Supporting Information

Elliott et al. 10.1073/pnas.1304838110

SI Materials and Methods

Our observations were made at the Coats Island west colony (62°57′N, 82°00′W), Nunavut, Canada, July 15 to August 10, 2006 ($n = 40$). Parental thick-billed murres (*Uria lomvia*) were caught at their nests with a noose pole. All birds were placed in a cloth bag and weighed using a Pesola spring balance (±10 g) during each capture period. Handling time, including bleeding time, was always less than 10 min and usually less than 5 min. In addition to doubly labeled water experiments, during each experimentation period 5–10 birds were simultaneously equipped with time/depth/temperature recorders (TDRs), but not injected or blood-sampled to examine the effect of these activities on murre time budgets.

TDR Observations. We attached Lotek LTD 1100 (Lotek Marine Technologies, Inc.) TDRs to plastic bands attached to the legs of murres. The TDRs were cylindrical (mass = 4.5 g; diameter = 1 cm; length = 3.3 cm; sampling interval = 3 s; depth precision = ±0.1 m; accuracy ~ ±2 m; 128-kB memory, recording continuously for 55 h). Back-mounted TDRs are known to impact murre provisioning rates, number of foraging trips, adult attendance, mass loss, and dive depth and duration. Though heavier, our leg-mounted TDRs had no impact on provisioning rates, trip duration, or mass loss (1–9). We used the pressure log from the TDR to determine time spent underwater, and the temperature log from the TDR to determine time in air, on land, and on water (7–9). In 2009, as part of a separate project, we attached TDRs and accelerometers to the same individuals. Activity budgets obtained from the two different instruments agreed within 3% for each activity.

Doubly Labeled Water. During the incubation and chick-rearing periods (July 18 to August 10, 2006), we injected intraperitoneally either 0.5 mL or 1.0 mL of doubly labeled water (50% $H_2{}^{18}O$ and 25% D_2O; see data table deposited in the Dryad database for exact enrichment) ~1-cm deep into the brood patch of 51 parental murres using a 27-gauge needle and a 1-mL syringe. Eight of the birds were weighed and killed after 240 min. The entire body was then freeze-dried to determine the body water content. Using those data, we were able to directly measure percent body water and show that it closely corresponded with the plateau method using values obtained 90–180 min after injection (10). The remaining birds were released following injection, and immediately returned to their breeding site. We failed to recapture 13 of the 51 birds within 90–180 min of injection. Equilibrium isotope values for these 13 birds were found using a regression between ^{18}O/deuterium and body mass for the 38 birds yielding plateau (initial) blood samples ($R^2 = 0.98$).

Initial blood samples were taken from the tarsal vein. Final blood samples were obtained 48 h ($n = 39$), 72 h ($n = 7$, including six birds sampled at 48 h), 96 h ($n = 5$, including three birds sampled at 48 h), and 120 h ($n = 1$) after the initial blood sample, and were taken from the brachial vein. Background blood samples were obtained from the tarsal vein in a separate group of 14 birds (seven on July 14 and seven on August 5) to avoid entering the brachial vein multiple times. The average background concentrations were 1993.4 ± 0.5 ppm of ^{18}O and 152.9 ± 0.5 ppm of deuterium, with no difference between dates ($P > 0.6$). All blood samples were collected into one, two, or three 60-μL capillary tubes that were flame-sealed and sent to the University of Aberdeen for laboratory processing after the field season. Only samples with final measurements above 20 ppm of ^{18}O were used (this cutoff is 10 SD above background; deuterium depleted at a slower rate and was consequently always >10 SD above background during final measurements when ^{18}O was >10 SD above background) (11). Final blood samples (F1 samples) >10 SD above background were obtained from 41 birds (of 43 birds that were injected—one bird was recaptured but not blood-sampled, and another was not recaptured), including 1 bird at 24 h, 38 at 48 h, 1 at 72 h, and 1 at 96 h. Second final blood samples (F2 samples) >10 SD above background were obtained from 10 birds, including 1 at 48 h, 5 at 72 h, and 4 at 96 h. For average estimates of murre energy consumption across the season or within breeding periods, single values for each individual (either F1 values or F2 values, where available) were averaged across all individuals.

We used a value of 0.809 for the respiratory quotient (25 kJ/L CO_2) based on a diet of 85% protein, 10% lipids, and 5% carbohydrates, as directly measured via nestling and adult dietary sampling (12). In contrast, the respiratory quotient for 11 thick-billed murres fasting for 4 h at the colony (i.e., using primarily lipids) measured using open-flow respirometry was 0.70 ± 0.05). Thus, it is unlikely that we underestimated the average respiratory quotient across activities, and therefore it is unlikely that daily energy expenditure was overestimated.

Activity-Specific Metabolic Rate. Complete activity records (i.e., birds recaptured before TDRs stopped recording) were obtained for 38 F1 records and four F2 records. Three of those birds provided both F1 and F2 measurements. Thus, our sample size for independent measurements of activity and metabolic rate was 39. We completed a second set of analyses ($n = 42$) that included the second set of F2 measurements as independent samples, assuming that most of the variation in metabolic rate is due to variation in activity rather than individual variation in activity-specific metabolic rates. Because we were interested in activity-specific metabolic rates (and had measured activity independently), we assumed that activity-specific metabolic rates did not vary significantly with time of day (i.e., we assumed that diel variation in energy expenditure was related primarily to diel variation in activity) and did not adjust values that were collected up to ±3 h from 24-h cycles.

To calculate activity-specific metabolic rates, we modeled total energy expenditure by the equation

$$EE = DMR \times T_d + FlMR \times T_f + RMR_w \times T_w + RMR_a \times T_a, \quad \textbf{[S1]}$$

where EE is energy expenditure during the sampling period (measured by doubly labeled water), DMR is diving metabolic rate, FlMR is flying metabolic rate, RMR_w is resting metabolic rate on the water, RMR_a is resting metabolic rate on land, T_d is time spent diving, T_f is time spent flying, T_w is time spent on the water, and T_a is time spent at the colony (the latter four measurements derived from TDR records). The metabolic rates for each activity were determined using a multiple linear regression to provide DMR, FlMR, RMR_w, and RMR_a. The values are therefore averages within each activity, and each includes relatively energy-intensive periods. For example, preening on the water would be included in RMR_w, preening at the colony in RMR_a, and active prey chasing in DMR. We also considered three other variations on the basic model, because dive costs are known to be nonlinearly related to dive duration and depth. By including all activities within the model and forcing an intercept of zero, we avoided underestimating activity-specific metabolic rates, which can occur when examining the slope of energy expenditure against each activity separately (13).

First, we considered a model where dive costs were proportional to the mechanical costs associated with overcoming buoyancy (14) rather than total dive time. To model those costs we used the equation

$$\mathrm{EE} = \eta \times \sum \mathrm{Ln}(\text{dive depth}) + C + \mathrm{FlMR} \times \mathrm{T_f} + \mathrm{RMR_w} \times \mathrm{T_w} + \mathrm{RMR_a} \times \mathrm{T_a}, \quad \textbf{[S2]}$$

where η is the efficiency of converting metabolic energy into mechanical energy (estimated as a coefficient in the multiple regression) and Σ Ln(dive depth) is the summation taken over all dives for the relevant period and represents mechanical buoyancy-related costs. Work done to overcome buoyancy is proportional to the expression Ln(dive depth) (14) up to a constant *C*, which was obtained assuming a surface buoyancy of 0.518 L (total buoyancy including air) as measured previously for thick-billed murres (15). We considered equations with and without the constant; we allowed the equation's intercept to vary to allow for the constant *C*.

Second, we considered a model where dive costs were proportional to the mechanical costs measured by accelerometers, and including drag, inertial, and buoyancy (5, 15) rather than total dive time. To construct that model, we used the equation

$$\mathrm{EE} = \eta \times \sum \text{mechanical costs} + \mathrm{FlMR} \times \mathrm{T_f} + \mathrm{RMR_w} \times \mathrm{T_w} + \mathrm{RMR_a} \times \mathrm{T_a}, \quad \textbf{[S3]}$$

where η is the efficiency of converting metabolic energy into mechanical energy (estimated as a coefficient in the multiple regression) and Σ mechanical costs is the summation taken over all dives for the relevant period for the predicted mechanical costs associated with diving to the depth of each dive. We estimated mechanical costs in 2009 by attaching accelerometers to 10 murres. Mechanical costs were estimated as the summation of buoyancy, inertial, and drag forces over each wing stroke (see ref. 15 for details of calculation and estimation of drag forces). The depth-specific value for the mechanical component of diving matched that obtained from the literature (5, 15). The model assumed that the mechanical costs for diving were primarily associated with descent.

Third, we considered a model where dive costs were determined by rate of oxygen use in the air sacs, with oxygen consumption rate declining exponentially through the dive because of shunting of blood away from nonessential organs. We used the functional relationship (an exponential model) described for penguins (16) and modeled the relationship using the equation

$$\mathrm{EE} = \eta \times \sum \left(1\text{-}e^{(-t/1.23)}\right) + \mathrm{FlMR} \times \mathrm{T_f} + \mathrm{RMR_w} \times \mathrm{T_w} + \mathrm{RMR_a} \times \mathrm{T_a}, \quad \textbf{[S4]}$$

where t is dive duration and η is the conversion factor from oxygen consumption rate to watts (energy expenditure rate); note that the equation keeps the form of oxygen utilization during the dive described by Knower Stockard (16), but the coefficient representing the absolute rate of oxygen consumption is subsumed into η. The summation was taken over all dives for the relevant period. The functional relationship assumed that the oxygen depletion rate followed the form described for deep-diving emperor penguins (*Aptenodytes forsteri*), the only deep-diving species where the relationship has been directly measured (figure 10 in ref. 16). We altered the time constant for emperor penguins from 1/2.91 (figure 10 in ref. 16) to 1/1.23 because the maximum dive duration for murres is ∼4.6 min (6), compared with 11 min for emperor penguins (16). Thus, oxygen stores in murres are used up approximately at 11/4.6 = 2.39× the rate of emperor penguins, and we adjusted the time constant to represent that change in rate (2.39/2.91 = 1/1.23). Because we were able to model directly the nonlinear change in oxygen consumption during the dive based on ref. 16 measurements, we did not need to combine oxygen consumption rates over the course of the entire bout (17).

We compared the effectiveness of different models using Akaike's information criterion (AIC), which penalizes models that increase the number of parameters without improvement in fit. Fig. S2 shows the expected shape of each of the four different models for diving metabolic rate.

To compare measured flight costs with mechanical flight costs predicted from aerodynamic theory, we calculated the mechanical flight costs for thick-billed murres using the program Flight.EXE 1.24 (18). We used the values 1 kg for body mass, 0.71 m for wingspan, and 0.051 m^2 for wing area. Results are shown in Fig. S3.

Weather and Controlling for the Effect of Temperature. Wind speed and ambient temperature were measured at sea level at the colony each day at 0800 and 2000 hours during the doubly labeled water experiment. Wind speed and temperature were also measured hourly at the Environment Canada weather station at the Coral Harbor airport, 100 km to the northwest. The foraging range of murres from the Coats Island colony encompasses the region between the colony and Coral Harbor, so these measurements bounded the foraging area used by murres in this study (19). We used the temperature logs from the TDRs to determine average temperature during flight and diving. We averaged both wind speed and temperature across each deployment and examined the regression of wind speed and temperature on energy expenditure. We also examined the residual of energy expenditure rate for each activity (i.e., after accounting for the other activities) and regressed that against wind speed and temperature.

To compare our values against other estimates of diving metabolic rate obtained for other species at other temperatures, we used linear relationships to convert all measured metabolic rates to the equivalent metabolic rate at 13 °C, following the protocol established by Enstipp et al. (figure 7 in ref. 20). We followed Enstipp et al. (20)—who actually chose 12.6 °C—and chose 13 °C because more previous studies used that water temperature than any other temperature; it is often equivalent to the typical temperature in a shallow pool used for respirometry, and therefore required transforming the fewest data. Diving metabolic rate decreases linearly with increasing temperature in endotherms diving in cold water (20–22), and we used linear relationships established for each taxonomic group to interpolate the value at 13 °C for those species that had not been previously measured at 13 °C. We also removed one early data point for African penguins (*Spheniscus demersus*) that had been flagged by other authors as too high (23).

Statistical Analyses. All statistical analyses were completed in R 3.2.1. Before using parametric statistics, we tested for normality (Shapiro–Wilk test) and homogeneity of variance (Levine's test). All percentages were arcsine-transformed before analysis. We only analyzed dives with maximum depth >3 m due to imprecision inherent in our TDRs. Body mass does not vary substantially in breeding murres (±15% from average value), and >90% of that variation occurs in metabolically inert lipid stores and total body water (10), so we did not expect a strong impact of body mass on metabolic rate. Nonetheless, we included body mass and body mass loss as covariates in the model. All values reported are means ± SE.

Are Our Values Realistic? Because past researchers have claimed that using doubly labeled water to measure activity-specific metabolic rates underestimates activity costs (e.g., ref. 13 criticized a different mathematical approach than the one we used),

we carefully considered whether our values obtained were realistic. First, we used a multiple regression rather than a simple regression to determine activity-specific metabolic rates (13). Simple regression tends to underestimate activity-specific metabolic rates by assuming that the intercept represents average activity-specific rates when none of the particular activity occurs; if different activities are intercorrelated, the slope can lead to erroneous measurements using simple regression. Second, we directly confirmed equations relating isotopic dilution to total body water content, and we obtained dietary samples from dozens of individuals so that we have a good representation of the respiratory quotient. Because energy expenditure measured via doubly labeled water is linearly (and mathematically) proportional to percent body water and respiratory quotient, by directly measuring those parameters in murres (10) we were able to reduce the absolute error in our energy expenditure measurements; most past studies that found discrepancies between doubly labeled water and direct respirometry did not measure percent body water directly (11). Third, we compared our values to those obtained on captive thick-billed murres diving in shallow dive tanks (21).

Croll and McLaren (21) measured metabolic rates of ~8 W for postabsorptive murres in the thermoneutral zone and resting in air, 28 W for murres preening in the thermoneutral zone, and 19 W for postabsorptive murres resting in water 0–5 °C (with an elevation to 24 W assuming additive effects of digestion and temperature, as much of the time resting on the surface in murres is believed to be associated with digestion). In comparison, our values of 9.2 W for resting on land and 26 W for resting on the water surface are quite similar, especially given that our measurements also included activities (such as preening). Diving metabolic rate in the thermoneutral zone for murres diving in shallow dive tanks was 21 W (21). Assuming a 14-W increase associated with diving in 0–5 °C water (21), the total diving metabolic rate in cold water was 35 W for an average dive duration of 41 s. In contrast, our value in the model assuming dive costs were proportional to dive time was 27 W, and our value (at 41 s) for the model assuming dive costs decreased exponentially with dive time was 50 W at 41 s, 40 W at 68 s (the average dive duration for our study population), and 13 W at 280 s, the maximum duration recorded for our study population (9). The differential effects of buoyancy in a shallow dive tank, the potential substitution of heat from exercise for thermoregulation, and the associated costs of returning peripheral body temperature to core temperature values meant that it was difficult to directly compare the diving metabolic rate from the earlier study (21) with our own, but values seemed to be relatively similar. Because there have been no past measurements of flight costs, we could not make comparisons between our costs and those measured previously.

Morphological Analyses. We use a multivariate approach to describe and compare the morphological space of 452 species of diving and nondiving birds (see ref. 24 for a similar approach). We collected data on wingspan, wing area, and body mass from two sources (ref. 25 and data appendix in ref. 26). We calculated aspect ratio (wingspan squared/wing area), wing loading (body mass/wing area), and a parameter we call the Pennycuick-induced drag coefficient (P_{ind}). P_{ind} is the induced drag coefficient from the Pennycuick model (18) divided by body mass:

$$P_{ind} = \frac{2.4 \text{ body mass}}{1.23\pi \text{ wingspan}^2}. \quad \textbf{[S5]}$$

The main difference between P_{ind} and wing loading is that P_{ind} uses the square of wingspan rather than wing area (18). Wingspan is a simpler parameter to measure than wing area. Because many of these parameters are intercorrelated and/or constructed from one another, we log-transformed each of the six parameters (body mass, wing area, wing-span, wing loading, aspect ratio, and P_{ind}) and performed a principal components analysis (PCA) to reduce the parameters to a single set of three orthogonal vectors that described over 99% of the morphological space. We used a multigroup discriminant analysis and a multivariate analysis of variance to determine whether five groups could be distinguished: flightless divers, plunge divers, foot-propelled divers, wing-propelled divers, and nondiving flyers. We classified shallow, wing-propelled divers, such as albatrosses and shearwaters (27), as nondiving flyers. A separate analysis that included both of those two groups showed no morphological differences between those two groups and other nondiving flyers.

In the PCA (Fig. S3*E*), PC1 explained 84.9% of the variation, PC2 explained 14.0% of the variation, and PC3 explained 1.1% of the variation. We ignored all further principal components because PC1–PC3 explained 99.99% of the variation. In brief, PC1 represented variation in body mass, whereas PC2 represented variation associated with the variation in wing area that is independent of body mass. Thus, wing- and foot-propelled divers (flightless or flying) were distinguished from nondivers primarily along PC2.

The discriminant analyses on PC1–PC3 correctly classified 100% of penguins, 96% of flyers, 93% of flying, wing-propelled divers, 46% of foot-propelled divers, but none of the plunge divers or shearwaters (Fig. S3*F*). A multivariate analysis of variance was highly significant ($F_{15,993} = 31.0$, $P < 0.0001$) with Tukey's post hoc tests showing no difference between shearwaters and flyers ($p_{adj} = 0.70$) or plunge divers and flyers ($p_{adj} = 0.94$), but significant differences among flyers and foot-propelled divers, flying wing-propelled divers, and penguins (all $p_{adj} < 0.0001$). Shearwaters, plunge divers, and flyers did not differ significantly from one another (all $p_{adj} > 0.5$), whereas foot-propelled divers, flying wing-propelled divers, and penguins did differ significantly from one another (all $p_{adj} < 0.001$). We used the statistical similarity of shearwaters, plunge divers, and flyers as further justification for considering all three as flyers. The discriminant analyses on PC1–PC3, with shearwaters and plunge divers categorized as flyers, correctly classified 100% of penguins, 98% of flyers, 93% of flying, wing-propelled divers, and 42% of foot-propelled divers (Fig. S3*G*). A multivariate analysis of variance was highly significant ($F_{9,999} = 46.4$, $P < 0.0001$) with Tukey's post hoc tests showing significant differences among flyers and foot-propelled divers, flying wing-propelled divers, and penguins (all $p_{adj} < 0.0001$).

We then correlated the morphological data (PC2 and P_{ind}) for species using primarily flapping flight against the residual of flight costs on the maximum output line (from Fig. 1*A*) for species using primarily flapping flight. We used the maximum output line as the expected value for flapping flight to account for body mass, because many birds using flapping flight cluster along that line (Fig. 1*A*). We chose P_{ind} because of its close connection to aerodynamic theory.

Energy expenditure in flight for a given body mass increased with both P_{ind} (Fig. S1*C*) and PC2 (Fig. S1*D*), with morphology (PC2/P_{ind}) explaining 30–40% of the variation once body mass was accounted for.

Scenario for the Evolution of Flightlessness in Diving Birds. Large body size allows for increased dive duration via (*i*) reduced mass-specific oxygen consumption rate; (*ii*) increased volume of oxygen stores; and (*iii*) increased density of oxygen stores (myoglobin concentrations) (28, 29). At the level of the individual, there is therefore a strong selective pressure for greater body size in diving birds, allowing for exploitation of deeper prey and more efficient diving (more time underwater for each surface pause) at all dive depths (28). In flying birds, those selective pressures are

balanced by selective pressures for reduced body mass to reduce flight costs (18). Species that nest on predator-free islands close to consistent, dense prey aggregations can become flightless and evolve higher body mass (27).

Flightlessness in diving birds appears to be associated with rapid evolution of large body size, as shown by the presence of early giant penguins. In support of this scenario, (*i*) extant and fossil wing-propelled diving birds are flightless above ∼1 kg, and flying below 1 kg; (*ii*) dive costs increase more slowly with body size in flightless than in flying birds (Fig. 1*B*); and (*iii*) the flying, wing-propelled diving bird space is strongly tilted along discriminant axis 1, which represents body mass (Fig. 1*D*), suggesting particularly strong morphological tradeoffs for heavy birds that use their wings both for flying and diving. We therefore speculate that once the benefits of deep-diving outweighed the benefits of flying, populations rapidly evolved larger individual body size, leading to increasing benefits in terms of dive efficiency. Development of a wing and feather structure optimized for underwater locomotion would have taken longer evolutionary time, partly because remodeling the wing may have required alterations in the kinematics of underwater flight. In contrast, flightless foot-propelled divers that have little use for wings (e.g., such as flightless cormorants *Phalacrocorax harrisi*, 3–5 kg) are only slightly larger than flying cormorants (e.g., usually 1–3 kg; Fig. 1*B*) and lack highly remodeled wing bones, but have very reduced primary feathers. We propose that flightlessness in wing-propelled diving birds therefore involves (*i*) inefficient foot-propelled or wing- and foot-propelled swimming at shallow depths among the first small, shallow-diving seabirds; (*ii*) the evolution of increased body size, reduced wing size, and increased dive efficiency in flying birds as they move toward the boundary of the adaptive valley in morphological space; (*iii*) the rapid "jump" to flightlessness, permitting an increase in body size and resulting dive efficiency (the great auk); and (*iv*) the slow remodeling of the wing architecture (bone and feathers) toward a form optimized for swimming.

1. Croll DA, Gaston AJ, Burger AE, Konnoff D (1992) Foraging behavior and physiological adaptation for diving in thick-billed murres. *Ecology* 73(1):344–356.
2. Falk K, Benvenuti S, Dall'Antonia L, Kampp K, Ribolini A (2000) Time allocation and foraging behaviour of chick-rearing Brünnich's guillemot *Uria lomvia* in high arctic Greenland. *Ibis* 142(1):82–92.
3. Falk K, Benvenuti S, Dall'Antonia L, Gilchrist G, Kampp K (2002) Foraging behaviour of thick-billed murres breeding in different sectors of the North Water polynya: An inter-colony comparison. *Mar Ecol Prog Ser* 231:293–302.
4. Paredes R, Jones IL, Boness DJ (2004) Reduced parental care, compensatory behaviour and reproductive costs of thick-billed murres equipped with data loggers. *Anim Behav* 69(1):197–208.
5. Watanuki Y, Niizuma Y, Gabrielsen GW, Sato K, Naito Y (2003) Stroke and glide of wing-propelled divers: Deep diving seabirds adjust surge frequency to buoyancy change with depth. *Proc Biol Sci* 270(1514):483–488.
6. Watanuki Y, et al. (2006) Swim speeds and stroke patterns in wing-propelled divers: A comparison among alcids and a penguin. *J Exp Biol* 209(Pt 7):1217–1230.
7. Elliott KH, Gaston AJ, Davoren GK (2008) Time allocation by a deep-diving bird reflects energy expenditure. *Anim Behav* 75(4):1310–1317.
8. Elliott KH, Gaston AJ, Davoren GK (2008) Time allocation by a deep-diving bird reflects energy gain and prey type. *Anim Behav* 75(4):1301–1310.
9. Elliott KH, Davoren GK, Gaston AJ (2007) The influence of buoyancy and drag on the dive behaviour of an arctic seabird, the thick-billed murre. *Can J Zool* 85(3): 352–361.
10. Jacobs SR, et al. (2012) Determining seabird body condition using nonlethal measures. *Physiol Biochem Zool* 85(1):85–95.
11. Speakman JR (1997) *Doubly-Labelled Water: Theory and Practice* (Chapman & Hall, London).
12. Gaston AJ, Woo K, Hipfner JM (2003) Trends in forage fish populations in northern Hudson Bay since 1981, as determined from the diet of nestling thick-billed murres *Uria lomvia*. *Arctic* 56(3):227–233.
13. Wilson RP, Culik BM (1993) Activity-specific metabolic rates from doubly labeled water studies: Are activity costs underestimated? *Ecology* 74(4):1285–1287.
14. Wilson RP, Hustler K, Ryan PG, Burger AE, Noldeke EC (1992) Diving birds in cold water: Do Archimedes and Boyle determine energetic costs? *Am Nat* 140(2):179–200.
15. Lovvorn JR, Watanuki Y, Kato A, Naito Y, Liggins GA (2004) Stroke patterns and regulation of swim speed and energy cost in free-ranging Brünnich's guillemots. *J Exp Biol* 207(Pt 26):4679–4695.
16. Knower Stockard T, et al. (2005) Air sac P_{O2} and oxygen depletion during dives of emperor penguins. *J Exp Biol* 208(Pt 15):2973–2980.
17. Culik BM, et al. (1996) Diving energetics in king penguins (*Aptenodytes patagonicus*). *J Exp Biol* 199(Pt 4):973–983.
18. Pennycuick CJ (2008) *Modelling the Flying Bird*. Theoretical Ecology Series (Academic, New York), Vol 5.
19. Elliott KH, et al. (2008c) Seabird foraging behaviour indicates prey type. *Mar Ecol Prog Ser* 354:289–303.
20. Enstipp MR, Grémillet D, Jones DR (2006) The effects of depth, temperature and food ingestion on the foraging energetics of a diving endotherm, the double-crested cormorant (*Phalacrocorax auritus*). *J Exp Biol* 209(Pt 5):845–859.
21. Croll DA, McLaren E (1993) Diving metabolism and thermoregulation in common and thick-billed murres. *J Comp Physiol B* 163(2):160–166.
22. Lovvorn JR (2007) Thermal substitution and aerobic efficiency: Measuring and predicting effects of heat balance on endotherm diving energetics. *Philos Trans R Soc Lond B Biol Sci* 362(1487):2079–2093.
23. Nagy KA, Sigefried WR, Wilson RP (1984) Energy utilization by free-ranging jackass penguins, *Spheniscus demersus*. *Ecology* 65(5):1648–1655.
24. Ricklefs RE (2012) Species richness and morphological diversity of passerine birds. *Proc Natl Acad Sci USA* 109(36):14482–14487, 10.1073/pnas.1212079109.
25. Osa Y (1994) Study of functional morphology for swimming and/or flying seabirds. PhD thesis (Tokyo Univ of Fisheries, Tokyo).
26. Kaiser GW (2007) *The Inner Bird* (Univ of British Columbia Press, Vancouver).
27. Burger AE (2001) Diving depths of shearwaters. *Auk* 118(3):755–759.
28. Mori Y (2002) Optimal diving behaviour for foraging in relation to body size. *J Evol Biol* 15(2):269–276.
29. Elliott KH, Shoji A, Campbell KL, Gaston AJ (2010) Oxygen stores and foraging behavior of two sympatric, planktivorous alcids. *Aquat Biol* 8:221–235.

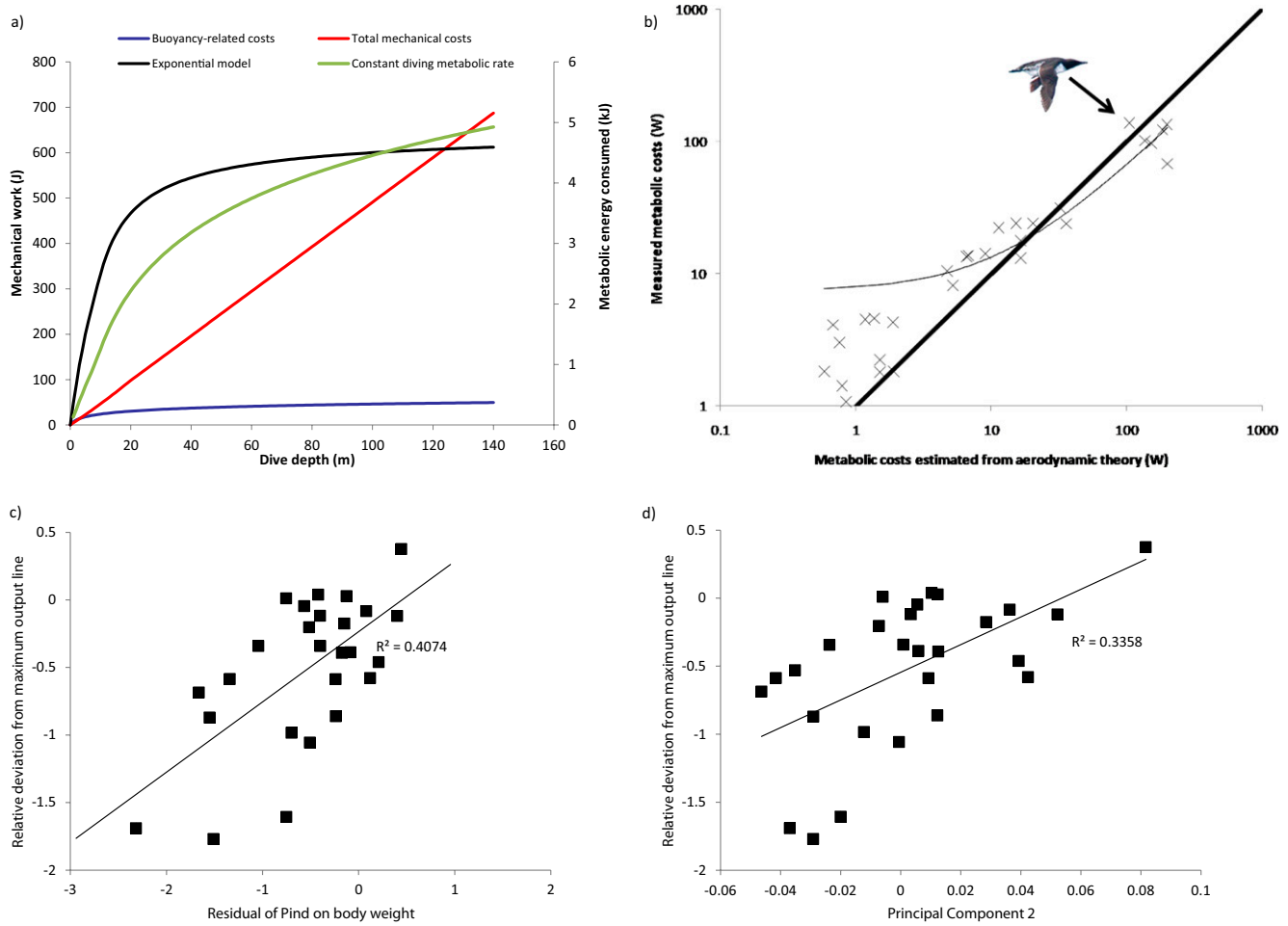

Fig. S1. Models used to estimate biomechanical costs and their comparisons with morphology. (*A*) Comparison among four different models describing how metabolic rate changes with dive depth. Buoyancy-related costs and total mechanical costs are represented as mechanical work (left scale) and converted to metabolic energy (right scale) assuming a constant efficiency. The exponentially declining and constant metabolic rate models are represented directly in metabolic energy consumed (in kilojoules, right scale). Because the data themselves generate the coefficient associated with converting the different functions into metabolic rates, the shapes of the functions rather than their absolute values are being compared. (*B*) Measured flight costs across 29 bird species compared with flight costs estimated from aerodynamic theory for the same bird species (Pennycuick model, as generated by the program Flight 1.24). The thin line represents the least-squares linear regression, which is curvilinear on the log-log graph. The solid line is the 1:1 line: the value if measured values were equal to the values predicted by the Pennycuick model. Murres had the highest absolute residual from the best-fit linear trend line (65 W) and, among large birds, had the highest absolute residual from the 1:1 line (33 W). Relative deviation from maximum output line [(measured value – expected value from maximum output line)/measured value] increased with both (*C*) P_{ind} and (*D*) PC2 (Fig. S4).

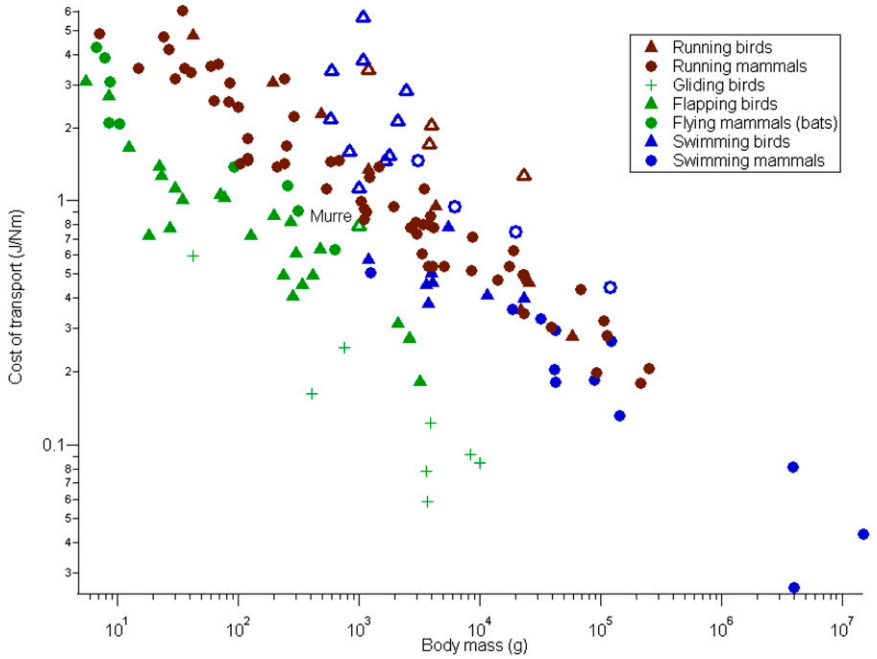

Fig. S2. Cost of transport for different modes of locomotion for birds and mammals. Open symbols represent animals that face functional tradeoffs with movement in different media: flying wing-propelled birds (when flying), human swimmers, sea otters and flying, diving birds (when swimming) and penguins and geese (when running).

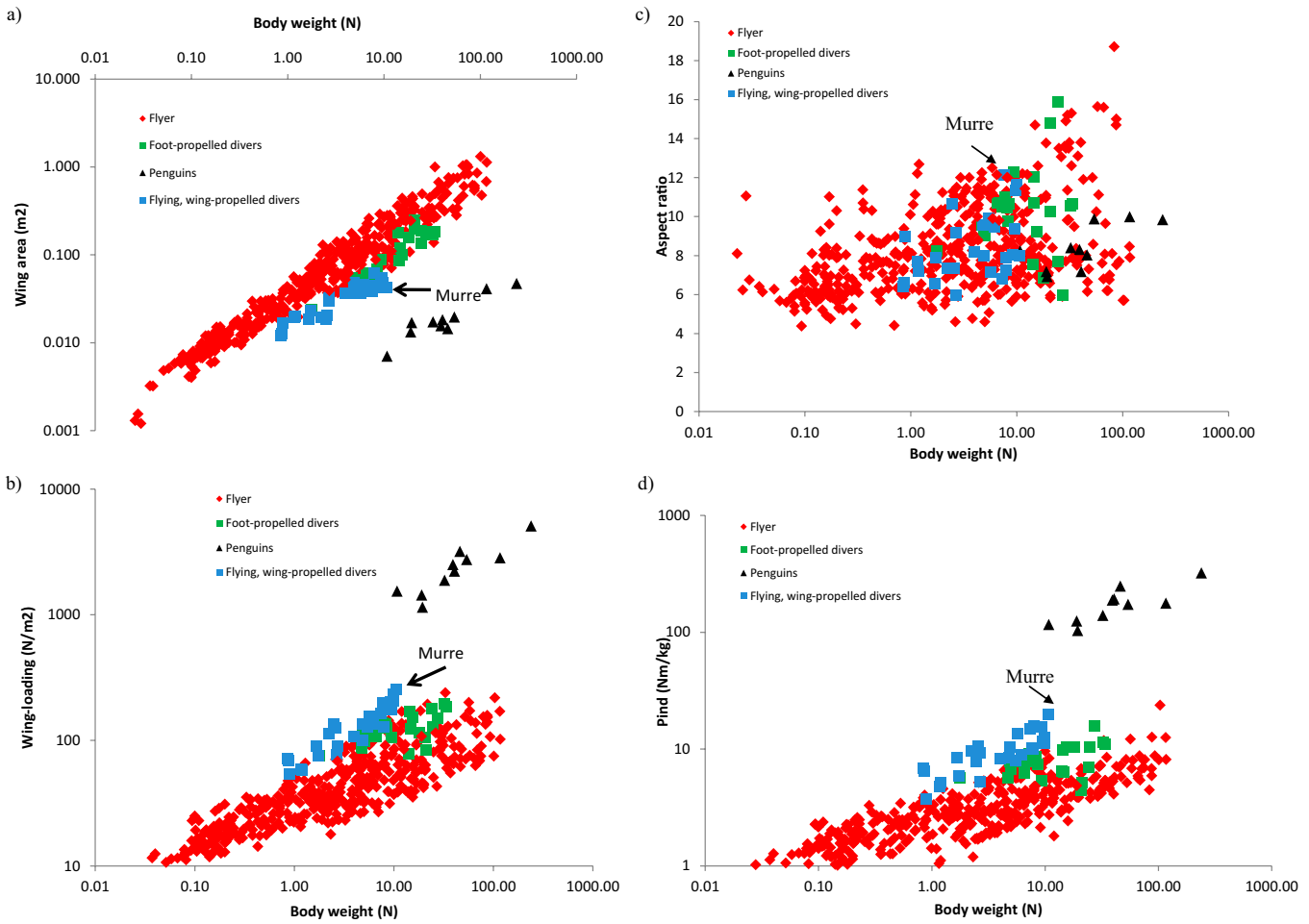

Fig. S3. (Continued)

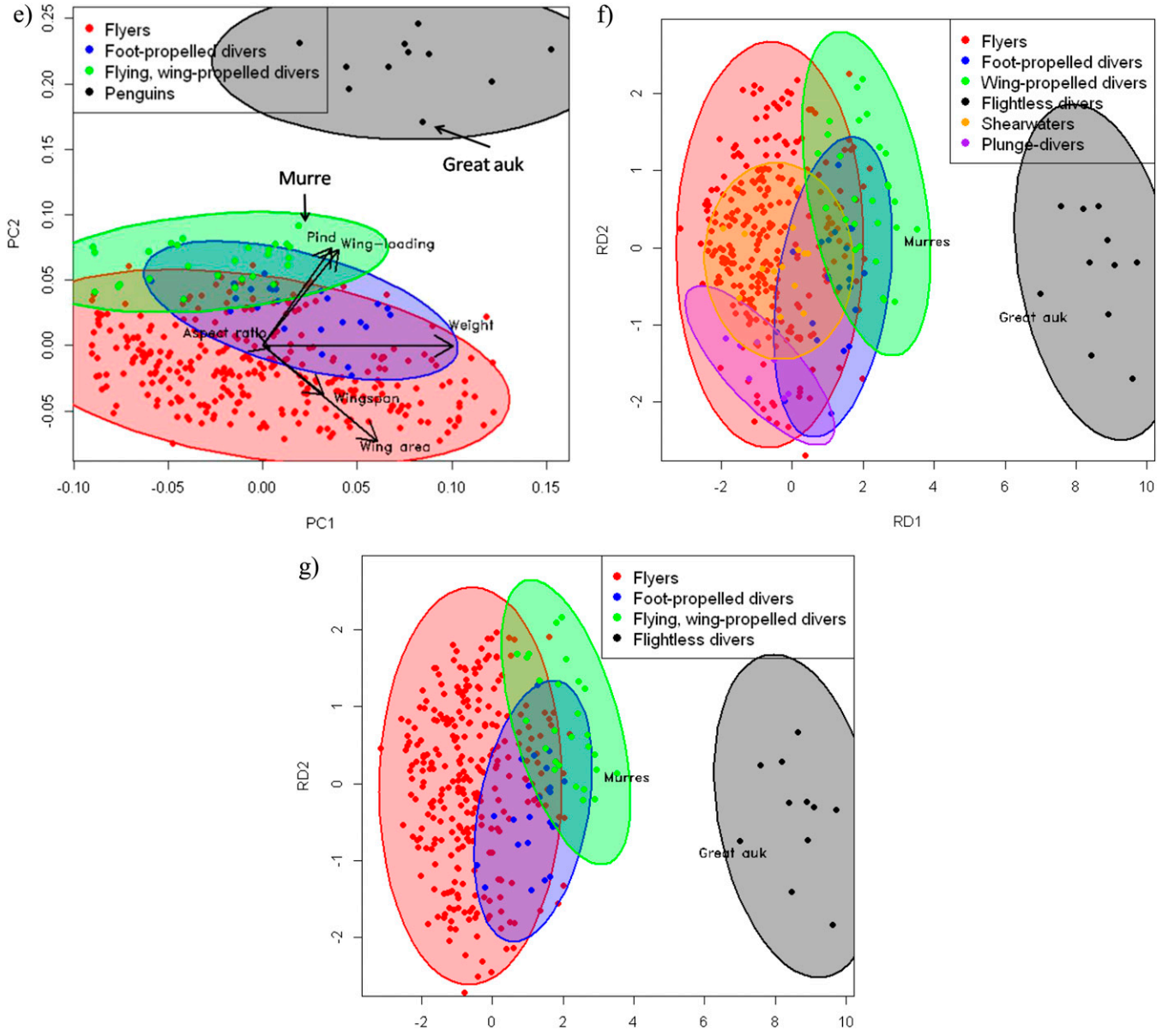

Fig. S3. Comparison of morphology across bird species. (*A*) Wing area; (*B*) wing-loading; (*C*) aspect ratio; and (*D*) P_{ind} as a function of body weight for bird species. (*E*) PCA of log-transformed avian morphology. The 95% confidence ellipses are shown for each group. The lengths of the arrows representing morphological variables were divided by 7.5 for ease of presentation. Penguins are excluded from wing-propelled divers. (*F* and *G*) Discriminant analysis of log-transformed avian morphology, with *Puffinus* shearwaters and plunge divers treated as separate groups (*F*), and *Puffinus* shearwaters and plunge divers treated as flyers (*G*). "Murres" and "Great auk" are printed immediately below the appropriate point.

Fig. S4. Humerus and wing of (*Left*) thick-billed murre; (*Center*) great auk; and (*Right*) Galapagos penguin. The value of PC2 (Fig. S3*E*) is low for the murre and high for the auk and penguin.