in the environment at a given time. Our results and observations are consistent with the hypothesis that insectivorous birds may be more reliant on interspecific foraging strategies during migration

and in early spring when resources are scarcer at upper elevations, but this seasonal dynamic warrants further investigation.

In this study we provide observational and experimental evidence that multiple insectivorous bird species are attracted to human disturbance in primary subalpine forest, where they gain foraging opportunities. The observed birds conspicuously and extensively foraged in the disturbed area, often responding within minutes of our initial disturbance. The ability to exploit novel foraging opportunities, specifically in association with humans, highlights behavioral flexibility within a bird community that experiences relatively little human disturbance. The relative speed and ability of birds to exploit novel foraging opportunities by coopting pre-existing interspecific associations could potentially facilitate species survival in the wake of environmental change and shifts in community composition, especially as humans increasingly modify the environment (Walther et al., 2002). Our observation that human disturbance provided foraging opportunities for birds in primary forest offers a promising direction for future research at the interface of adaptive behavior, environmental change, conservation biology, and biodiversity science.

3 Age- and sex-structured variations in flight muscle and behavior underlies classic life history strategies in birds

3.1 Introduction

The avian flight apparatus has enabled the evolution of diverse life history strategies in birds. Flight muscles not only facilitate a remarkable range of functions associated with powered flight, but also play an essential role in thermoregulation, as birds generate endogenous heat primarily through muscle shivering (Marsh and Dawson, 1989; Block, 1994). As the dominant muscle in the flight apparatus, the *pectoralis major* is responsible for the downward wing stroke, lift, and thrust, and accounts for the majority of muscle mass in birds. The *pectoralis* is inextricably linked to diverse functional tasks, like seasonal movements, foraging/prey capture, predator avoidance, social interactions, and thermogenesis. Variation in *pectoralis* phenotypes, like muscle size and fiber composition, has been linked to diverse ecological strategies among species (Rosser and George, 1985; Choi et al., 1993; Scott et al., 2009; Dakin et al., 2018). As a classic multifunctional organ, the *pectoralis* must balance the diverse demands of powered flight and thermoregulation, and is thus deeply integrated into avian life history and ecology (Rosser and George, 1985; Scott et al., 2009; Segre et al., 2015; Wright et al., 2016; Dakin et al., 2018). Here, we expand beyond species-level comparisons, combining physiological and behavioral data across age and sex classes to study how demographic-dependent selection on function shapes flight muscle phenotypes. Specifically, we test how competing functional demands within species are resolved in the *pectoralis*.

Comparative methods have uncovered striking relationships between flight muscle form and function that reflect taxon-specific selection on flight muscle phenotypes (Rosser and

George, 1985; Dakin et al., 2018). For example, bar-headed goose (*Anser indicus*) populations that migrate over the Himalayas have accumulated changes in their *pectoralis* muscle fibers and vasculature that appear to enhance endurance flight and aerobic capacity under extreme hypoxia (Scott et al., 2009; Hawkes et al., 2011). In hummingbirds, increases in flight muscle capacity and size translate to functional increases in performance and acceleration, facilitating dynamic flight ecologies (Segre et al., 2015; Dakin et al., 2018). In comparison, island birds have evolved smaller *pectoralis* muscles than continental relatives with reduced predation pressure, underscoring the functional role of flight muscles in predator avoidance (Wright et al., 2016). Comparisons between species with contrasting ecologies have been foundational in establishing broad connections between flight muscle phenotypes and ecology (Rosser and George, 1985; Dakin et al., 2018), but these studies only address form and function along a single axis of variation – among species. Given that selection on flight muscle phenotypes and function also varies along demographic axes like age and sex, there exist fundamental axes of life history variation within species that remain to be explored (Ricklefs and Wikelski, 2002).

A powerful, yet underutilized approach to understand how diverse selective environments shape flight muscle phenotypes is include demographic axes, like age and sex, in studies of phenotypic variation. This approach controls for confounding factors associated with general differences among species, and when used in concert with species comparisons, can provide a more detailed picture of how flight muscle phenotypes are optimized by age, sex, and species to meet demographic-specific demands on functionality. Without considering selection on flight muscle function along demographic axes of age and sex, we are limited in our ability to understand how flight muscle phenotypes are integrated into an individual's life history and ecology. Males, for example, might experience stronger selective pressures to invest resources in

flight muscle traits that enhance territoriality or, in migratory birds, early arrival to breeding grounds, which is a general tendency in migratory birds (Morbey and Ydenberg, 2001). Females, however, should benefit from directly investing resources in reproductive traits, like egg formation, which could potentially draw from protein stores in the *pectoralis* (Raveling, 1979). By accounting for variation along demographic axes, we can better understand how diverse selective environments shape flight muscle phenotypes to balance competing functional demands.

The functional demands on the *pectoralis* muscle are met in part by modifying underlying muscle components, such as muscle fiber identity and tissue vasculature, both of which are shaped by the distinct ecological pressures experienced by species and individuals (Leon-Velarde et al., 1993; Scott et al., 2009). The *pectoralis* of most volant birds, for example, is comprised of two fast-twitch skeletal muscle fiber types: fast oxidative-glycolytic fibers and fast glycolytic fibers (hereafter referred to as fast oxidative and glycolytic fibers, respectively; Rosser et al., 1994; Meyers and Stakebake, 2005; Welch and Altshuler, 2009). The *pectoralis* of small birds, however, can be comprised exclusively of fast oxidative fibers, while non-volant and soaring species can have a third skeletal muscle fiber type, slow oxidative fibers, underscoring the demand on fast-type fibers for powered flight (Rosser et al., 1994; Meyers and Stakebake, 2005; Welch and Altshuler, 2009). Fiber types differ markedly in their physiological and functional attributes (Peter et al., 1972), and their relative abundance within the *pectoralis* fundamentally impacts an individual's capacity for powered flight and thermogenesis. Fast oxidative fibers are aerobic and can sustain muscle contractions over longer periods of time. By increasing these fatigue-resistant fast oxidative fibers, the *pectoralis* can enhance functions associated with shivering thermogenesis and sustained powered flight. In contrast, glycolytic

fibers are anaerobic, predominately metabolize carbohydrates, and can generate high mechanical power in short bursts. Glycolytic fibers lack the necessary endurance for sustained shivering thermogenesis, but enhance functions associated with short-burst flight capacity, like predator avoidance, takeoff, and other rapid flight maneuvers (Dial et al., 1987; Askew et al., 2001). Past work has suggested that *pectoralis* phenotypes in smaller birds are constrained by thermogenic demands and less constrained by the need for burst fiber types for takeoff, which explains the propensity for the *pectoralis* of small birds to contain only fast oxidative fibers (Welch and Altshuler, 2009). Fast oxidative fibers have thus been hypothesized to meet all the functional demands of thermogenesis and powered flight in small birds (Welch and Altshuler, 2009; Dakin et al., 2018). However, few species have been surveyed for these phenotypes.

We studied three species of *Tarsiger* bush-robins (*Tarsiger indicus*, *T. chrysaeus*, and *T. rufilatus*) that demonstrate a number of ecological differences within and among species that is tied to life along an elevational gradient. Bush-robins also demonstrate common life history strategies that provide the necessary context to interpret patterns of flight muscle variation. Bush-robins are small passerines (<17 g) that are sexually dichromatic and socially monogamous (Morimoto et al., 2005; Satio et al., 2006). Within species, they show delayed plumage maturation in which first-year males (subadults) retain female-like plumage despite being fully capable of breeding (de Hoyo et al., 2005). Delayed plumage maturation is a widespread, adaptive life-history strategy in which subadult males signal their inferior competitive status to avoid the costs associated with direct competition with socially dominant adult males (Rowher et al., 1980; Lyon and Montgomerie, 1986; Morimoto et al., 2006; reviewed in Hawkins et al., 2012). Moreover, between species, bush-robins show striking differences in migratory behavior, seasonal distributions, and the environments they experience. Our focal species breed in

sympatry at upper elevations above 3000 m in the Himalayas and Hengduan mountains of temperate Asia, but migrate downslope to different elevations for the winter (Figure 19). These species thus experience dramatic differences in environmental pressures over the course of a year. Furthermore, within each species, we discovered age- and sex-structured differences in the timing of migration, reflecting another widespread, adaptive life history strategy in which differences among age and sex classes in the timing of life history events have evolved to optimize lifetime reproduction (Mayr, 1939; Morbey and Ydenberg, 2001). The distinct selective environments within and among species impose different functional demands on flight muscle phenotypes that must be balanced within an individual's lifetime and between species.

We leveraged within- and among-species comparisons in *Tarsiger* bush-robins, integrating flight muscle physiology, behavior, and environmental data within the context of age- and sex-specific selection to ask how diverse functional demands on flight muscle are reconciled in this multifunctional organ. Using an unprecedented dataset, we specifically ask how variation in flight muscle phenotypes is associated with age- and sex-structured migration patterns and delayed plumage maturation, two classic life-history strategies that have driven decades of research (Mayr, 1939; Morbey and Ydenberg, 2001; Hawkins et al., 2011). Analyzing flight muscles in the context of life history provides the necessary framework to understand selective forces driving and maintaining phenotypic variation, while also revealing the underlying functional basis of life history variation. In this study we leverage an intersectional framework in which we assert that flight muscle phenotypes are a manifestation of the combined and balanced selective pressures that individuals experience as a result of multi-class identity (i.e. species, age, and sex).

3.2 Results

We caught bush-robins (n = 329) over three consecutive spring seasons as they migrated upslope to breeding elevations in Hailuogou Valley, Sichuan, China, which is situated on the eastern edge of the Tibetan Plateau (Figure 19). In this region, the three species form a clade, breeding in sympatry above 3000 m elevation (Figure 19). The three species show no evidence of interbreeding throughout their range. We caught all individuals within one kilometer of a local climate station in subalpine forest at 3000 m as they moved between elevations, either executing annual spring migration or short-term elevational movements in response to acute early-spring cold weather events (Figure 19). Breeding territories were established above the sampling elevation. Maximum and minimum temperatures for the date of capture were recorded for each individual. For a subset of individuals of each species (n = 129), we surveyed *pectoralis* muscle

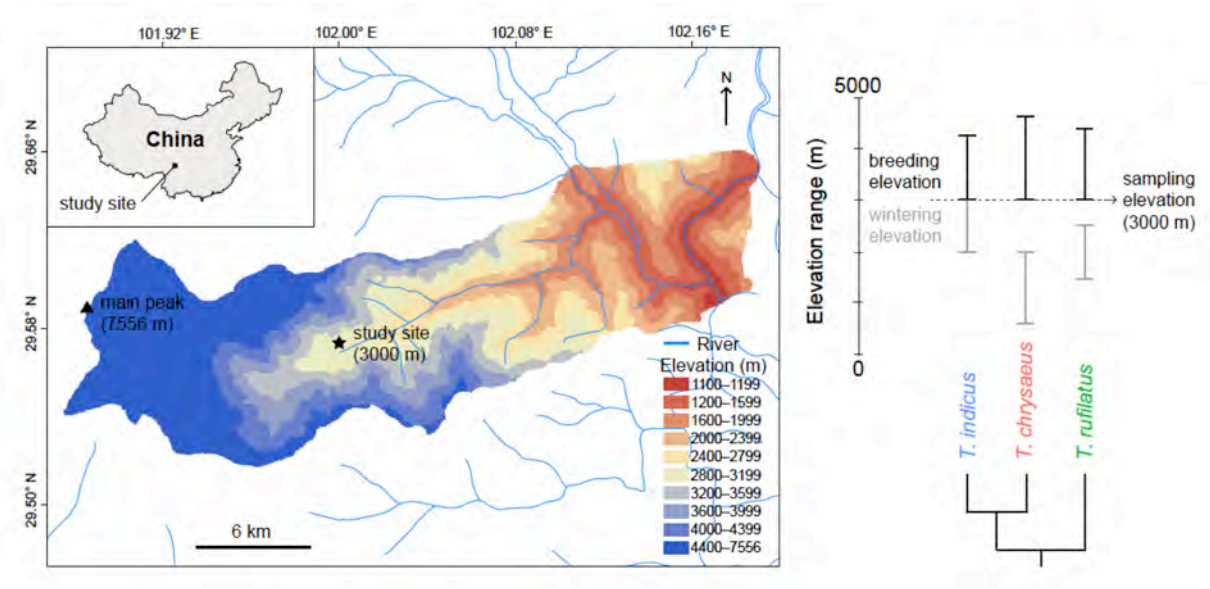

Figure 19 | Map of Hailuogou Valley, Sichuan, China (study site), phylogenetic relationships, and seasonal elevation distributions of study taxa. Phylogenetic relationships are derived from Sangster et al. (2010), and elevational distributions are from Handbook of Birds of the World (de Hoyo et al. 2005).

phenotypes, quantifying muscle fiber type and muscle capillarity in three demographic classes: adult males, subadult males, and females. We treated females as a single group because female age classes are indistinguishable by plumage. We analyzed all flight muscle traits as raw values and relative values (as a proportion of body mass). The interpretations of results remain the same with and without controlling for differences in body mass. We tested for statistical differences among species, between sexes, and between male age classes with analysis of variance and covariance tests, which are reported in Tables 1-3.

3.2.1 Variation in flight muscle phenotype

Pectoralis mass and body mass varied foremost by species and sex class – *pectoralis* muscles and body mass were larger in *T. indicus* and *T. rufilatus* than in *T. chrysaeus*, and larger in males than females (Figure 20; Tables 1 and 2). Despite the small body size of bush-robins (<17 g), we discovered the presence of both fast types of muscle fibers (fast oxidative and fast glycolytic) in the *pectoralis* of all three species (Figure 21). When we considered the relative contributions of each fiber type to total *pectoralis* mass, the total masses of both fast oxidative fibers (Figure 22a,c) and glycolytic fibers (Figure 22b,c) – calculated as the product of muscle mass and the areal density of each fiber type –were larger in males than females (Table 2). Oxidative fiber mass did not differ between male age classes, but glycolytic fiber masses were consistently larger in adult males than subadult males (Figure 22b,c; Table 3), reflecting increased investment by adult males in glycolytic fibers. Among species, *T. rufilatus* had the largest oxidative fiber mass (Figure 22a,c), which is consistent with a more flighted ecology – *T. rufilatus* has longer wings and shorter legs than the other two species (Figure 23). *T. indicus* and *T. chrysaeus*, however, have similar wing shapes and tarsus lengths (Figure 23), and had strikingly similar

oxidative fiber mass, despite *T. indicus* having larger *pectoralis* muscles (Figure 22a,c; Table 1). The differences in *pectoralis* mass between *T. indicus* and *T. chrysaeus* are due to increased investment of *T. indicus* in glycolytic fibers, rather than changes in oxidative fiber mass (Figure 22; Table 1).

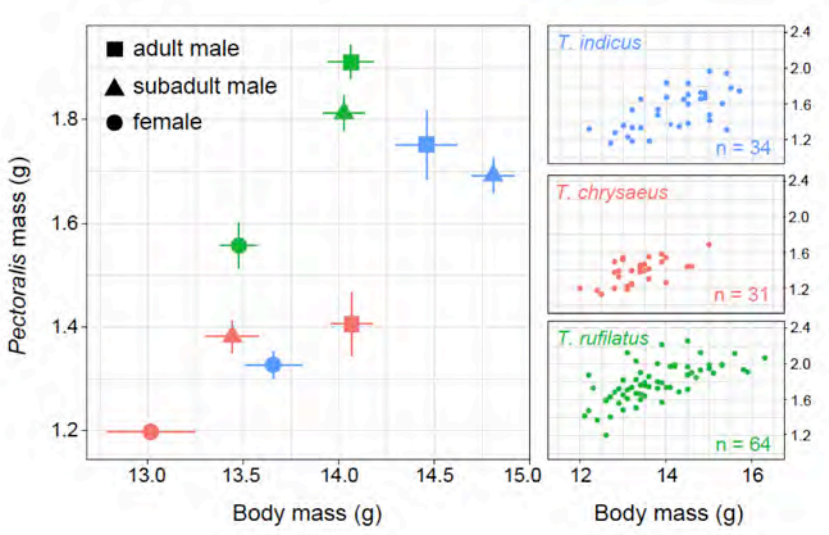

Figure 20 | *Pectoralis* muscle mass plotted against body mass. Left panel shows species means (with standard error bars), plotted by age and sex classes. Species are plotted by color (*T. indicus* = blue, *T. chrysaeus* = red, *T. rufilatus* = green). Right panels are broken down by species to show values for each individual. Each point in the right panels represents an individual in our sample.

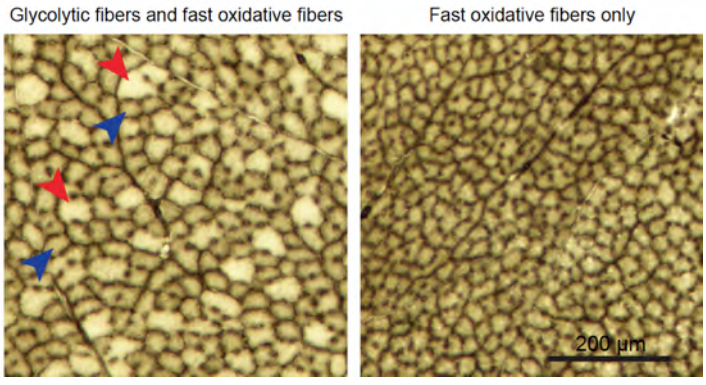

Figure 21 | Histochemical staining of bush-robin *pectoralis* muscle, showing representative cross-sections of individuals with both glycolytic fibers and fast oxidative fibers (left panel) or only fast oxidative fibers (right panel). Muscles were stained with Myosin-ATPase to identify fiber types. Red arrows show examples of fast glycolytic fibers. Blue arrows show examples of fast oxidative fibers. The *pectoralis* cross-section in the right image is comprised of only fast oxidative fibers.

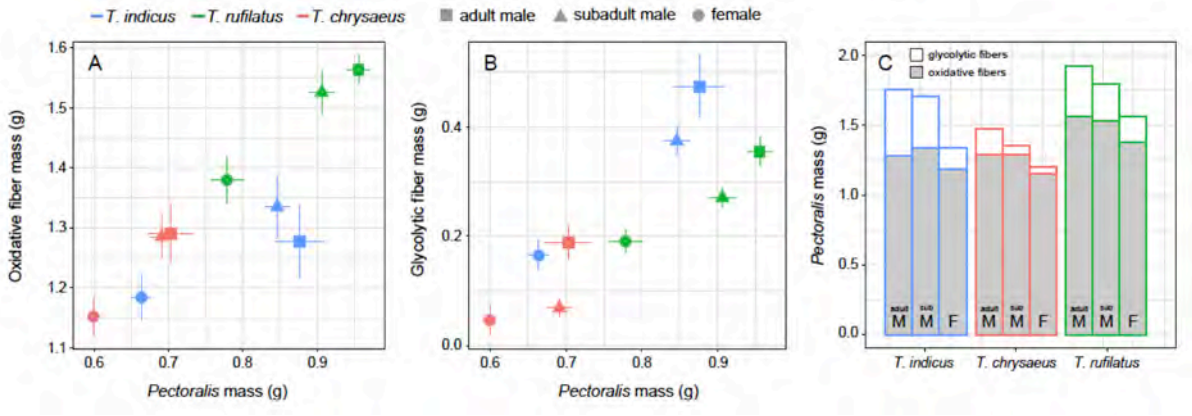

Figure 22 | The relative contributions of fast oxidative fiber mass and glycolytic fiber mass to overall *pectoralis* mass. Panels A and B show species means (with standard error bars), plotted by age and sex classes. Species are plotted by color. Panel C show the ratio of each fiber type by age and sex class within each species (the mean of each fiber type was used).

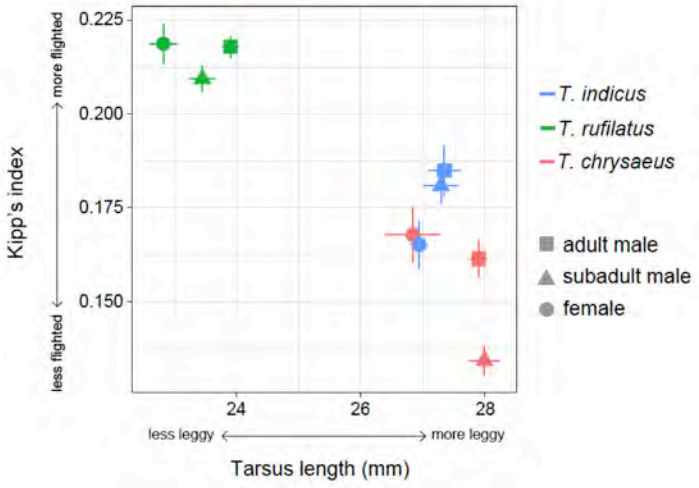

Figure 23 | Morphometric data showing that *T. rufilatus* (green) has longer, pointer wings and shorter legs than *T. indicus* (blue) and *T. chrysaeus* (red), suggesting different ecologies. Species means (with standard error bars) were plotted by age and sex classes (adult males = squares, subadult males = triangles, females = circles) for tarsus and Kipp's index, which is also known as the hand-wing index. Kipp's index is a common measurement of wing aspect ratio, which is strongly linked to flight capacity (Lockwood et al., 1998). Kipp's index is reported on a size-independent scale (0-1) from short, rounded wings to long narrow wings (Kipp, 1959). Kipp's index is calculated as: (length of longest primary feather – length of first secondary feather) / length of longest primary feather.

Table 1 | Species differences in body mass, flight muscle phenotypes, and experienced temperature, analyzed using a one-way analysis of covariance (ANCOVA, type III sums of squares) with class identity (adult male, subadult male, female) as a covariate. Pairwise species comparisons were performed using Tukey's post-hoc tests for each trait, and p-values are reported in the table. P-values are bolded if <0.05. "raw" = mass of the specified trait uncorrected for body mass. "relative" = mass of the specified trait analyzed as a proportion of body mass.

	Pairwise species comparison (p-value)		
trait	*T. indicus – T. chrysaeus*	*T. indicus – T. rufilatus*	*T. rufilatus – T. chrysaeus*
body mass	**2.90E-06**	**0.0001**	0.1659
pectoralis mass (raw)	**0.0000**	**5.00E-07**	**0.0000**
pectoralis mass (relative)	**4.16E-05**	**0.0000**	**0.0000**
fast oxidative fiber mass (raw)	0.5520	**0.0000**	**0.0000**
fast oxidative fiber mass (relative)	0.1482	**0.0000**	**0.0000**
fast glycolytic fiber mass (raw)	**0.0000**	0.0650	**0.0000**
fast glycolytic fiber mass (relative)	**0.0000**	**0.0000**	0.1451
maximum daily temperature	**0.0061**	0.5920	**0.0139**
minimum daily temperature	**0.0002**	**0.0035**	0.2130

Table 2 | Sex differences in body mass, flight muscle phenotypes, and experienced temperature, analyzed with a one-way analysis of variance (ANOVA, type III sums of squares). Means and standard errors for each trait are reported for by sex each species. P-values are bolded if <0.05. "raw" = mass of the specified trait uncorrected for body mass. "relative" = mass of the specified trait analyzed as a proportion of body mass.

trait	*T. indicus* female	*T. indicus* male	F	P	*T. chrysaeus* female	*T. chrysaeus* male	F	P	*T. rufilatus* female	*T. rufilatus* male	F	P
body mass (g)	13.66 ± 0.15	14.70 ± 0.10	37.24	**5.63E-08**	13.02 ± 0.23	13.88 ± 0.09	7.15	**0.0092**	13.48 ± 0.10	14.05 ± 0.08	19.58	**1.68E-05**
pectoralis mass (g) (raw)	1.327 ± 0.019	1.709 ± 0.020	90.41	**7.67E-11**	1.198 ± 0.013	1.391 ± 0.018	6.10	**0.0196**	1.557 ± 0.021	1.864 ± 0.016	38.83	**4.51E-08**
pectoralis mass (%) (relative)	9.755 ± 0.139	11.673 ± 0.135	47.27	**8.84E-08**	9.253 ± 0.074	10.477 ± 0.095	8.80	**0.0061**	11.825 ± 0.118	13.339 ± 0.097	26.28	**3.22E-06**
fast oxidative fiber mass (g) (raw)	1.184 ± 0.025	1.317 ± 0.025	5.64	**0.0249**	1.152 ± 0.025	1.287 ± 0.016	3.90	0.0599	1.380 ± 0.017	1.548 ± 0.012	16.09	**0.0002**
fast oxidative fiber mass (%) (relative)	8.708 ± 0.177	8.933 ± 0.147	0.40	0.5309	8.891 ± 0.121	9.643 ± 0.104	2.74	0.1107	10.430 ± 0.100	11.176 ± 0.089	6.86	**0.0115**
fast glycolytic fiber mass (g) (raw)	0.165 ± 0.018	0.406 ± 0.017	38.22	**1.31E-06**	0.046 ± 0.022	0.112 ± 0.011	1.90	0.1812	0.191 ± 0.009	0.319 ± 0.011	14.61	**0.0003**
fast glycolytic fiber mass (%) (relative)	1.216 ± 0.132	2.774 ± 0.128	29.23	**1.02E-05**	0.362 ± 0.171	0.832 ± 0.079	1.77	0.1964	1.436 ± 0.062	2.244 ± 0.072	13.56	**0.0005**
maximum daily temperature (C)	10.37 ± 0.78	7.96 ± 0.67	5.23	**0.0256**	11.07 ± 1.98	11.51 ± 0.71	0.03	0.8611	8.50 ± 0.53	10.44 ± 0.43	8.03	**0.0051**
minimum daily temperature (C)	1.24 ± 0.42	–0.56 ± 0.23	15.20	**0.0002**	1.75 ± 1.22	1.52 ± 0.25	0.06	0.8138	0.92 ± 0.22	1.37 ± 0.21	2.06	0.1530

Table 3 | Male age class differences in body mass, flight muscle phenotypes, and experienced temperature, analyzed with a one-way analysis of variance (ANOVA, type III sums of squares). Means and standard errors for each trait are reported age class for each species. P-values are bolded if <0.05. "raw" = mass of the specified trait uncorrected for body mass. "relative" = mass of the specified trait analyzed as a proportion of body mass.

trait	*T. indicus* adult male	*T. indicus* subadult male	F	P	*T. chrysaeus* adult male	*T. chrysaeus* subadult male	F	P	*T. rufilatus* adult male	*T. rufilatus* subadult male	F	P
body mass (g)	14.46 ± 0.16	14.80 ± 0.12	2.69	0.1092	14.07 ± 0.10	13.44 ± 0.14	11.04	**0.0014**	14.06 ± 0.12	14.04 ± 0.11	0.03	0.8708
pectoralis mass (g) (raw)	1.752 ± 0.043	1.693 ± 0.023	0.77	0.3923	1.406 ± 0.027	1.382 ± 0.026	0.15	0.7014	1.911 ± 0.023	1.813 ± 0.021	4.37	**0.0421**
pectoralis mass (%) (relative)	12.27 ± 0.26	11.44 ± 0.14	3.97	0.0636	10.69 ± 0.14	10.36 ± .15	0.95	0.3383	13.68 ± 0.17	12.98 ± 0.08	6.12	**0.0172**
fast oxidative fiber mass (g) (raw)	1.277 ± 0.039	1.335 ± 0.033	0.43	0.5243	1.290 ± 0.019	1.285 ± 0.029	0.01	0.9355	1.564 ± 0.016	1.526 ± 0.019	0.85	0.3634
fast oxidative fiber mass (%) (relative)	8.942 ± 0.222	8.930 ± 0.195	0.00	0.9821	9.519 ± 0.143	9.715 ± 0.181	0.24	0.6302	11.263 ± 0.148	11.065 ± 0.103	0.44	0.5115
fast glycolytic fiber mass (g) (raw)	0.475 ± 0.036	0.375 ± 0.017	3.27	0.0921	0.188 ± 0.012	0.069 ± 0.014	13.92	**0.0013**	0.355 ± 0.019	0.270 ± 0.009	5.96	**0.0192**
fast glycolytic fiber mass (%) (relative)	3.328 ± 0.256	2.522 ± 0.126	4.14	0.0613	1.386 ± 0.092	0.516 ± 0.100	13.29	**0.0016**	2.477 ± 0.124	1.948 ± 0.056	5.40	**0.0255**
maximum daily temperature (C)	5.62 ± 0.98	8.79 ± 0.80	4.60	**0.0389**	10.82 ± 0.83	13.22 ± 1.30	2.17	0.1455	9.34 ± 0.61	11.32 ± 0.58	5.50	**0.0208**
minimum daily temperature (C)	–1.34 ± 0.30	–0.28 ± 0.28	4.25	**0.0466**	1.21 ± 0.28	2.30 ± 0.52	3.53	0.0647	0.91 ± 0.30	1.74 ± 0.47	4.02	**0.0473**

3.2.2 Differential timing of arrival to breeding elevations

Bush-robins differed within and among species in their timing of migration to breeding elevations. Arrival order among species was associated with wintering elevation and seasonal temperature – species that wintered higher arrived earlier to breeding elevations (Figures 24a and 25), and birds that arrived earlier experienced colder temperatures (Figures 24b and 26; Tables 1-3). *T. indicus* arrived first, followed by *T. rufilatus* and then *T. chrysaeus*. Within each species, we observed the same qualitative patterns of age- and sex-structured migration, in which adult males arrived before subadult males, which together arrived before females (Figure 24b). In each sampling year, some males of *T. indicus* were already singing above 3000 m on the breeding grounds when we began sampling, indicating that these individuals had already passed through our sampling area. This observation suggests that we are conservatively estimating the arrival

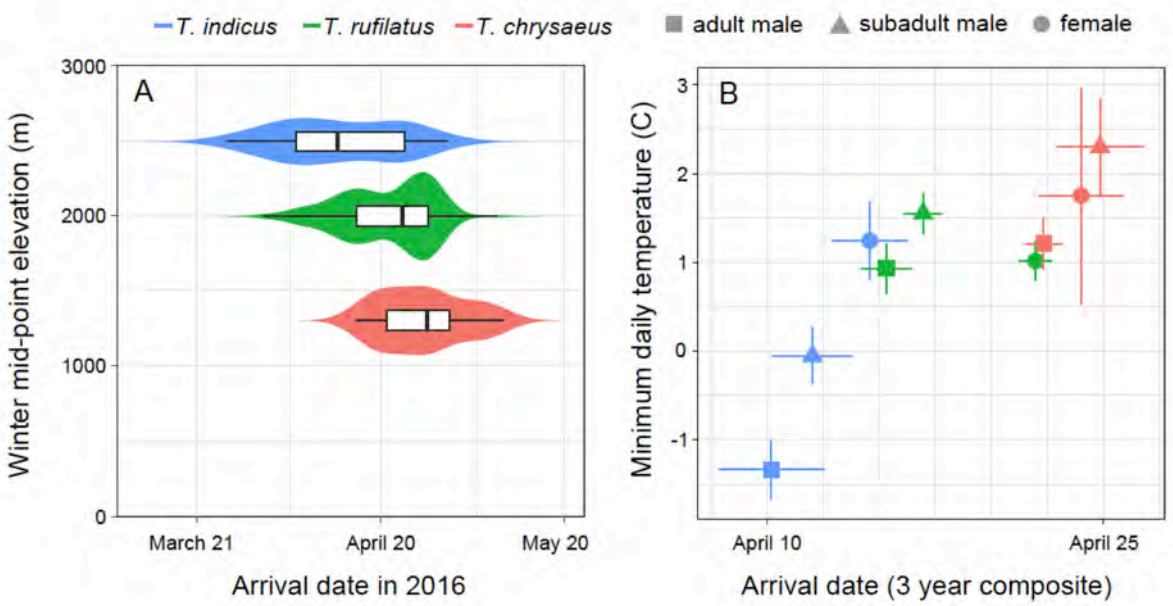

Figure 24 | Birds that winter higher arrived first to breeding elevations when temperatures were coldest. Species are plotted by color. (A) Elevational mid-point of species wintering distribution plotted against the distribution of arrival date to 3000 m elevation for each species in 2016. Similar plots for years 2014 and 2015 can be found in Figure 25. Box plots show median date and quartiles of data. Violin plots show distribution of the data. (B) Minimum daily temperature on the day each individual was captured, averaged over the three sampling years. Species means (with standard error bars) are plotted by age and sex classes.

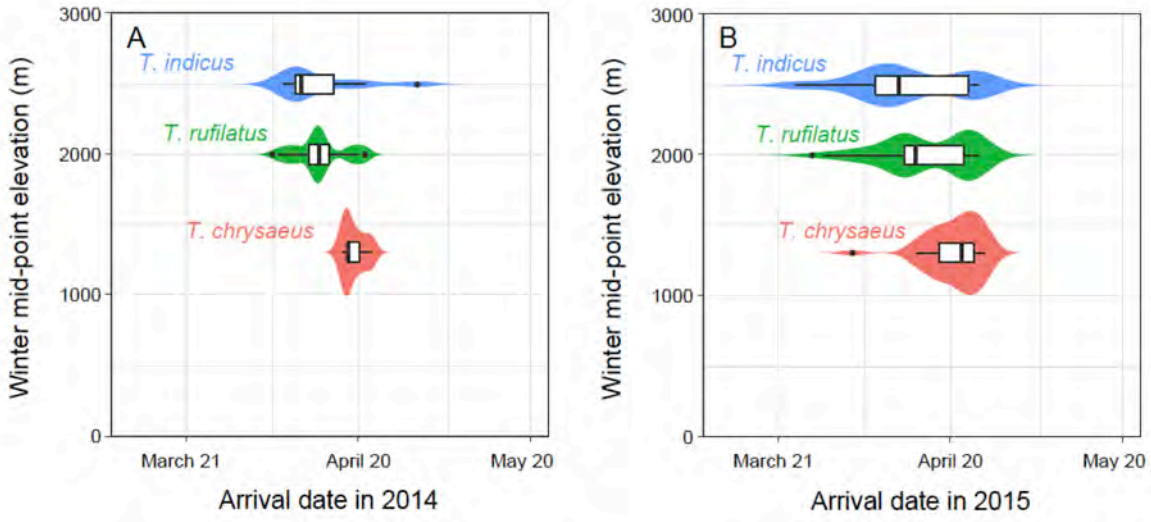

Figure 25 | Birds that winter higher arrived first to breeding elevations. Species are plotted by color. Elevational mid-point of species wintering distribution plotted against the distribution of arrival date to 3000 m elevation for each species in 2014 (A) and 2015 (B). Box plots show median date and quartiles of data. Violin plots show distribution of the data.

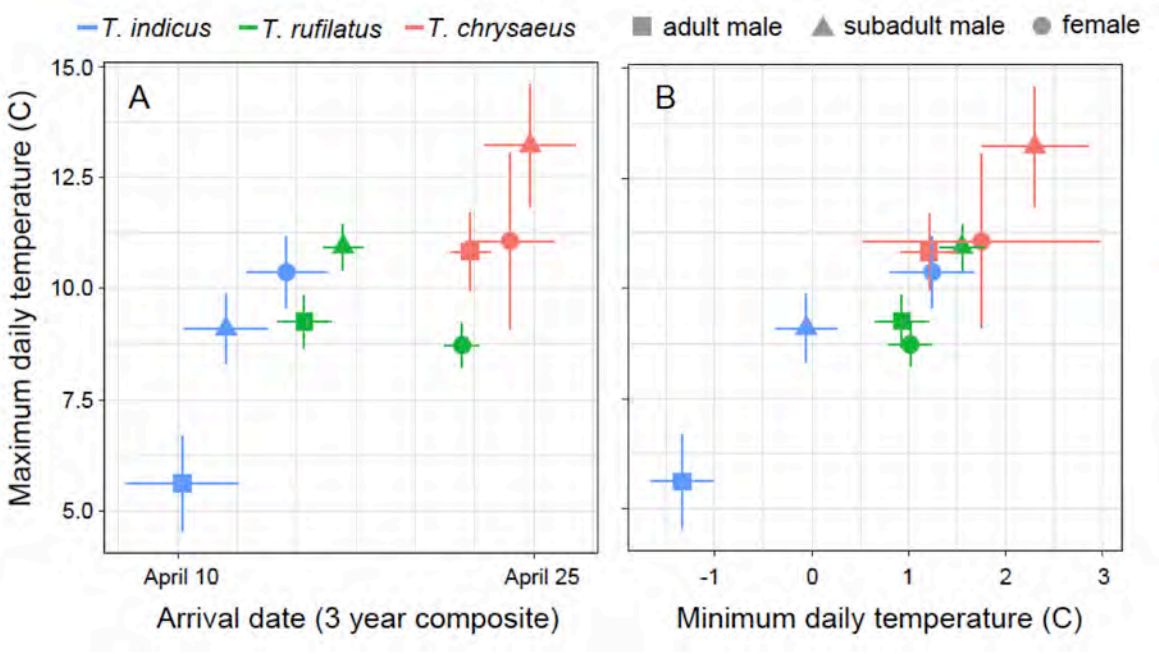

Figure 26 | Birds that arrived first to breeding elevations experienced colder temperatures. Species means (with standard error bars) are plotted by age and sex classes. (A) Maximum daily temperature on the day each individual was captured, averaged over the three sampling years. (B) Maximum daily temperature plotted over minimum daily temperature on the day each individual was captured.

time of *T. indicus* males as well as the disparity in arrival time and experienced temperature between *T. indicus* age and sex classes. Despite this conservative estimate of male arrival in *T. indicus*, we still recovered a signal of age- and sex-structured variation in arrival time and experienced temperature.

3.2.3 Association of flight muscle phenotypes with arrival and seasonal temperature

Species, age, and sex classes caught earlier in spring, when temperatures were colder, generally had larger body mass and *pectoralis* muscles (Figure 27a-b). The association between *pectoralis* mass and temperature was largely driven by differences in glycolytic fiber mass among and within species. That is, species, age, and sex classes that were caught earlier have increased glycolytic fiber mass in their *pectoralis* (Figure 27c). This result is counterintuitive, as glycolytic

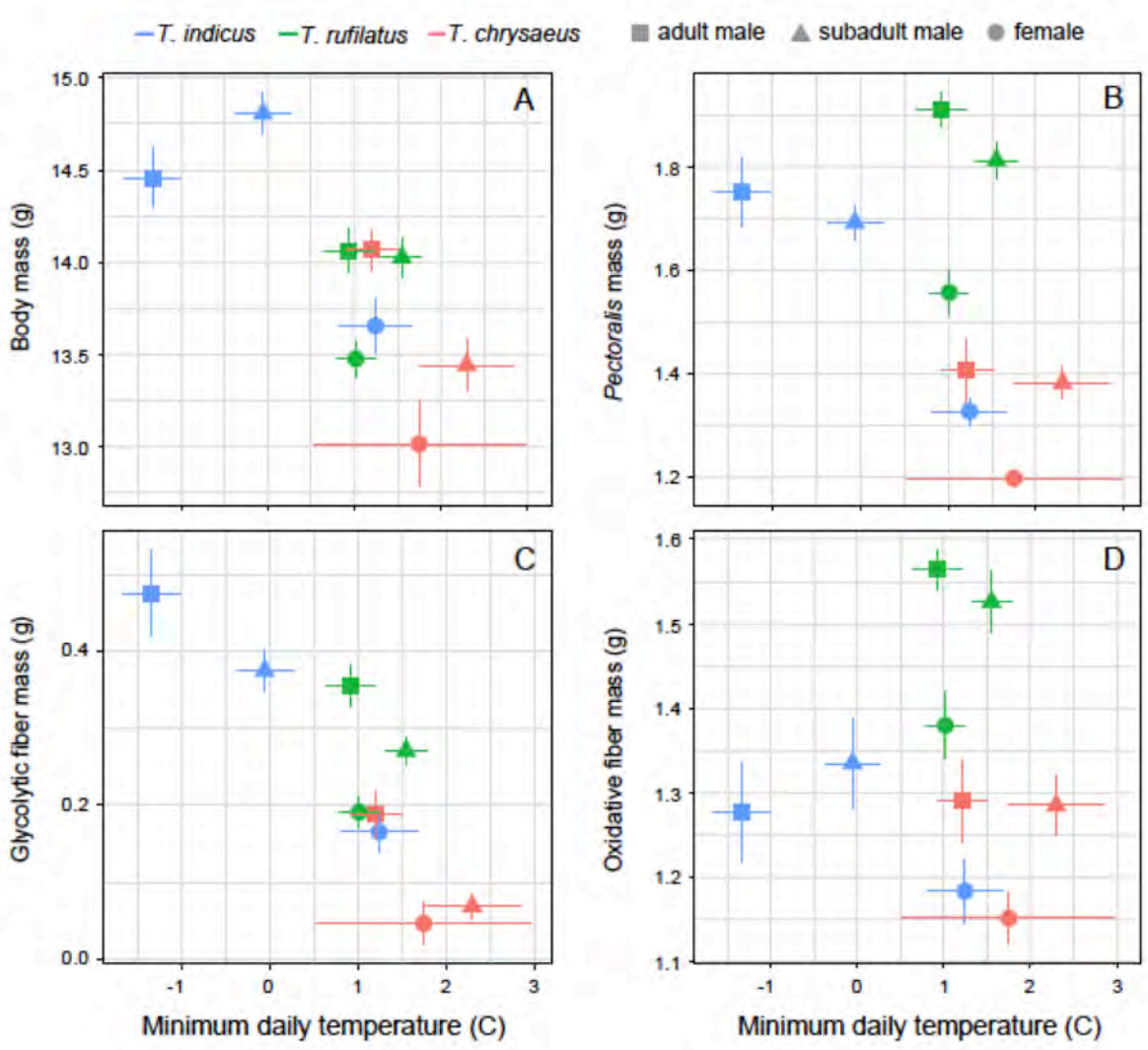

Figure 27 | Mass and *pectoralis* phenotypes plotted against minimum daily temperature on the day each individual was caught. Species means (with standard error bars) are plotted by age and sex classes. Species are plotted by color. Panels C and D plot the fiber-specific masses of the *pectorals*.

fibers likely contribute little to thermogenesis (Marsh and Dawson, 1989), yet individuals that experience colder temperatures are investing resources in glycolytic fibers, which could be otherwise invested in oxidative fibers. The larger oxidative fiber mass in males than females was generally associated with earlier arrival and colder temperatures, but differences among species in oxidative fiber mass were not (Figure 27d).

3.2.4 Differential sensitivity to cold stress and extreme weather events

We found that bush-robins have the ability to abandon breeding territories in early spring in response to acute cold weather events. In response to acute cold weather events, bush-robins shifted downslope or halted migration upslope, returning to breeding elevations as soon as temperatures warmed (Figure 28). The propensity to abandon breeding elevations, however, varied by species, age, and sex in association with wintering elevation and arrival order. Species that winter higher were more likely to persist at breeding elevations through acute cold weather events, as were adult males (Figure 28). This behavioral flexibility confounds and introduces

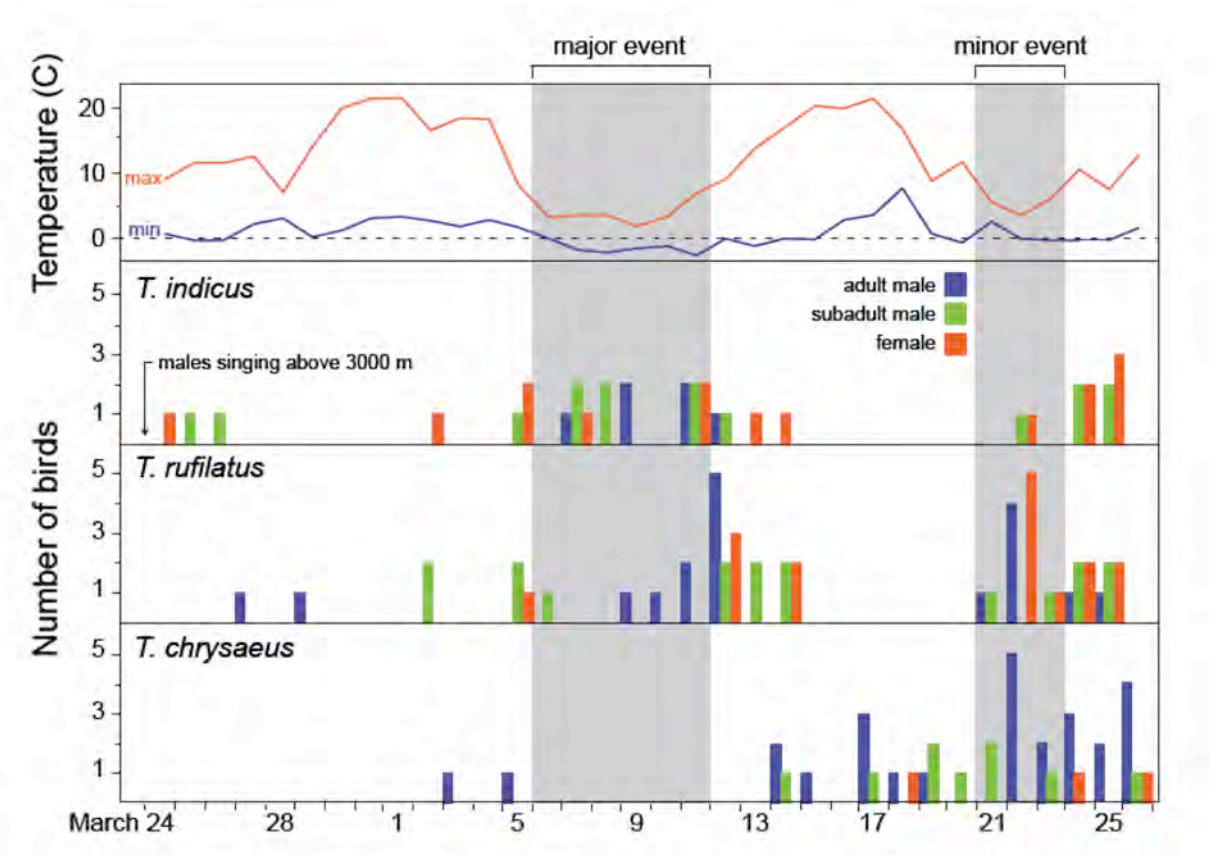

Figure 28 | Differential behavioral response to cold weather events in 2015 in the three bush-robin species. The colored bars denote the number of individuals caught at 3000 m elevation on a given day. Shaded areas denote the two acute cold weather events. *T. indicus* males were already singing above 3000 m elevation when we began netting on March 24th.

noise into our arrival time data because not all captures at 3000 m reflect an individual's first arrival for the given season. Nevertheless, despite this noise, we were still able to recover species-, age-, and sex-structured differences in occurrence patterns at 3000 m elevation (the lower elevational breeding limit) that are associated with wintering elevations and temperature.

3.3 Discussion

By integrating seasonal climate data with physiological and behavioral data across demographic axes, we discovered age- and sex-structured variation in flight muscle phenotypes that is associated with the distinct functional demands and selective pressures imposed on each demographic class. Furthermore, the *pectoralis* of bush-robins contained both fast oxidative and glycolytic fibers, implying that the *pectoralis* of small birds may not always be constrained by thermogenic demands to contain only fast oxidative fibers, as previously thought (Welch and Altshuler, 2009; Dakin et al., 2018). In fact, the species that winters the highest, arrives first, and experiences the coldest temperatures had the highest contribution of glycolytic fibers to overall *pectoralis* mass, suggesting that the selective benefits of accruing glycolytic fibers in bush-robins outweigh investment in oxidative fibers that may directly enhance cold tolerance. The prevalence of glycolytic fibers varied by age and sex within each species, further suggesting a functional role of glycolytic fibers in small passerines that has been previously unrecognized. By analyzing flight muscle phenotypes within the context of adaptive life history strategies, like age- and sex-structured migration and delayed plumage maturation, we can better understand the selective environments under which variation in flight muscle phenotypes arise and are maintained, which in turn provides insight into the mechanisms underlying these life history strategies.

3.3.1 A novel function of glycolytic fibers

The discovery of glycolytic fibers in the *pectoralis* of bush-robins is striking given their small body size (<17 g) because the *pectoralis* of most small birds had previously been found to contain only fast oxidative fibers (Rosser and George, 1985; Welch and Altshuler, 2009), and the presence of glycolytic fibers in bush-robins suggests an unrecognized functional role for these fibers. In larger birds, glycolytic fibers are essential for explosive tasks, like takeoff, because high mechanical power is required to lift larger birds into the air (Dial et al., 1987; Askew et al., 2001). Smaller birds, however, are less constrained by mass-specific demands on fiber type for takeoff, and fast oxidative fibers were previously thought to meet all of their functional demands of powered flight (Welch and Altshuler, 2009). To date, glycolytic fibers have been observed in the *pectoralis* of one other species under 20 g, the European Robin (*Erithacus rubecula*), a close relative of *Tarsiger* bush-robins (Sangster et al., 2010). It is possible, however, that these fibers may be more widespread since fiber composition has been characterized in less than 5% of bird species. If glycolytic fibers are not necessary for takeoff in small birds, then their presence should be associated with alternative functional demands on flight, like burst flight associated with competitive interactions. The high mechanical power of glycolytic fibers would enhance short-burst flight capacity, improving an individual's ability to secure and defend breeding territories and mates. Under this hypothesis, we would expect to find increased glycolytic fibers in dominant and territorial individuals, which we found support for in all three bush-robin species (Figure 28b). Glycolytic fibers may thus be recruited for competitive interactions in small birds, representing a unique functional context in which variation in muscle phenotypes in small birds is tied to life-history trade-offs and social structure.

3.3.2 Sex differences in selection on flight muscle function

Male bush-robins arrived to breeding elevations before females (Figure 24b), which is a common migratory strategy among birds. This life-history strategy is assumed to arise through sex-specific selection that increases reproductive success in one or both sexes (Morbey and Ydenberg, 2001). A number of hypotheses have been proposed to explain the adaptive significance of sex-structured migration. For example, territorial males may be under selection to arrive earlier to breeding grounds to establish territories (Ketterson and Nolan, 1976; Myers, 1981). This hypothesis is supported by a preponderance of work showing early-arriving males achieve higher reproductive success (e.g., Perrins ,1970; Aebischer et al., 1996; Lozano et al., 1996). In addition, selection to arrive later to avoid unfavorable environmental conditions in early spring could act more strongly on females (Ketterson and Nolan, 1983; Francis and Cooke, 1986). This hypothesis is particularly notable in our temperate, high-elevation context because female bush-robins should disproportionately benefit from investing resources directly into reproductive traits, such as egg formation, rather than investing in muscle traits that enhance cold tolerance. Although the fitness consequences of sex-structured migration in bush-robins remain unknown, this adaptive life history strategy exposes males and females to different environmental and social pressures (Figure 24b). Male bush-robins should thus benefit from flight muscle traits that enhance thermogenic capacity and competitive ability, facilitating early arrival and territoriality, while females should benefit from later arrival when conditions are more favorable on breeding grounds, allowing them to maximize direct investment in reproductive traits.

The sex differences we observed in flight muscle phenotypes are consistent with the sex-dependent selection associated with sex-structured migration. Male bush-robins showed adaptive

increases in fatigue-resistant oxidative fibers, which enhance shivering thermogenesis. Males also showed increases in powerful glycolytic fibers, which enhance short-burst flight capacity and territoriality. These results support the hypothesis that glycolytic fibers play an important role in social dynamics and competitive interactions in bush-robins. Increases in both muscle fiber types have evolved in concert with the environmental and social demands on males, highlighting how competing selective pressures on function can be balanced in the flight muscle. In contrast, females show reduced investment in both fast oxidative and glycolytic fibers, and have correspondingly smaller flight muscles, in association with their later arrival to breeding elevations. By arriving when spring temperatures are warmer, and after males have established territories, females experience reduced pressure to allocate resources into flight muscles that increase cold tolerance and competitive traits. Instead, females can invest resources in reproductive traits outside of the flight muscle. These results underscore how the tensions shaping flight muscle phenotypes differ between sexes. In males, tension exists within the flight muscle between traits that enhance thermogenesis and competitive ability, but in females greater tension exists between the flight muscle and reproductive traits outside of the muscle.

3.3.3 Age differences in social dominance and flight muscle function

The observed differences in glycolytic fiber mass between male age classes further support the hypothesis that glycolytic fibers are important in social dynamics and competitive interaction, providing functional evidence of reduced competitive ability in subadult males. Adult males showed increased glycolytic fiber mass compared to subadult males (Figure 22b). These findings are consistent with delayed plumage maturation being an honest signal of competitive ability, and uncover age-structured variation in muscle morphology that likely translates to reduced

competitive function in subadult males. If males are under selection to arrive early to secure breeding territories, then traits that enhance territory defense and competitive ability, like glycolytic fibers, should also be under selection to increase. The reduced level of glycolytic fibers in subadult males likely reduces their competitive ability, which is further supported by their migratory behavior and plumage phenotype. If subadult males are competitively inferior to adult males as a consequence of differences in glycolytic capacity, then subadult males would benefit from avoiding costly interactions with dominant males, which they appear to achieve by arriving later to breeding elevations and retaining female-like plumage. Delayed plumage maturation allows subadult males to effectively signal their inferior competitive ability to avoid aggressive interactions (Hawkins et al., 2012). Our results contend that it is not only experience in breeding and fighting that makes subadult males inferior (Forslund and Pärt, 1995), but that they invest less in various flight muscle traits, rendering them less competitively able than their adult counterparts. These findings provide some of the first evidence for variation in performance traits associated with delayed plumage maturation in which age-structured variation in flight muscle phenotypes and behavior are associated with the inferior competitive ability of subadult males.

3.3.4 Behavioral flexibility in response to acute cold stress

If males are under selection to arrive to breeding elevations early, then the behavioral flexibility in bush-robins to respond to acute climatic stress would facilitate survival at upper elevation earlier in spring when environmental conditions are more variable. In early spring when environmental conditions are unpredictable, individuals can migrate upslope to breeding elevations, but drop to lower elevations if conditions turn unfavorable. This behavioral flexibility

blurs the traditional views of arrival time, but highlights the dynamic nature of temperate montane systems. In early spring, males are establishing territories and learning neighbors, so there is a real benefit to arriving early to breeding elevations if the costs of abandoning a territory are not large. Because altitudinal migrants, like bush-robins, can access optimal climates over short geographic space, these dynamic, short-distance movements are a viable strategy that facilitates early arrival to breeding grounds in unpredictable environments (O'Neill and Parker, 1978; Hahn et al., 2004). The fact that adult males of *T. indicus* remained at breeding elevations during one of the observed cold weather events suggests a cost to abandoning territories. Furthermore, the differential response between species, age, and sex classes to acute cold stress suggests that demographic-based differences in thresholds for cold avoidance drive decision-making. These thresholds could be temporal, physiological, or a combination of both. For example, individuals might be less willing to abandon a territory in early spring if they have invested more time and energy into establishing that territory, and may only leave once they can no longer persist survive the cold temperatures. Given the behavioral patterns that we observed, these thresholds appear to vary in accordance with flight muscle traits along taxonomic and demographic axes of selection.

3.3.5 Species differences in ecology and flight muscle function

We found that species that arrive earlier in spring have larger *pectoralis* muscles than later arrivals, which is consistent with past work showing that cold tolerance increases with flight muscle size (Swanson et al., 2013; Petit and Vezina, 2014). This relationship between flight muscle mass and cold tolerance, however, has been previously defined in species whose *pectoralis* muscles contain only fast oxidative fibers. In bush-robins, when we consider the

relative contributions of fiber type to overall *pectoralis* mass and function, *T. indicus* and *T. chrysaeus* have strikingly similar oxidative fiber mass (Figure 22a,c), despite experiencing dramatically different thermal environments. The differences in *pectoralis* mass between *T. indicus* and *T. chrysaeus* are thus not due to changes in oxidative fiber mass, but rather differences in glycolytic fiber mass, which likely contribute little to sustainable thermogenesis (Marsh and Dawson, 1989). These results underscore the need to account for fiber-type composition when interpreting flight muscle function.

Among species, the staggered pattern of migration and the presumed differences in cold sensitivity that it reflects, suggest that there is a connection between seasonal distributions, phenological events, and thermal tolerance. These data suggest that *T. indicus* is more cold tolerant than *T. rufilatus*, which is more cold tolerant than *T. chrysaeus*. Under thermogenic predictions alone, we would then expect *T. indicus* to have *pectoralis* phenotypes that confer adaptive increases in cold tolerance, such as increases in the mass of fast oxidative fibers. *T. rufilatus*, however, had the highest oxidative fiber mass. This discrepancy can be reconciled when we consider species-specific ecology and that a number of other traits within the flight muscle and organism work in concert to define whole-organism metabolism and capacities for thermogenic and locomotory performance. First, *T. rufilatus* has a more flighted ecology than the other two species, as suggested by morphometric data (Figure 23), which likely drives increased oxidative fiber mass in *T. rufilatus*. Second, the predicted differences in cold tolerance among species may be achieved through changes in other traits that work synergistically with muscle fibers to impact metabolic capacity at upper elevations, like hemoglobin oxygen-binding affinity (Projecto-Garcia et al., 2013; Zhu et al., 2018), cardiorespiratory O_2 transport, the activity of enzymes in metabolic pathways (Stager et al., 2014), and the position of mitochondria within the

muscle fibers (Scott et al., 2009), among many others. If *T. indicus* has adaptive modifications in other traits that confer increases in cold tolerance, then there might be less constraint against fiber identity evolving along other functional axes. Modifications in these other traits could also vary within species along demographic axes, thus offering promising directions for future research. Incorporating a broader suite of traits that interact with flight muscle phenotypes will further shape our understanding of how organisms balance the competing functional demands on flight muscle imposed by thermogenesis and powered flight.

3.3.6 Seasonal plasticity in flight muscle phenotype

Species that experience colder temperatures showed increased glycolytic fiber mass, which is striking given that glycolytic fibers contribute little to sustained shivering thermogenesis (Marsh and Dawson, 1989). This finding highlights the potential importance of glycolytic fibers for social interactions, and may reflect seasonal fluctuations in hormones that underlie plasticity of flight muscle phenotypes. In birds, muscle growth is directly linked to breeding hormones, like testosterone, which become elevated in the breeding season (Wingfield et al., 1987). If glycolytic fibers are important for securing territories and mates, as our data suggest, then males might be incorporating more glycolytic fibers into their *pectoralis* as breeding activities are initiated. Under this hypothesis, we may have captured a temporal snapshot of muscle phenotypes where each species is sampled at a different relative breeding stage, with *T. indicus* having already initiated breeding activities, and *T. rufilatus* and *T. chrysaeus* each lagging behind.

Past work on male European Robins, a close relative to bush-robins, supports the above hypothesis, showing that testosterone levels and glycolytic fiber mass both increase during the breeding season compared to other times of the year, such as during migration (Lundgren, 1988;

Schwabl, 1992). This seasonal plasticity in flight muscle phenotypes further highlights the power of within-species comparisons and integrative datasets in helping interpret the functional role of phenotypic variation. Our behavioral data suggest that the three bush-robins were at different stages of breeding activity at the time of capture: *T. indicus* males were already singing on territories at the start of each sampling period, while *T. chrysaeus* had not yet reached breeding elevations. Under this hypothesis, *T. rufilatus* and *T. chrysaeus* should undergo similar increases in glycolytic fibers as their respective breeding seasons progressed, and although seasonal lags in glycolytic fiber development might also explain differences between age classes within each species, subadults would still have lower levels of glycolytic fibers when territories are first being established.

3.4 Materials and methods

3.4.1 Study site, climate data, and sampling birds

We conducted this study at 3000 m elevation in Hailuogou Valley, Sichuan, China (29.576096°N, 101.998624°E). Hailuogou Valley drains the eastern slope of Mount Gongga, the easternmost peak in Asia above 7000 m elevation, spanning a sharp elevational gradient from 1400-7556 m in under 25km (Figure 19). Birds were caught using mist-nets over three consecutive spring seasons, from April 5–April 22 in 2014, March 24–April 26 in 2015, and March 24–May 10 in 2016. During these days, 8-13 mist nets were open from sun-up (+/- 1hr) to sundown (+/- 1hr). Maximum and minimum daily temperatures for our study site were collected from the climate station at 3000 m run by the Institute of Mountain Hazards and Environment, Chinese Academy of Science. Birds were netted and collected with authorization from Gongga

Shan National Nature Reserve and Hailuogou Forestry Bureau. Birds were handled in accordance with the Institutional Animal Care and Use Committee at the University of Chicago.

Mass was recorded for all individuals upon capture, after which birds were either sacrificed for characterization of *pectoralis* phenotype or immediately released back into the environment at 3000 m. The maximum and minimum temperatures were recorded for each individual from the day they were captured. For all birds that were released, we uniquely trimmed one to two tail feathers (<1cm of the tip) to identify if they were recaptured in a given year. Bush-robins annually molt tail feathers, so it is unknown if released individuals were recaptured across years. Within a given year, however, recapture rates were relatively low – only 10% of released individuals were caught more than once (averaged over the 3 years). For individuals that were caught more than once in a year, we report body mass from the day the individual was first caught. However, we report temperature values from the coldest day an individual was caught (determined from the minimum daily temperature) to specifically test hypotheses associated with cold sensitivity.

3.4.2 *Pectoralis* mass and histology

For a subset of captured birds (n =129), we measured *pectoralis* mass and quantified fiber composition within the *pectoralis* muscle. In these individuals, the right flight muscles were dissected out within minutes of being euthanized. The *pectoralis* was then separated from the *supracoracoideus* and the entire right *pectoralis* was weighed. We report total organism *pectoralis* mass as the right *pectoralis* multiplied by two. We then dissected out the middle third of the right *pectoralis* (in relation to the length of the keel), coating this muscle section in O.C.T. embedding medium, which was then flash-frozen in liquid nitrogen-cooled isopentane.

Embedded muscle was then sealed in aluminum foil and stored in liquid nitrogen until being transferred into a -80 C freezer, where they were stored until being sectioned.

Muscle fiber types were assessed in the *pectoralis* using methods described by Deveci et al. (2001). *Pectoralis* muscles were sectioned (10 um) in a cryostat at -20 C, transverse to the length of the muscle fibers. Sections were then stained for myosin-ATPase (pre-incubation at pH 4.4) and succinate dehydrogenase (SDH) activity to distinguish muscle fiber types (Deveci et al., 2001). We imaged the stained sections with a light microscope, and we used stereological methods described by Weibel (1979) and Egginton (1990) to make unbiased measurements of areal fiber density. For each sample, we analyzed 12 images to account for fiber heterogeneity within the muscle, which we determined in preliminary measurements to be a sufficient number of images to achieve a stable mean value for an individual bird. The mean areal densities of each fiber type for an individual (i.e. the relative areal ratios of each fiber type) were then multiplied by the *pectoralis* mass to get the mass of each fiber type in the *pectoralis*. Fiber type was preferentially assessed using myosin-ATPase stained sections unless these sections were unavailable (due to tissue damage during staining procedure). In these cases we quantified fiber type from SDH stained sections. We confirmed that fiber-typing outcomes were consistent between the two staining methods by comparing a subset of individuals in which both staining methods were analyzed.

3.4.3 Vouchered specimens and morphometric measurements

All birds collected for *pectoralis* phenotyping are vouchered as specimens at the National Zoological Museum of China or the Nature Museum of Sichuan University, except for six individuals that were stolen by a Siberian weasel (*Mustela sibirica*) at our study site in 2014.

These specimens were measured for tarsus length and Kipp's index (i.e. wing aspect ratio) as a proxy for ecological differences among species. Tarsus length and Kipp's index are standard morphological measurements that are strongly linked to ecology (Lockwood et al., 1998). Kipp's index is reported on a size-independent scale (0-1) from short, rounded wings to long narrow wings (Kipp, 1959). Kipp's index is calculated as: (length of longest primary feather – length of first secondary feather) / length of longest primary feather.

3.4.4 Statistical analyses

We generally report means and standard errors for each trait by species, age, and sex class unless otherwise specified. *Pectoralis* mass, oxidative fiber mass, and glycolytic fiber mass were all analyzes and raw masses and relative masses, corrected for body mass. Relative masses were calculated as the mass of the given trait, divided by the individual's body mass. We tested for differences among species in each trait using one-way ANCOVAs (Type III sums of squares) with class identity (i.e. adult male, subadult male, female) as a covariate. We then performed a Tukey's post-hoc test to determine which pairwise species comparisons differed significantly for each trait. A summary of p-values for these analyses can be found in Table 1. Within each species, we then tested for sex differences in each trait with one-way ANOVAs (Type III sums of squares). For these analyses, adult males and subadult males were analyzed together in respect to females. Analyses of sex differences were conducted for each species independently, and a summary of p-values can be found in Table 2. We then tested for age differences in males for each trait with one-way ANOVAs (Type III sums of squares). Analyses of age differences were conducted for each species independently, and a summary of p-values can be found in Table 3. In all analyses, differences were treated as significant if p-values were < 0.05.

4 Physiological flexibility and seasonal movements across elevational gradients

4.1 Introduction

Mountains exhibit steep environmental gradients over short geographic distances, providing excellent systems to examine physiological adaptation to local environmental pressures (Monge and Leon-Velarde, 1991; Rezende et al., 2005; Projecto-Garcia et al., 2013). Animal habitats in montane regions can extend from sea-level to over 5000 m elevation, spanning atmospheric pressures of 760-420 mmHg and a temperature range over 30°C (West, 1996). At the top of this gradient, endothermic animals face considerable reductions in temperature and partial pressure of oxygen (PO_2) that present challenges for aerobic metabolism and thermoregulation (Rosenmann and Morrison, 1974; Hayes and Chappell, 1986; Ward et al., 1995; Chappell and Hammond, 2004; Dillon et al., 2006). High-elevation adapted animals undergo dramatic compensatory changes in respiratory physiology in response to cold, hypoxic environments. These responses extend to altered protein chemistry (Weber, 2007; Storz et al., 2009; Projecto-Garcia et al., 2013; Zhu et al., 2018), morphology (Kiyamu et al., 2012), and gene regulation (Simonson et al., 2010; Cheviron et al., 2012). A large body of literature focuses on the adaptive mechanisms that facilitate living at high elevations (reviewed in Storz et al., 2010), but this work has largely ignored the preponderance of animals that seasonally migrate in and out of high elevation environments. Because physiological adaptions to high elevation environments are often maladaptive at low elevation, and vice versa (Projecto-Garcia et al., 2013), migrants are presented with distinct metabolic challenges as they shift seasonally between disparate environments.

In birds, which employ flight to access different habitats, two behavioral strategies have evolved that enable breeding at upper elevations in seasonal, temperate mountains: (1) species remain at upper elevations through the winter, or (2) they migrate to lower elevations to escape harsh alpine winter conditions. Species exhibit a continuum of migratory behaviors spanning these two alternative strategies, which in turn impose dramatically disparate environmental selective pressures. By vacating upper elevations in the winter, elevational migrants experience reduced seasonal variation in temperature but increased variation in PO_2. In contrast, species that remain at upper elevations through the winter experience dramatic seasonal variation in temperature but minimal variation in PO_2. A key to understanding the consequences of altitudinal migration or remaining at upper elevations is to study the adaptive physiological mechanisms and trade-offs that are associated with these different strategies.

Physiological flexibility and associated shifts in gene expression mitigates the stress of shifting elevations between seasons, but the mechanisms remain to be tested. Species that experience high levels of environmental variability should evolve high plasticity to cope with the wide range of environmental pressures they experience, as the benefits of occupying multiple environments outweigh the costs of plasticity (Lande, 2014). For example, a species that breeds at high elevation but winters at sea-level experiences dramatic variation in PO_2 through the year. This species should thus show increased flexibility in physiological traits and metabolic pathways that affect O_2 flux through the aerobic respiration pathway to mitigate the stress of shifting elevations. A number of diffusive steps in the cardiorespiratory pathway determine an animal's physiological disposition at a given elevation (Storz et al., 2010), including breathing, O_2 diffusion across the lung/blood barrier, O_2 transport and delivery, O_2 diffusion across the blood/tissue barrier, and tissue-O_2 consumption to generate ATP. Adjustments in physiological

traits and associated shifts in gene expression at different stages along the respiration pathway can alter O_2 flux, energy production, and an organism's metabolic performance (Cheviron et al., 2014). Physiological flexibility along the cardiorespiratory pathway would alleviate harmful effects of seasonality and increase a migratory animal's metabolic performance across an elevational gradient.

In contrast to elevational migrants, animals that breed and winter at high elevations experience little variation in PO_2 throughout the year. If physiological flexibility comes with costs (DeWitt et al., 1998; Lande, 2014), high-elevation residents should specialize on hypoxic, high-elevation environments, and show reduced capacity to cope with high PO_2 (Cheviron et al., 2014). Importantly, mechanisms that increase metabolic performance at high elevation are generally the same as those that increase cold tolerance (Cheviron et al., 2014). Thus, high-elevation specialization on low PO_2 conditions may increase whole-organism metabolism and confer adaptive increases in cold tolerance, which would be critical to year-round high elevation living.

Here, we conducted common garden experiments to ask two related questions: (1) do elevational migrants show increased physiological flexibility in metabolic traits to mitigate the stress of shifting environments, and (2) do species that remain at upper elevations throughout the year show reduced flexibility and increased canalization in metabolic traits, implying a cost of flexibility? To address these questions, we compared physiological flexibility in a pair of closely related species that exhibit differing degrees of elevational movement. We identified the bush-robins, *Tarsiger indicus* and *T. chrysaeus*, as an ideal species pair to test the relationship between physiological flexibility and seasonal movement because they breed in sympatry above 3000 m, but *T. indicus* remains at upper elevations through the winter (above 2000 m), while *T. chrysaeus*

migrates to lower elevations (between 600-2000 m) (Hoyo et al., 2005). We quantified the extent of physiological flexibility in these species through common garden experiments, in which we manipulated temperature and PO_2. We then measured a suite of physiological traits that are known to affect aerobic metabolic function and cold tolerance, including *pectoralis* muscle mass, lung mass, heart mass, and hematological parameters that affect blood-oxygen carrying capacity. We predict that *T. indicus* would exhibit reduced flexibility in traits and a reduced capacity to shift trait values in response to changes in temperature and PO_2. In contrast, because *T. chrysaeus* experiences a wider range of environmental variability in a given year, we predict that *T. chrysaeus* exhibits increased flexibility in metabolic traits in response to shifting environmental pressures.

4.2 Materials and methods

4.2.1 Study animals

The two species in this study, *T. indicus* and *T. chrysaeus* were caught from a wild, sympatric population in Hailuogou Valley, Sichuan, China (29.576096°N, 101.998624°E). We caught all birds between 3000-3400 m elevation as birds moved upslope to breeding grounds in April and early May. Only males in definitive adult plumage were used in this study because we previously found substantial physiological differences within bush-robin species across age and sex classes (chapter 3). We conducted experiments over two years (2015 and 2016) to obtain sufficient sample sizes. Prior to experiments, we kept birds in captivity at 3000 m elevation in individual mesh enclosures for a period between 5 and 35 days. During this period, birds acclimated to captivity under ambient environmental conditions at 3000 m, after which they were transported

to Sichuan University in Chengdu, Sichuan where we conducted common garden experiments. All birds were collected for this study with authorization from Gongga Shan National Nature Reserve and Hailuogou Forestry Bureau. Birds were handled in accordance with the Institutional Animal Care and Use Committee at the University of Chicago (IACUC Protocol #72394).

4.2.2 Common garden experiments

We exposed bush-robins to simulated environmental conditions using a two-by-two experimental design, manipulating temperature and PO_2. Four to six individuals per species were randomly assigned to one of four experimental treatments: (1) warm temperature and normoxia, (2) cold temperature and normoxia, (3) warm temperature and hypoxia, and (4) cold temperature and hypoxia (warm temp. = 27°C (+/- 2°C), cold temp. = 8°C (+/- 2°C), normoxia = 718 mmHg, hypoxia = 439 mmHg). Warm treatments correspond to a general thermoneutral zone in which an adult bird can maintain body temperature with limited energy expenditure. Normoxia and hypoxia treatments correspond to PO_2 at 500 m and 4650 m elevation, respectively. Photoperiod was kept constant to match ambient photoperiod for Hailuogou Valley at the time when experiments were conducted (13.5 light hours/10.5 dark hours). Birds were housed individually in experimental chambers without visual access to other individuals. Chambers were constructed out of circular PVC pipe, stood on end with acrylic floors and lids. Each chamber was approximately 30 cm tall and 40 cm wide. Sample sizes in each treatment for *T. indicus* and *T. chrysaeus*, respectively, were: warm temperature–normoxia = 4 and 5, cold temperature–normoxia = 4 and 5, warm temperature–hypoxia = 6 and 6, and cold temperature–hypoxia = 4 and 5.

Upon arrival to Sichuan University, birds were given a two-day buffer period before experimental treatments began to allow for recovery from any stress induced from transport. After this adjustment period, birds were weighed and placed into their experimental treatments. For the first two days of the experiment, birds in hypoxia treatments were kept at 540 mmHg (the equivalent of 3000 m elevation) to reduce any initial shock and injury to the birds, after which PO_2 was reduced to 439 mmHg, the full hypoxia treatment. Birds acclimated to their full treatment environments for two weeks. Birds in cold and hypoxia treatments were removed from their treatment environments for <15 minutes every two days when we fed birds and cleaned chambers. Birds in warm and normoxia treatments experienced constant treatment conditions during feeding and chamber cleaning because these were the ambient conditions outside of the experimental chambers.

4.2.3 Post-treatment physiological measurements

Birds acclimated to their treatment environment for two weeks, after which we weighed each bird and measured three hematological parameters associated with blood-oxygen carrying capacity: hematocrit (Hct), hemoglobin ([Hb]) concentration, and mean corpuscular hemoglobin concentration (MCHC). Hct (%) is measured as the proportion of red blood cells to total blood volume, and [Hb] is the mass of hemoglobin per deciliter of blood (g/dL). Hct and [Hb] were measured following methods described in DuBay and Witt (2014). MCHC (g/dl) was calculated as the ratio of [Hb] to Hct to test whether more or less hemoglobin is being packed into each red blood cell than expected, irrespective of red blood cell size (Campbell & Ellis, 2007). Each bird was then euthanized and additional physiological measurements were taken.

Immediately after a bird was euthanized, the left *pectoralis* muscle was dissected out and flash frozen in liquid nitrogen. This muscle was stored on liquid nitrogen for less than two weeks before being transferred to a -80°C freezer for long term storage. The right *pectoralis* muscle, left lung, and heart were then dissected out and weighed separately. These masses were taken as wet measurements within minutes of the bird being sacrificed. We calculated total organ mass for both the *pectoralis* muscles and the lungs by multiplying the right *pectoralis* and left lung by two, respectively. For the *pectoralis*, lung, and heart, we analyzed raw mass and mass as a proportion of body mass to control for differences associated with body size. Body mass-corrected values were calculated as the raw mass of the organ divided by body mass at time of collection. All birds in this study are vouchered as specimens at the Nature Museum of Sichuan University.

4.2.4 Statistical analyses

We report means and standard errors for each trait. For each species and trait, we also report, for comparison, the mean values from individuals caught in the wild at 3000 m elevation between March-May in 2014-2016, at the start of the breeding season (data from previous chapter). For all statistical analyses, trait values were log transformed. Within each species, we tested the main effects of temperature and PO_2, and their interaction, on a given trait with a two-way ANOVA or ANCOVA (Type III sums of squares). In final analyses we removed the interaction of temperature and PO_2 from the model because we found no significant effects of this interaction. In addition, we plotted temperature and PO_2 treatments independently to visualize their respective impacts. For Hct, [Hb], and MCHC, we treated body mass (from the end of the treatment) as a covariate. All other traits were analyzed without covariates, except heart mass,

which we analyzed both without covariates and with Hct as a covariate. Between species, we tested for differences in a given trait with a two-way ANOVA, where species identity was treated as an interacting variable with both temperature and PO_2. In all analyses, differences were treated as significant if p-values were < 0.05.

4.3 Results

At the end of the two-week experimental period, we found significant differences within and among species in hematological parameters and heart mass in response to changes in temperature and PO_2. However, we found no significant differences among treatments in body mass, *pectoralis* mass, or lung mass in either species. Body mass was not significantly different among treatments within each species (ANOVA temperature and PO_2, $p > 0.05$ for both species) or between species (ANOVA species, $p > 0.05$), and body mass was similar to the range of body mass that exists for both species in the wild (Figure 29), suggesting that experimental animals were able to maintain body mass in captivity. Raw p*ectoralis* mass and lung mass, and their

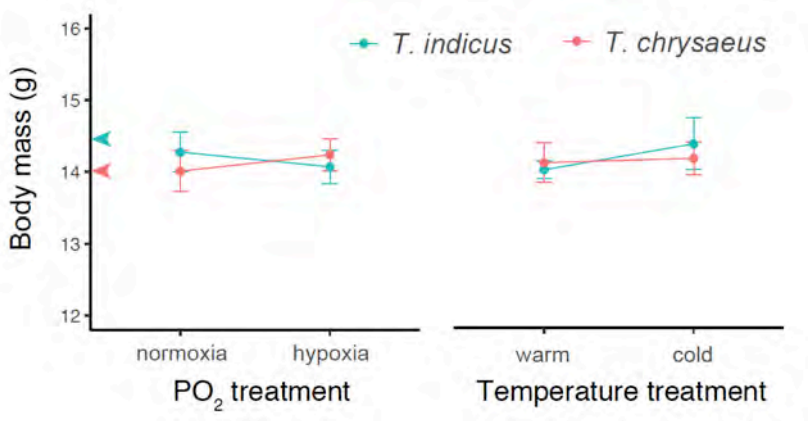

Figure 29 | Body mass after two-week acclimation treatments. Species means (with standard error bars) are plotted. Temperature and PO_2 treatments were plotted separately because we found not significant interactions between these environmental variables. Colored arrows correspond to species means from adult males caught in the wild at 3000 m at the start of the breeding season (data from previous chapter). There are no significant differences in body mass among treatments or between species. N = 18 *T. indicus*, 21 *T. chrysaeus*.

relative trait values (as a proportion of body mass), were also not significantly different within each species (ANOVA temperature and PO_2, $p > 0.05$ for both species) or between species (ANOVA species, $p > 0.05$) (Figures 30a,b and 31a,b). In *T. indicus*, however, raw and relative *pectoralis* masses were substantially lower after the experiment than masses taken from *T. indicus* in the wild (Figures 30a and 31a).

In *T. chrysaeus*, cold-acclimated individuals had significantly larger raw and relative heart masses than warm-acclimated individuals (ANOVA temperature, raw: $p = 0.0096$, relative: $p = 0.0206$) (Figures 30c and 31c). Because high Hct can be associated with maladaptive increases in heart mass through hypertrophy of the right ventricle, we also tested the effects of temperature and PO_2 while controlling for Hct as a covariate. We found that heart mass still increased in cold-acclimated *T. chrysaeus*, irrespective to changes in Hct (ANCOVA temperature, raw: $p = 0.0093$, relative: $p = 0.0232$). We did not find a significant difference in the heart mass of *T. chrysaeus* in response to changes in PO_2 (ANOVA PO_2, raw and relative: $p > 0.05$) (Figures 30c and 31c). Heart mass in *T. indicus* showed a similar qualitative response as *T. chrysaeus* to changes in temperature, but this response was muted and not statistically significant (ANOVA temperature, raw and relative: $p > 0.05$) (Figures 30c and 31c). *T. indicus* did not show a significant difference in heart mass in response to changes in PO_2 (ANOVA PO_2, raw and relative: $p > 0.05$) (Figures 30c and 31c). We found, however, a species effect in raw and relative heart mass in which *T. indicus* had a significantly smaller heart (ANOVA species, raw: $p = 0.0093$; relative: $p = 0.0109$), despite generally having a larger heart in the wild.

We found significant changes in hematological parameters in both species in response to PO_2 but not temperature. Hypoxia-acclimated birds of both species had significantly higher Hct

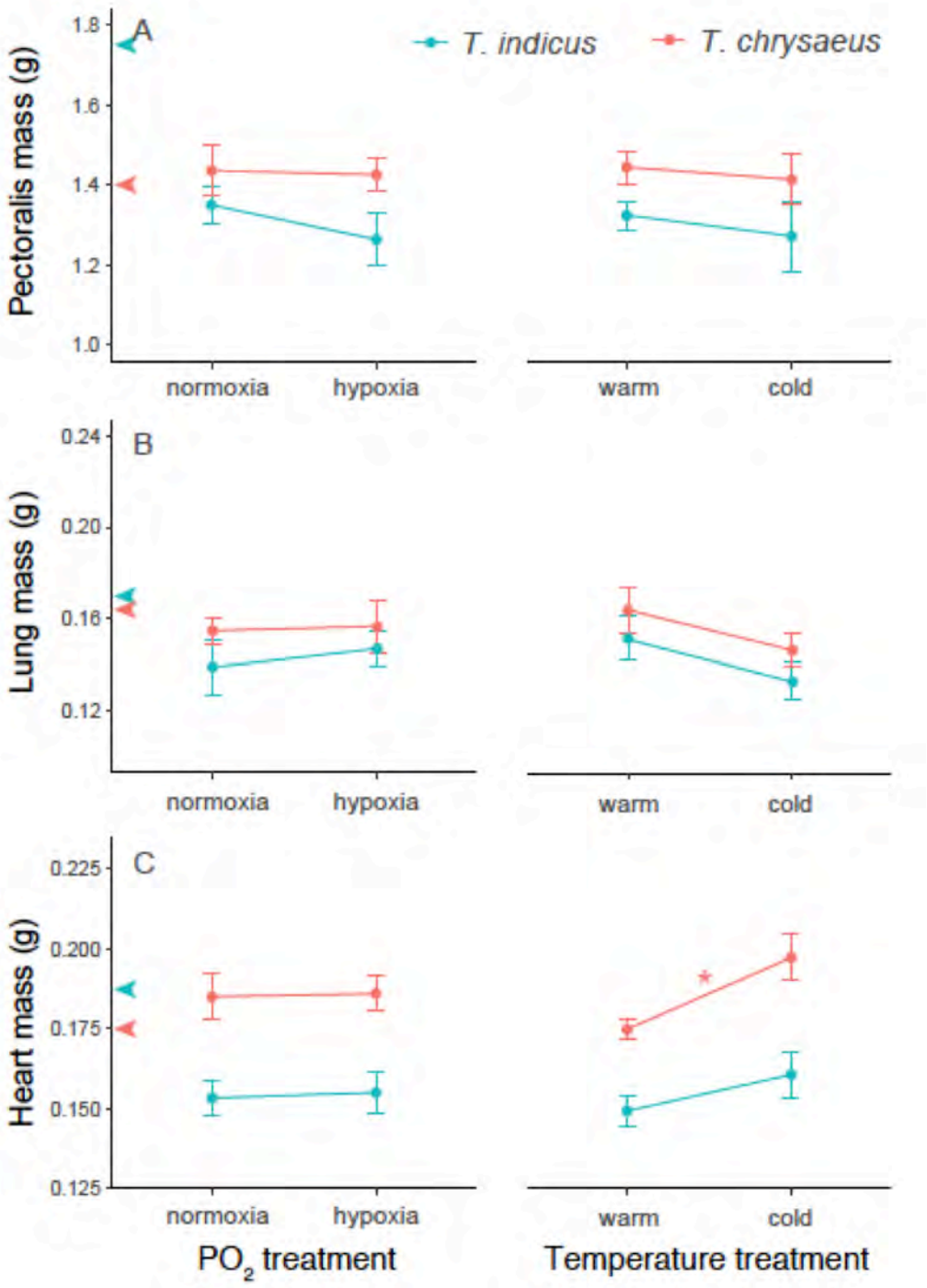

Figure 30 | Raw organ masses after two-week acclimation treatments. Species means (with standard error bars) are plotted. Temperature and PO_2 treatments were plotted separately because we found not significant interactions between these environmental variables. Colored arrows correspond to species means from adult males caught in the wild at 3000 m at the start of the breeding season (data from previous chapter). Asterisks denote statistically significant differences between treatments ($p < 0.05$). N = 18 *T. indicus*, 21 *T. chrysaeus*.

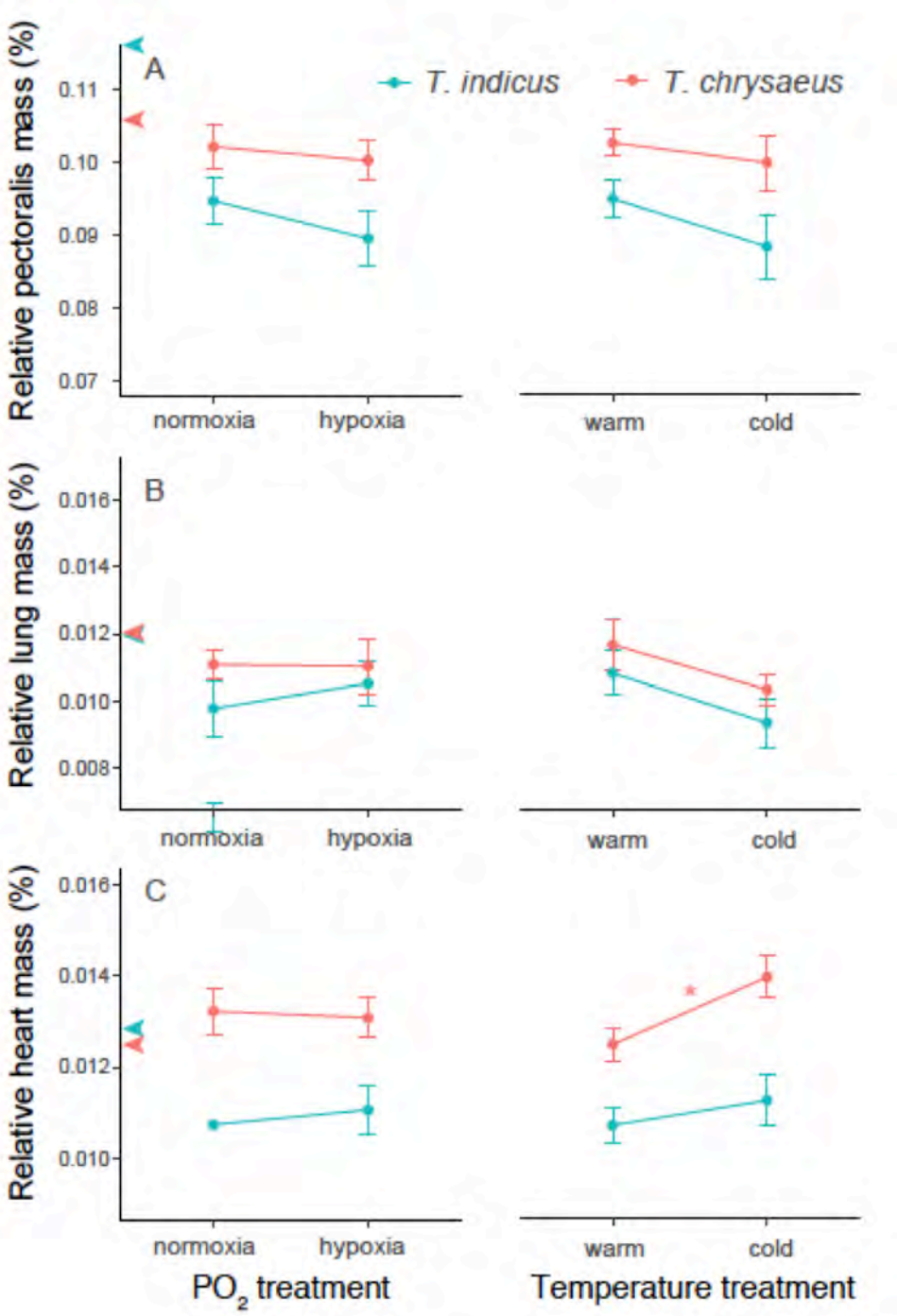

Figure 31 | Relative organ masses after two-week acclimation treatments, calculated as raw organ mass divided by body mass for an individual. Species means (with standard error bars) are plotted. Temperature and PO_2 treatments were plotted separately because we found not significant interactions between these environmental variables. Colored arrows correspond to species means from adult males caught in the wild at 3000 m at the start of the breeding season (data from previous chapter). Asterisks denote statistically significant differences between treatments ($p < 0.05$). N = 18 *T. indicus*, 21 *T. chrysaeus*.

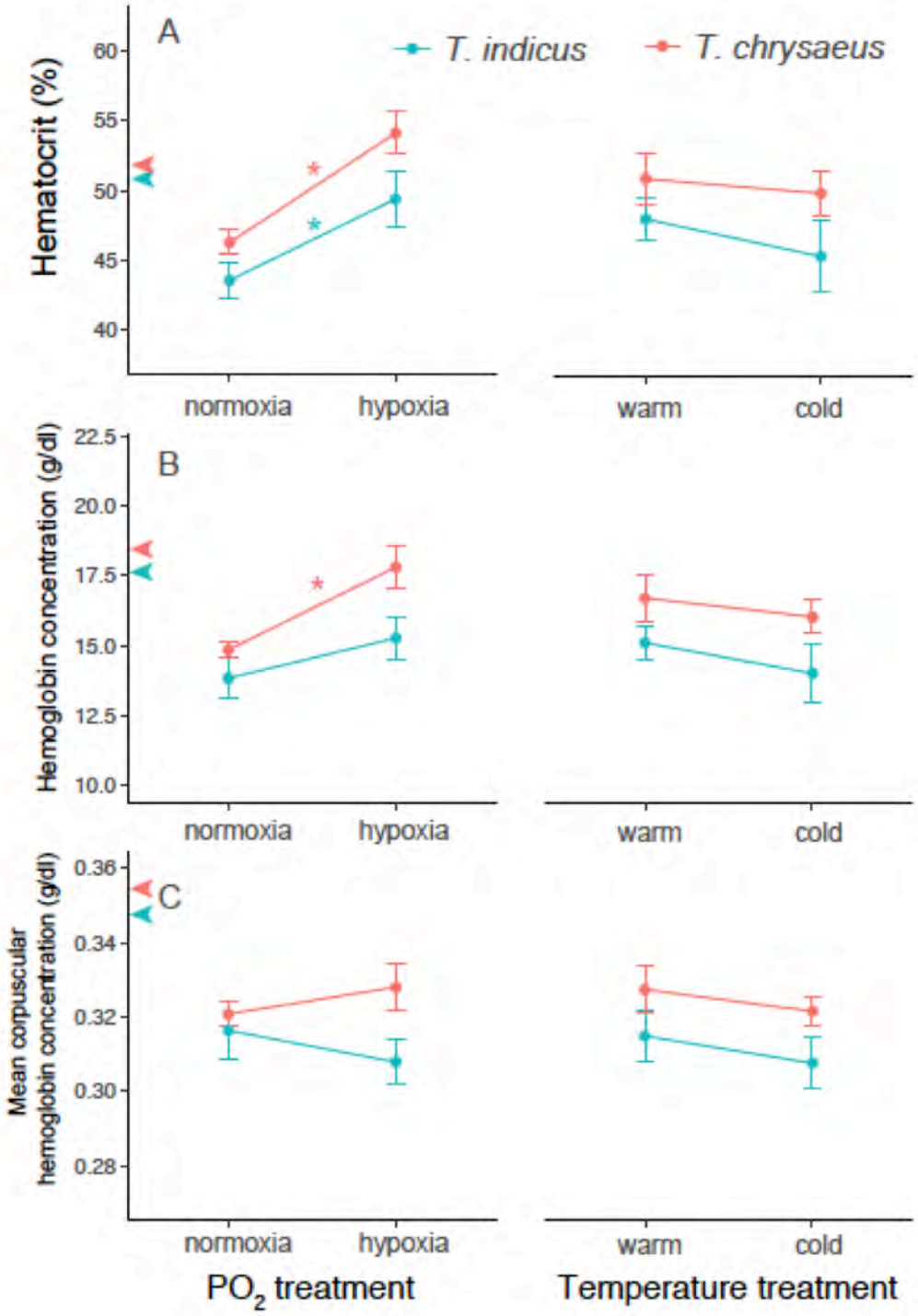

Figure 32 | Blood parameters after two-week acclimation treatments. Species means (with standard error bars) are plotted. Temperature and PO_2 treatments were plotted separately because we found not significant interactions between these environmental variables. Colored arrows correspond to species means from adult males caught in the wild at 3000 m at the start of the breeding season (data from previous chapter). Asterisks denote statistically significant differences between treatments ($p < 0.05$). N = 18 *T. indicus*, 21 *T. chrysaeus*.

than normoxia-acclimated birds (ANCOVA PO_2, *T. indicus*: p = 0.0151, *T. chrysaeus*: p = 0.0006) (Figure 34a). In *T. chrysaeus*, hypoxia-acclimated birds had significantly higher [Hb] than normoxia-acclimated birds (ANCOVA PO_2, p = 0.0018) (Figure 32b). [Hb] in *T. indicus* showed a similar qualitative response as *T. chrysaeus* to changes in PO_2, but this response was muted and not significant (ANCOVA PO_2, p > 0.05) (Figure 32b). We did not find a species effect on Hct or [Hb] (ANCOVA species, p > 0.05), but *T. indicus* had lower Hct and [Hb] in all treatments. A direct comparison between hypoxia-acclimated *T. indicus* and normoxia-acclimated *T. chrysaeus* revealed no significant difference in Hct or [Hb] (t-test, Hct: p = 0.2088, [Hb]: p =0.7698). Neither species showed a significant response in Hct or [Hb] to changes in temperature (ANCOVA temperature, p > 0.05 for both species) (Figure 32a), and we found no significant differences in MCHC among treatments within each species (ANCOVA temperature and PO_2, p > 0.05 for both species) or between species (ANCOVA species, p > 0.05).

4.4 Discussion

Our results support the hypothesis that increased physiological flexibility is associated with higher degrees of seasonal movement, suggesting a mechanism to mitigate the physiological stress of shifting environments. *T. chrysaeus* exhibits increased flexibility in metabolic traits in response to changes in temperature and PO_2, while *T. indicus* showed a more blunted response. The species that remains at higher elevations through the year and experiences reduced annual variation in PO_2 showed qualitatively similar directional changes in metabolic traits in responses to temperature and PO_2 compared to the more extreme migrant. These changes, however, were blunted and often non-significant. This result highlights greater metabolic canalization in *T. indicus*, implying a cost to flexibility. Overall, our results suggest that the capacity for

physiological shifts is optimized to the conditions experienced by a species. Furthermore, these results provide empirical evidence for the role of metabolic flexibility in facilitating seasonal movements in and out of high elevation environments.

T. chrysaeus showed increases in Hct and [Hb] in response to hypoxia, and increased heart mass in response to cold. Both of these responses highlight an increased capacity for this species to shift physiology traits in response to disparate environments, but it is unclear whether or not this flexibility is adaptive. For example, increases in blood parameters under hypoxia are generally thought to be a misdirected response to environmental hypoxia that arose as a response to anemia (Hebbel et al., 1978; Storz, 2010). Low ambient PO_2 and anemia both result in low tissue oxygenation, which triggers red blood cell production. However, when an animal is not anemic, elevated Hct increases blood viscosity beyond normal levels, stressing the heart. As Hct increases, the heart must work harder to pump blood, increasing cardiorespiratory stress. Chronic exposure to hypoxic conditions and the associated increases in Hct and [Hb] can have adverse long-term consequences on the heart and oxygen delivery (Storz et al., 2010). *T. chrysaeus* individuals, however, spend less than half of each year under hypoxic environmental condition at upper elevations. It is possible that *T. chrysaeus* thus benefits from temporary increases in Hct and [Hb] to increase oxygen supply to metabolic tissues, without suffering the long-term consequences of elevated Hct on the heart. This hypothesis could be tested in the future with comparative lifespan data.

The difference in heart mass in *T. chrysaeus* between temperature treatments appears to be an adaptive response to cold stress, rather than a maladaptive response to increasing Hct. When increased heart mass is maladaptive, the right ventricles, which is responsible for pumping deoxygenated blood to the lungs, enlarges to compensate for increases in Hct. Growth of the

heart muscle, known as cardiac hypertrophy, is associated with increased hypertension and risk of heart failure when it occurs in the right ventricle (Voelkel et al., 2006; Julian et al., 2007; Ostadal and Kolar, 2007). Cardiac hypertrophy of the left ventricle, however, appears to be an adaptive response that increases metabolic performance under cold stress. Without weighing the ventricles independently, we cannot be certain which ventricle in *T. chrysaeus* is responsible for the changes in heart mass, but the lack of association between Hct and heart mass among experimental treatments suggests an adaptive increase in left ventricle mass in cold-acclimated birds. Heart mass appears to be responding to changes in temperature, not PO_2 (and Hct), which would be expected if heart mass was increasing through right ventricle hypertrophy. When we accounted for the effects of Hct and [Hb] on heart mass, we found that shifts in temperature still had a significant effect on heart mass.

The reduced flexibility that we observed in *T. indicus*, which spends its entire life above 2000 m, suggests specialization and optimization of traits that have evolved in association with life at upper elevations and restricted elevational movements. Compared to *T. chrysaeus*, *T. indicus* showed a qualitatively similar shift in [Hb] in response to PO_2 and a shift in heart mass in response to temperature, but both responses were blunted in *T. indicus*. These results imply either a cost in flexibility or genetic assimilation, in which species with reduced elevational movements benefit from increased specialization on local environmental conditions. Interestingly, hypoxia-acclimated *T. indicus* exhibited similar levels of Hct and [Hb] to normoxia-acclimated *T. chrysaeus*, which has been observed in other high-elevation adapted birds and mammals (Storz et al., 2010). This result further suggests that *T. indicus* is specialized to high-elevation environments because high-elevation specialists under hypoxia often exhibit trait values similar to lowlanders at sea-level (Storz et al., 2010). Furthermore, heart mass in

general was lower in *T. indicus* than *T. chrysaeus*, despite being generally larger than *T. chrysaeus* in the wild. This result may reflect a response in *T. indicus* to lower overall Hct levels.

We found that the organ masses in *T. indicus* were systematically lower at the end of the experiment than in individuals caught at 3000 m elevation in the wild. In contrast, *T. chrysaeus* showed similar trait values between experimental birds and birds measured in the wild at 3000 m elevation. It is possible that *T. indicus* either could not maintain organ mass or that larger organs were not necessary under experimental treatments and uniformly got smaller. We know from previous work (Chapter 3) that the *pectoralis* muscle in adult male *T. indicus* can be comprised of >25% fast glycolytic fibers, which facilitate short-burst flight capacity. *T.* chrysaeus, on the other hand, has many fewer fast glycolytic fibers. Upon visual inspection of *pectoralis* muscle fiber types in experimental birds, it appears that *T. indicus* had a reduced number of fast glycolytic fibers, which may be a consequence of being held in captive, low use conditions. However, a detailed analysis of muscle fiber composition will be necessary to quantify the reduction of fast glycolytic fibers in captive birds. This analysis will be possible because histology samples for all experimental birds we also collected for future work. It is also worth noting that the low elevation PO_2 environment (i.e. normoxic treatment) extended well beyond PO_2 conditions that *T. indicus* would naturally experience in the wild. It is possible that these conditions hindered the maintenance of organ mass. Alternatively, the cold treatment was not as severe as temperatures regularly experienced by *T. indicus* at upper elevations. It is possible that under these experimental conditions, *T. indicus* no longer needed to maintain muscle mass, and associated organ mass, for thermogenesis.

The differences we observed within and between bush-robin species in response to shifting environmental pressures are highly suggestive of changes in underlying gene expression.

In this study, we did not find a significant interacting effect of temperature and PO_2, despite the fact that both pressures impact whole organism metabolism and oxygen flux through the cardiorespiratory system. It is possible that gene expression analyses will provide increased power to detect potential interactions of temperature and PO_2 on metabolic flexibility. Moving forward, it will be important to incorporate gene regulatory analyses and genomic analyses to determine the role of regulatory flexibility and genomic specialization, respectively, in shaping seasonal elevational distributions and migratory behavior.

References

Aebischer, A., Perrin, N., Krieg, M., Studer, J. and Meyer, D.R., 1996. The role of territory choice, mate choice and arrival date on breeding success in the Savi's Warbler Locustella luscinioides. *Journal of Avian Biology*, pp.143-152.

Askew, G.N., Marsh, R.L. and Ellington, C.P., 2001. The mechanical power output of the flight muscles of blue-breasted quail (Coturnix chinensis) during take-off. *Journal of Experimental Biology*, *204*(21), pp.3601-3619.

Bachmann, J., 2007. Will the circle be unbroken: a history of the US National Ambient Air Quality Standards. *Journal of the Air & Waste Management Association*, *57*(6), pp.652-697.

Beasley, L.E. and Carothers, S.W., 1974. Unusual feeding habits in two species of blackbirds. *The Wilson Bulletin*, pp.478-479.

Beisiegel, B.D.M., 2007. Foraging association between coatis (Nasua nasua) and birds of the Atlantic Forest, Brazil. *Biotropica*, *39*(2), pp.283-285.

Block, B.A., 1994. Thermogenesis in muscle. *Annual review of physiology*, *56*(1), pp.535-577.

Bond, T.C., Bhardwaj, E., Dong, R., Jogani, R., Jung, S., Roden, C., Streets, D.G. and Trautmann, N.M., 2007. Historical emissions of black and organic carbon aerosol from energy-related combustion, 1850–2000. *Global Biogeochemical Cycles*, *21*(2).

Bond, T.C., Doherty, S.J., Fahey, D.W., Forster, P.M., Berntsen, T., DeAngelo, B.J., Flanner, M.G., Ghan, S., Kärcher, B., Koch, D. and Kinne, S., 2013. Bounding the role of black carbon in the climate system: A scientific assessment. *Journal of Geophysical Research: Atmospheres*, *118*(11), pp.5380-5552.

Bond, T.C., Streets, D.G., Yarber, K.F., Nelson, S.M., Woo, J.H. and Klimont, Z., 2004. A technology-based global inventory of black and organic carbon emissions from combustion. *Journal of Geophysical Research: Atmospheres*, *109*(D14).

Buffon, G ,1771–1783. Histoire naturelle des oiseaux. Imprimerie royale, Paris

Campbell, T.W. and Ellis, C.K., 2013. *Avian and exotic animal hematology and cytology*. John Wiley & Sons.

Chappell, M.A. and Hammond, K.A., 2004. Maximal aerobic performance of deer mice in combined cold and exercise challenges. *Journal of Comparative Physiology B*, *174*(1), pp.41-48.

Cheviron, Z.A., Bachman, G.C., Connaty, A.D., McClelland, G.B. and Storz, J.F., 2012. Regulatory changes contribute to the adaptive enhancement of thermogenic capacity in

high-altitude deer mice. *Proceedings of the national academy of sciences*, *109*(22), pp.8635-8640.

Cheviron, Z.A., Connaty, A.D., McClelland, G.B. and Storz, J.F., 2014. Functional genomics of adaptation to hypoxic cold-stress in high-altitude deer mice: transcriptomic plasticity and thermogenic performance. *Evolution*, *68*(1), pp.48-62.

Choi, I.H., Ricklefs, R.E. and Shea, R.E., 1993. Skeletal muscle growth, enzyme activities, and the development of thermogenesis: a comparison between altricial and precocial birds. *Physiological Zoology*, *66*(4), pp.455-473.

Chuang, Y.H., Mazumdar, S., Park, T., Tang, G., Arena, V.C. and Nicolich, M.J., 2011. Generalized linear mixed models in time series studies of air pollution. *Atmospheric Pollution Research*, *2*(4), pp.428-435.

Dakin, R., Segre, P.S., Straw, A.D. and Altshuler, D.L., 2018. Morphology, muscle capacity, skill, and maneuvering ability in hummingbirds. *Science*, *359*(6376), pp.653-657.

Davidson, C.I., 1979. Air pollution in Pittsburgh: A historical perspective. *Journal of the Air Pollution Control Association*, *29*(10), pp.1035-1041.

Dean, W.R.J. and MacDonald, I.A.W., 1981. A review of African birds feeding in association with mammals. *Ostrich*, *52*(3), pp.135-155.

del Hoyo, J., Elliott, A. and Christie, D., 2005. Handbook of the birds of the world. Volume 10. Cuckoo-shrikes to Thrushes. *Lynx Edicions, Barcelona*.

Deveci, D., Marshall, J.M. and Egginton, S., 2001. Relationship between capillary angiogenesis, fiber type, and fiber size in chronic systemic hypoxia. American Journal of Physiology-Heart and Circulatory Physiology, 281(1), pp.H241-H252.

DeWitt, T.J., Sih, A. and Wilson, D.S., 1998. Costs and limits of phenotypic plasticity. *Trends in ecology & evolution*, *13*(2), pp.77-81.

Dial, K.P., Kaplan, S.R., Goslow, G.E. and Jenkins, F.A., 1987. Structure and neural control of the pectoralis in pigeons: implications for flight mechanics. *The Anatomical Record*, *218*(3), pp.284-287.

Dillon, M.E., Frazier, M.R. and Dudley, R., 2006. Into thin air: physiology and evolution of alpine insects. *Integrative and Comparative Biology*, *46*(1), pp.49-61.

DuBay, S.G. and Witt, C.C., 2014. Differential high-altitude adaptation and restricted gene flow across a mid-elevation hybrid zone in Andean tit-tyrant flycatchers. *Molecular ecology*, *23*(14), pp.3551-3565.

Dugatkin, L.A., 2013. *Principles of Animal Behavior: Third International Student Edition*. WW Norton & Company.

Fjeldså, J., Bowie, R.C. and Rahbek, C., 2012. The role of mountain ranges in the diversification of birds. *Annual Review of Ecology, Evolution, and Systematics*, *43*, pp.249-265.

Forslund, P. and Pärt, T., 1995. Age and reproduction in birds—hypotheses and tests. *Trends in Ecology & Evolution*, *10*(9), pp.374-378.

Francis, C.M. and Cooke, F., 1986. Differential timing of spring migration in wood warblers (Parulinae). *The Auk*, pp.548-556.

Grinder, R.D., 1870. The battle for clean air: The smoke problem in post-civil war America. *Pollution and Reform in American Cities*, *1930*, pp.83-103.

Hahn, T.P., Sockman, K.W., Breuner, C.W. and Morton, M.L., 2004. Facultative altitudinal movements by mountain white-crowned sparrows (Zonotrichia leucophrys oriantha) in the Sierra Nevada. *The Auk*, *121*(4), pp.1269-1281.

Hansen, J., Sato, M., Ruedy, R., Lacis, A. and Oinas, V., 2000. Global warming in the twenty-first century: An alternative scenario. *Proceedings of the National Academy of Sciences*, *97*(18), pp.9875-9880.

Hansen, J.E., Sato, M.K.I., Ruedy, R., Nazarenko, L., Lacis, A., Schmidt, G.A., Russell, G., Aleinov, I., Bauer, M., Bauer, S. and Bell, N., 2005. Efficacy of climate forcings. *Journal of Geophysical Research: Atmospheres*, *110*(D18).

Hardy, E , 1937. Polluted wild life. Country Life 81:676.

Harrison, C.J.O., 1963. Industrial'discoloration of house sparrows and other birds. *Br Birds*, *56*, pp.296-297.

Hartshorne, R., 1936. A new map of the manufacturing belt of North America. *Economic Geography*, *12*(1), pp.45-53.

Hawkes, L.A., Balachandran, S., Batbayar, N., Butler, P.J., Frappell, P.B., Milsom, W.K., Tseveenmyadag, N., Newman, S.H., Scott, G.R., Sathiyaselvam, P. and Takekawa, J.Y., 2011. The trans-Himalayan flights of bar-headed geese (Anser indicus). *Proceedings of the National Academy of Sciences*, *108*(23), pp.9516-9519.

Hawkins, G.L., Hill, G.E. and Mercadante, A., 2012. Delayed plumage maturation and delayed reproductive investment in birds. *Biological Reviews*, *87*(2), pp.257-274.

Hayes, J.P. and Chappell, M.A., 1986. Effects of cold acclimation on maximum oxygen consumption during cold exposure and treadmill exercise in deer mice, Peromyscus maniculatus. *Physiological zoology*, *59*(4), pp.473-481.

Hebbel, R.P., Eaton, J.W., Kronenberg, R.S., Zanjani, E.D., Moore, L.G. and Berger, E.M., 1978. Human llamas: adaptation to altitude in subjects with high hemoglobin oxygen affinity. *The Journal of clinical investigation*, *62*(3), pp.593-600.

Heymann, E.W. and Hsia, S.S., 2015. Unlike fellows–a review of primate–non-primate associations. *Biological Reviews*, *90*(1), pp.142-156.

Hickey, J.J. and Anderson, D.W., 1968. Chlorinated hydrocarbons and eggshell changes in raptorial and fish-eating birds. *Science*, *162*(3850), pp.271-273.

Ito, A. and Penner, J.E., 2005. Historical emissions of carbonaceous aerosols from biomass and fossil fuel burning for the period 1870–2000. *Global Biogeochemical Cycles*, *19*(2).

Ives, J.E., Britten, R., Armstrong, D.W., Gill, W.A. and Goldman, F.H., 1936. Atmospheric Pollution of American Cities for the Years 1931 to 1933 with Special Reference to the Solid Constituents of the Pollution. *Atmospheric Pollution of American Cities for the Years 1931 to 1933 with Special Reference to the Solid Constituents of the Pollution.*

Jacobson, M.Z., 2000. A physically-based treatment of elemental carbon optics: Implications for global direct forcing of aerosols. *Geophysical Research Letters*, *27*(2), pp.217-220.

Jacobson, M.Z., 2004. Climate response of fossil fuel and biofuel soot, accounting for soot's feedback to snow and sea ice albedo and emissivity. *Journal of Geophysical Research: Atmospheres*, *109*(D21).

Johnson, K.G., Schaller, G.B. and Jinchu, H., 1988. Comparative behavior of red and giant pandas in the Wolong Reserve, China. *Journal of mammalogy*, *69*(3), pp.552-564.

Julian, R.J., McMillan, I. and Quinton, M., 1989. The effect of cold and dietary energy on right ventricular hypertrophy, right ventricular failure and ascites in meat-type chickens. *Avian Pathology*, *18*(4), pp.675-684.

Ketterson, E.D. and Nolan, V., 1976. Geographic Variation and Its Climatic Correlates in the Sex Ratio of Eastern-Wintering Dark-Eyed Juncos (Junco Hyemalis Hyemalis). *Ecology*, *57*(4), pp.679-693.

Ketterson, E.D. and Nolan, V.A.L., 1983. The evolution of differential bird migration. In *Current ornithology* (pp. 357-402). Springer, Boston, MA.

King, A.J. and Cowlishaw, G., 2009. Foraging opportunities drive interspecific associations between rock kestrels and desert baboons. *Journal of Zoology*, *277*(2), pp.111-118.

Kiyamu, M., Bigham, A., Parra, E., León-Velarde, F., Rivera-Chira, M. and Brutsaert, T.D., 2012. Developmental and genetic components explain enhanced pulmonary volumes of

female Peruvian Quechua. *American journal of physical anthropology*, *148*(4), pp.534-542.

Koch, D., Schulz, M., Kinne, S., McNaughton, C., Spackman, J.R., Balkanski, Y., Bauer, S., Berntsen, T., Bond, T.C., Boucher, O. and Chin, M., 2009. Evaluation of black carbon estimations in global aerosol models. *Atmospheric Chemistry and Physics*, *9*(22), pp.9001-9026.

Kuniy, A.A., de Morais, M.M. and Gomes, E.P.C., 2003. Association between olivaceous woodcreeper (Sittasomus griseicapillus) and golden lion tamarin (Leontopithecus rosalia) at Uniao Biological Reserve, Rio das Ostras, Brazil. *Acta Biologica Leopoldensia*, *25*, pp.261-264.

Lack, D., 1948. Notes on the ecology of the robin. *Ibis*, *90*(2), pp.252-279.

Lande, R., 2014. Evolution of phenotypic plasticity and environmental tolerance of a labile quantitative character in a fluctuating environment. *Journal of evolutionary biology*, *27*(5), pp.866-875.

Leon-Velarde, F., Sanchez, J., Bigard, A.X., Brunet, A., Lesty, C. and Monge-C, C., 1993. High altitude tissue adaptation in Andean coots: capillarity, fibre area, fibre type and enzymatic activities of skeletal muscle. *Journal of Comparative Physiology B*, *163*(1), pp.52-58.

Lockwood, R., Swaddle, J.P. and Rayner, J.M., 1998. Avian wingtip shape reconsidered: wingtip shape indices and morphological adaptations to migration. *Journal of Avian Biology*, pp.273-292.

Loucks, C.J., Lü, Z., Dinerstein, E., Wang, H., Olson, D.M., Zhu, C. and Wang, D., 2001. Giant pandas in a changing landscape.

Lozano, G.A., Perreault, S. and Lemon, R.E., 1996. Age, arrival date and reproductive success of male American redstarts Setophaga ruticilla. *Journal of Avian Biology*, pp.164-170.

Lundgren, B.O. and Kiessling, K.H., 1988. Comparative aspects of fibre types, areas, and capillary supply in the pectoralis muscle of some passerine birds with differing migratory behaviour. *Journal of Comparative Physiology B*, *158*(2), pp.165-173.

Lyon, B.E. and Montgomerie, R.D., 1986. Delayed plumage maturation in passerine birds: Reliable signaling by subordinate males?. *Evolution*, *40*(3), pp.605-615.

Marsh, R.L. and Dawson, W.R., 1989. Avian adjustments to cold. In *Animal Adaptation to Cold* (pp. 205-253). Springer, Berlin, Heidelberg.

Mayr, E., 1939. The sex ratio in wild birds. *The American Naturalist*, *73*(745), pp.156-179.

McConnell, J.R., Edwards, R., Kok, G.L., Flanner, M.G., Zender, C.S., Saltzman, E.S., Banta, J.R., Pasteris, D.R., Carter, M.M. and Kahl, J.D., 2007. 20th-century industrial black carbon emissions altered arctic climate forcing. *Science*, *317*(5843), pp.1381-1384.

McKay, B.D., 2013. The use of digital photography in systematics. *Biological Journal of the Linnean Society*, *110*(1), pp.1-13.

Meyers, R.A. and Stakebake, E.F., 2005. Anatomy and histochemistry of spread-wing posture in birds. 3. Immunohistochemistry of flight muscles and the "shoulder lock" in albatrosses. *Journal of Morphology*, *263*(1), pp.12-29.

Monge, C. and Leon-Velarde, F., 1991. Physiological adaptation to high altitude: oxygen transport in mammals and birds. *Physiological Reviews*, *71*(4), pp.1135-1172.

Morbey, Y.E. and Ydenberg, R.C., 2001. Protandrous arrival timing to breeding areas: a review. *Ecology letters*, *4*(6), pp.663-673.

Morimoto, G., Yamaguchi, N. and Ueda, K., 2006. Plumage color as a status signal in male–male interaction in the red-flanked bushrobin, Tarsiger cyanurus. *Journal of ethology*, *24*(3), pp.261-266.

Myers, J.P., 1981. A test of three hypotheses for latitudinal segregation of the sexes in wintering birds. *Canadian Journal of Zoology*, *59*(8), pp.1527-1534.

Novakov, T., Ramanathan, V., Hansen, J.E., Kirchstetter, T.W., Sato, M., Sinton, J.E. and Sathaye, J.A., 2003. Large historical changes of fossil-fuel black carbon aerosols. *Geophysical Research Letters*, *30*(6).

O'Neill, J.P. and Parker, T.A., 1978. Responses of birds to a snowstorm in the Andes of southern Peru. *The Wilson Bulletin*, *90*(3), pp.446-449.

Ostadal, B. and Kolar, F., 2007. Cardiac adaptation to chronic high-altitude hypoxia: beneficial and adverse effects. *Respiratory physiology & neurobiology*, *158*(2-3), pp.224-236.

Peng, R.D., Dominici, F. and Louis, T.A., 2006. Model choice in time series studies of air pollution and mortality. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *169*(2), pp.179-203.

Penner, J.E. and Novakov, T., 1996. Carbonaceous particles in the atmosphere: A historical perspective to the Fifth International Conference on Carbonaceous Particles in the Atmosphere. *Journal of Geophysical Research: Atmospheres*, *101*(D14), pp.19373-19378.

Perrins, C.M., 1970. The timing of birds 'breeding seasons. *Ibis*, *112*(2), pp.242-255.

Peter, J.B., Barnard, R.J., Edgerton, V.R., Gillespie, C.A. and Stempel, K.E., 1972. Metabolic profiles of three fiber types of skeletal muscle in guinea pigs and rabbits. *Biochemistry*, *11*(14), pp.2627-2633.

Petit, M. and Vézina, F., 2014. Phenotype manipulations confirm the role of pectoral muscles and haematocrit in avian maximal thermogenic capacity. *Journal of Experimental Biology*, *217*(6), pp.824-830.

Platt, H.L., 1991. *The Electric City: Energy and the Growth of the Chicago Area, 1880-1930*. University of Chicago Press.

Projecto-Garcia, J., Natarajan, C., Moriyama, H., Weber, R.E., Fago, A., Cheviron, Z.A., Dudley, R., McGuire, J.A., Witt, C.C. and Storz, J.F., 2013. Repeated elevational transitions in hemoglobin function during the evolution of Andean hummingbirds. *Proceedings of the National Academy of Sciences*, *110*(51), pp.20669-20674.

Pyle, P, 1997. Identification Guide to North American Birds: Columbidae to Ploceidae (Slate Creek Press, Point Reyes Station, CA).

Ramanathan, V. and Carmichael, G., 2008. Global and regional climate changes due to black carbon. *Nature geoscience*, *1*(4), p.221.

Rand, A.L., 1953. Factors affecting feeding rates of anis. *The Auk*, *70*(1), pp.26-30.

Rezende, E.L., Gomes, F.R., Ghalambor, C.K., Russell, G.A. and Chappell, M.A., 2005. An evolutionary frame of work to study physiological adaptation to high altitudes. *Revista Chilena de Historia Natural*, *78*(2).

Ricklefs, R.E. and Wikelski, M., 2002. The physiology/life-history nexus. *Trends in Ecology & Evolution*, *17*(10), pp.462-468.

Rosenmann, M. and Morrison, P., 1974. Maximum oxygen consumption and heat loss facilitation in small homeotherms by He-O2. *American Journal of Physiology-Legacy Content*, *226*(3), pp.490-495.

Rosser, B.W. and George, J.C., 1986. The avian pectoralis: histochemical characterization and distribution of muscle fiber types. *Canadian Journal of Zoology*, *64*(5), pp.1174-1185.

Rosser, B.W., Waldbillig, D.M., Wick, M. and Bandman, E., 1994. Muscle fiber types in the pectoralis of the white pelican, a soaring bird. *Acta Zoologica*, *75*(4), pp.329-336.

Rohwer, S., Fretwell, S.D. and Niles, D.M., 1980. Delayed maturation in passerine plumages and the deceptive acquisition of resources. *The American Naturalist*, *115*(3), pp.400-437.

Sangster, G., Alström, P., Forsmark, E. and Olsson, U., 2010. Multi-locus phylogenetic analysis of Old World chats and flycatchers reveals extensive paraphyly at family, subfamily and genus level (Aves: Muscicapidae). *Molecular Phylogenetics and Evolution*, *57*(1), pp.380-392.

Satio, D.S., Morimoto, G., Fukunaga, A. and Ueda, K., 2006. Isolation and characterization of microsatellite markers in red-flanked bushrobin, Tarsiger cyanurus (Aves: Turdidae). *Molecular Ecology Resources*, *6*(2), pp.425-427.

Schwabl, H., 1992. Winter and breeding territorial behaviour and levels of reproductive hormones of migratory European robins. *Ornis Scandinavica*, pp.271-276.

Scott, G.R., Egginton, S., Richards, J.G. and Milsom, W.K., 2009. Evolution of muscle phenotype for extreme high altitude flight in the bar-headed goose. *Proceedings of the Royal Society of London B: Biological Sciences*, p.rspb20090947.

Segre, P.S., Dakin, R., Zordan, V.B., Dickinson, M.H., Straw, A.D. and Altshuler, D.L., 2015. Burst muscle performance predicts the speed, acceleration, and turning performance of Anna's hummingbirds. *ELife*, *4*.

Sheldon, W. G., 1937. Notes on the Giant Panda. Journal of Mammalogy 18:13–19.

Shindell, D. and Faluvegi, G., 2009. Climate response to regional radiative forcing during the twentieth century. *Nature Geoscience*, *2*(4), p.294.

Shindell, D., Kuylenstierna, J.C., Vignati, E., van Dingenen, R., Amann, M., Klimont, Z., Anenberg, S.C., Muller, N., Janssens-Maenhout, G., Raes, F. and Schwartz, J., 2012. Simultaneously mitigating near-term climate change and improving human health and food security. *Science*, *335*(6065), pp.183-189.

Simonson, T.S., Yang, Y., Huff, C.D., Yun, H., Qin, G., Witherspoon, D.J., Bai, Z., Lorenzo, F.R., Xing, J., Jorde, L.B. and Prchal, J.T., 2010. Genetic evidence for high-altitude adaptation in Tibet. *Science*, *329*(5987), pp.72-75.

Stern, A.C. and Professor, E., 1982. History of air pollution legislation in the United States. *Journal of the Air Pollution Control Association*, *32*(1), pp.44-61.

Stevens, M., PARraga, C.A., Cuthill, I.C., Partridge, J.C. and Troscianko, T.S., 2007. Using digital photography to study animal coloration. *Biological Journal of the Linnean Society*, *90*(2), pp.211-237.

Storz, J.F., 2010. Genes for high altitudes. *Science*, *329*(5987), pp.40-41.

Storz, J.F., Runck, A.M., Sabatino, S.J., Kelly, J.K., Ferrand, N., Moriyama, H., Weber, R.E. and Fago, A., 2009. Evolutionary and functional insights into the mechanism underlying

high-altitude adaptation of deer mouse hemoglobin. *Proceedings of the National Academy of Sciences*, *106*(34), pp.14450-14455.

Storz, J.F., Scott, G.R. and Cheviron, Z.A., 2010. Phenotypic plasticity and genetic adaptation to high-altitude hypoxia in vertebrates. *Journal of Experimental Biology*, *213*(24), pp.4125-4136.

Stradling, D. and Thorsheim, P., 1999. The smoke of great cities: British and American efforts to control air pollution, 1860-1914. *Environmental History*, *4*(1), pp.6-31.

Swanson, D.L., Zhang, Y. and King, M.O., 2013. Individual variation in thermogenic capacity is correlated with flight muscle size but not cellular metabolic capacity in American goldfinches (Spinus tristis). *Physiological and Biochemical Zoology*, *86*(4), pp.421-431.

Tarr, J.A., 1981. Changing fuel use behavior: The Pittsburgh smoke control movement, 1940–1950. *Technological Forecasting and Social Change*, *20*(4), pp.331-346.

Tarr, J.A. and Zimring, C., 1997. The struggle for smoke control in St. Louis. *Common fields: an environmental history of St. Louis. Missouri Historical Society Press, St. Louis*, pp.199-220.

Thompson, D.R., Furness, R.W. and Monteiro, L.R., 1998. Seabirds as biomonitors of mercury inputs to epipelagic and mesopelagic marine food chains. *Science of the Total Environment*, *213*(1-3), pp.299-305.

Voelkel, N.F., Quaife, R.A., Leinwand, L.A., Barst, R.J., McGoon, M.D., Meldrum, D.R., Dupuis, J., Long, C.S., Rubin, L.J., Smart, F.W. and Suzuki, Y.J., 2006. Right ventricular function and failure: report of a National Heart, Lung, and Blood Institute working group on cellular and molecular mechanisms of right heart failure. *Circulation*, *114*(17), pp.1883-1891.

Walther, G.R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J., Fromentin, J.M., Hoegh-Guldberg, O. and Bairlein, F., 2002. Ecological responses to recent climate change. *Nature*, *416*(6879), p.389.

Ward, M.P., Milledge, J.S. and West, J.B., 1995. Acute and subacute mountain sickness. *High Altitude Medicine and Physiology*, pp.366-387.

Weber, R.E., 2007. High-altitude adaptations in vertebrate hemoglobins. *Respiratory physiology & neurobiology*, *158*(2-3), pp.132-142.

Welch Jr, K.C. and Altshuler, D.L., 2009. Fiber type homogeneity of the flight musculature in small birds. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, *152*(4), pp.324-331.

West, J.B., 1996. Prediction of barometric pressures at high altitudes with the use of model atmospheres. *Journal of Applied Physiology*, *81*(4), pp.1850-1854.

Westland, S., Ripamonti, C. and Cheung, V., 2012. *Computational colour science using MATLAB*. John Wiley & Sons.

Wingfield, J.C., Ball, G.F., Dufty, A.M., Hegner, R.E. and Ramenofsky, M., 1987. Testosterone and aggression in birds. *American Scientist*, *75*(6), pp.602-608.

Wood, S.N., 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *73*(1), pp.3-36.

Wu, Y., DuBay, S.G., Colwell, R.K., Ran, J. and Lei, F., 2017. Mobile hotspots and refugia of avian diversity in the mountains of south-west China under past and contemporary global climate change. *Journal of biogeography*, *44*(3), pp.615-626.

Zhang, S. and Wang, L., 2000. Following of brown capuchin monkeys by white hawks in French Guiana. *The Condor*, *102*(1), pp.198-201.

Zhou H.M., Li Y.D., Jiang Y., Yuan Y.L., Zhou Z.R., Yang C.M., Deng M.J., Shao F.L., Wu M., Wang S.P., Li Y., Wei C.Y., Qin H., Zhang H., Huang F., Yu G.B., Li Z.M., He L.Y., Yang Y., Chen X.W., Zhao C.Q., Shi L., Li L.L., Li L., 2014. Birds of Gonggashan Reserve. University of Electronic Sciences and Technology of China Press, Chengdu.

Zhu, X., Guan, Y., Signore, A.V., Natarajan, C., DuBay, S.G., Cheng, Y., Han, N., Song, G., Qu, Y., Moriyama, H. and Hoffmann, F.G., 2018. Divergent and parallel routes of biochemical adaptation in high-altitude passerine birds from the Qinghai-Tibet Plateau. *Proceedings of the National Academy of Sciences*, *115*(8), pp.1865-1870.