How hibernation in frogs drives brain and reproductive evolution in opposite directions

Wen Bo Liao[1,2,*], Ying Jiang[1,2], Long Jin[1,2] & Stefan Lüpold[3,*]

[1] Key Laboratory of Southwest China Wildlife Resources Conservation (Ministry of Education), China West Normal University, Nanchong, Sichuan, China

[2] Key Laboratory of Artificial Propagation and Utilization in Anurans of Nanchong City, China West Normal University, Nanchong, Sichuan, China

[3] Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland

*Corresponding authors: Wen Bo Liao, Stefan Lüpold

Email: liaobo_0_0@126.com (WBL) and stefan.luepold@ieu.uzh.ch (SL)

Author Contributions: W.B.L. conceptualized the study; W.B.L., Y.J., and L.J. collected the data; S.L. analyzed the data; and W.B.L. and S.L. wrote the manuscript.

Competing Interest Statement: The authors declare that they have no conflict of interest.

Keywords: Anura | Brain size | Fat storage | Resource allocation | Trade-off | Testis size

Abstract

Environmental seasonality can promote the evolution of larger brains through cognitive and behavioral flexibility but also hamper it when temporary food shortage is buffered by stored energy. Multiple hypotheses linking brain evolution to resource acquisition and allocation have been proposed, albeit separately for different groups of birds or mammals rather than being directly compared within any single group. Here, using direct tissue measurements and experimentally validated brumation ('hibernation') parameters, we integrated these hypotheses across frogs in the context of varying brumation duration and its environmental correlates. We show that protracted brumation reduces brain size and instead promotes reproductive investments, likely in response to brumation-dependent changes in the socio-ecological context that ultimately affect the operation of sexual selection and evolution of mating systems. Our results reveal novel insight into the complex processes of brain and reproductive evolution in organisms whose 'cold-blooded' metabolism is particularly susceptible to environmental seasonality.

Introduction

Seasonal fluctuations in climatic conditions and primary productivity can result in temporary food limitation, imposing energetic constraints on animals. Maintaining a positive energy balance across seasons, or at least minimizing the negative balance during lean periods, can be achieved by more constant net energy intake than predicted solely by food abundance (Sol, 2009), or by reduced investments in costly organs (Heldstab et al., 2018). At the center of both strategies is the size of the brain. On the one hand, a relatively large brain can improve the cognitive ability and behavioral flexibility (Reader and Laland, 2002; Lefebvre et al., 2004; Sol et al., 2005; Benson-Amram et al., 2016) to locate more diverse and dispersed food sources to

buffer the environmental fluctuations of seasonal habitats (‘cognitive buffer hypothesis’) (Allman et al., 1993; Sol, 2009; van Woerden et al., 2012). On the other hand, as the high metabolic costs of brain tissue (Mink et al., 1981; Aiello and Wheeler, 1995; Lukas and Campbell, 2000) may not be temporarily reducible (Mink et al., 1981), periodic food scarcity is expected to constrain brain size evolution (‘expensive brain hypothesis’) (Isler and van Schaik, 2009).

Extreme fluctuations in resource acquisition and total metabolic activity are found in hibernating species, whose basal metabolic rate may drop by over 90% during hibernation (Ruf and Geiser, 2015). Such a radical reduction in metabolism is likely to limit investments in the maintenance of brain function and thus the support of a large brain. Indeed, mammals with some period of hibernation tend to have relatively smaller brains than their non-hibernating counterparts (Heldstab et al., 2018).

In addition to cognitive responses (van Woerden et al., 2010), periods of food scarcity can also promote the evolution of physiological responses. For example, a longer digestive tract may permit more efficient resource accumulation during a short active period (Sibly, 1981) and so could be favored by selection in species with prolonged hibernation. This response could parallel brain size evolution or result in an evolutionary trade-off between the two organs (‘expensive tissue hypothesis’) (Aiello and Wheeler, 1995), similar to predicted trade-offs between brain size and sexually selected traits (‘expensive sexual tissue hypothesis’) (Pitnick et al., 2006) or other costly organs more generally (‘energy trade-off hypothesis’) (Isler and van Schaik, 2006). Further, physiological buffering is often accompanied by a seasonal reduction in the metabolic rate or activity (e.g., hibernation), with energy drawn from stored fat reserves (Heldstab et al., 2016). Even though hibernation does not preclude benefits of cognitive abilities (e.g., to efficiently accumulate fat reserves before hibernation), the evolution of a relatively larger brain could be hampered by seasonally alternating between cognitive benefits during the

active period and maintenance costs at no obvious benefit during hibernation. In fact, it has been proposed that the benefits of maximal fat stores and minimal metabolic expenditure on somatic maintenance during hibernation, on average, are likely to outweigh those of a large brain ('fat−brain trade-off hypothesis') (Navarrete et al., 2011; Heldstab et al., 2016). By comparison, non-hibernators may always gain a positive net benefit of a relatively large brain, possibly even enhanced if it mitigates resource acquisition when food is seasonally scarce (Allman et al., 1993; Sol, 2009; van Woerden et al., 2012). If so, the probability of positive selection on brain size should be higher in the absence of hibernation, providing one explanation for the relatively smaller brains in hibernating compared to non-hibernating mammals (Heldstab et al., 2018).

Allocating a fixed energy budget to the demands of different organs throughout prolonged hibernation might also have important reproductive consequences. Particularly for species that reproduce almost immediately after emerging from hibernation, as in some mammals (Psenner, 1957; Place et al., 2002) and many amphibians (Wells, 1977; Fei and Ye, 2001), the seasonal recrudescence of their reproductive tissue necessarily occurs before emergence when the stored resources are most limited (Isler and van Schaik, 2009; Isler, 2011). Reproductive investments, however, are intimately linked to fitness, such as testis size that may be under intense selection by sperm competition resulting from female multiple mating (Lüpold et al., 2020). Female promiscuity itself is prevalent where males are less able to effectively monopolize their mates (Lüpold et al., 2014), and this would seem particularly likely when the breeding activity is highly synchronized in dense aggregations (Lüpold et al., 2014, 2017). Indeed, across anurans (frogs and toads) that are often bound to small water bodies for reproduction, males invest relatively more in their testes and less in their forelimbs (used in pre-mating competition) as population density increases (Buzatto et al., 2015; Lüpold et al., 2017). If

the breeding activity is more synchronized because of a shortened active period, thus increasing the risk of either competitive fertilization or simply of sperm depletion by a high mating rate (Vahed and Parker, 2012), selection for relatively larger testes would be stronger precisely where the available fat stores need to last longer, with likely consequences for resource demands and allocation while overwintering. Including reproductive investments and breeding patterns in studies of allocation trade-offs in response to hibernation and environmental seasonality would thus seem critical but remains to be done, particularly in the context of brain evolution.

The opposing selection pressures on brain or gonad size (i.e., cognitive or fitness benefits *versus* metabolic costs), varying degrees of seasonality and diverse strategies of buffering periodic food scarcity (e.g., cognitive *versus* physiological) between species render environmental fluctuations an ideal context to study brain size evolution. The different hypotheses invoked to explain the coevolution of brain size with other organs were each independently developed for a separate set of mammalian or avian taxa, the two vertebrate classes with the relatively largest brains (Jerison, 1973). These hypotheses have yet to be directly tested against one another in a single taxon and ideally in the immediate context of seasonal activity, considering the extent, rather than the mere presence/absence, of hibernation. This last point is important to the extent that the classification of ‘hibernation’ could range anywhere between one or multiple brief inactive bouts (i.e., minimally different from non-hibernating species) and spending most months of the year in dormancy, with severe energetic and life-history constraints despite again simply being classified as hibernating. Further, understanding the generality of the patterns reported in the large-brained mammals or birds requires validation in other taxa, ideally with smaller brains and different overall energetic demands. Such generalization would also permit contextualizing brain evolution in the two

largest-brained taxa in terms of the strength of selection on encephalization relative to potential metabolic constraints.

A particularly suitable system is presented in ectothermic ('cold-blooded') species whose metabolism and activity are tightly linked to their ambient temperature and considerably hampered outside a species-specific temperature range (Wells, 2007). Hence, temperature constraints on metabolism and activity patterns potentially set stricter physical boundaries to the ability to buffer food scarcity through behavioral flexibility than in the endothermic ('warm-blooded') mammals or birds. Additionally, the different groups of mammals, for which most hypotheses on brain evolution were proposed, exhibit extremely diverse Bauplans and lifestyles that could confound overall conclusions. By contrast, anurans are relatively homogeneous in body size and shape, diet or locomotion (Kardong, 2019), but still markedly divergent in their relative brain size (Liao et al., 2022) or in reproductive (Lüpold et al., 2017) and other investments (Wells, 2007) in response to environmental variation. Consequently, any resource trade-offs around the evolution of brain size could, if they exist, be easier to isolate across anurans than across mammals.

Here, we examine variation in male brain and testis size relative to body fat, limb muscles and the main visceral organs (see Fig. 1) across the males of 116 anuran species in the context of the vastly varying 'hibernation' periods and their environmental correlates. Frogs differ from mammals in their physiology of hibernation, in that the seasonal inactivity is a consequence of ambient temperatures dropping below the activity range rather than of active metabolic depression (referred to here as 'brumation' for distinction) (Pinder et al., 1992; Wells, 2007). Yet, important parallels remain in terms of resource allocation. In both taxa, hibernating animals spend extended periods depleting fixed, previously accumulated energy stores across sequential investments, albeit at a reduced metabolic rate (Staples, 2016). By contrast, non-

hibernators can at least partially compensate for spent resources as they go, but their higher metabolic rate requires continued resource acquisition even when food is seasonally scarce (Heldstab et al., 2016). Such differences between strategies are likely to affect the relative costs and benefits of different organs, and thus how species optimally allocate resources between them. Relatively larger traits (e.g., brain) can evolve—possibly at the cost of other traits—if they confer some net fitness benefits.

Brumating anurans drop their heart rate, become sluggish, draw the nictitating membranes across the eyes for protection, spread their legs for stability, and undergo multiple physiological changes to protect against freezing or to switch from pulmonary to cutaneous gas exchange or from aerobic to anaerobic metabolism (Pinder et al., 1992; Fei and Ye, 2001; Wells, 2007; Tattersall and Ultsch, 2008). Yet, several anuran species are known to regularly move in their burrows or underwater hibernacula in response to changes in soil temperature or oxygen concentration (van Gelder et al., 1986; Stinner et al., 1994; Holenweg and Reyer, 2000), or when disturbed (Tattersall and Ultsch, 2008; Niu et al., 2022). Brumating frogs can also exhibit higher levels of brain cell renewal than active ones, possibly to avoid brain damage (Cerri et al., 2009), and it seems plausible that the total investment in such renewal would increase with the amount of brain tissue. Further, larger brains could also be less tolerant to the often hypoxic brumation conditions (Pinder et al., 1992; Wells, 2007; Tattersall and Ultsch, 2008) if the findings from other ectotherms extend to brumating anurans (Sukhum et al., 2016). Overall, maintaining a relatively larger brain while brumating is likely to come with greater costs that could constrain brain evolution compared to species that show only brief or no seasonal inactivity.

Instead of the typical indirect proxies or scores (Navarrete et al., 2011; Heldstab et al., 2016, 2018; Luo et al., 2017), we directly quantified seasonal changes in tissue sizes and estimated

the brumation duration as the period during which temperatures were continuously below the species-specific activity threshold. These estimated brumation durations corresponded to the periods of no frog activity detected in field surveys. Our results provide robust evidence that the duration of brumation under varying climates modifies how frogs allocate their limited resources. We then resolve the direct and indirect links between these variables to test the different hypotheses on brain evolution, which have been independently proposed for different avian and mammalian taxa, within our single set of anurans. We further integrate variation in breeding contexts and reproductive evolution in this context of resource allocation. Our broad comparative approach permits disentangling important evolutionary responses to environmental seasonality with its far-reaching behavioral and physiological consequences for species during both their active and inactive periods.

Results

Determinants of brumation duration. For a mean (±SD) of 3.41 ± 0.95 males in each of 116 anuran species, we combined experimentally determined thermal activity thresholds with multi-year temperature fluctuations at their collection sites to estimate the species-specific brumation periods (details and validation in *Material and Methods*). These periods averaged between 0.6 ± 0.5 and 250.5 ± 16.7 days across the five years examined, with high repeatability within species (R = 0.95 [95%CI: 0.93, 0.96]; Fig. S1; Table S1). Twenty-two species were predicted to overwinter for ≤9 days, another three species for 18−27 days, and the remaining 91 species for ≥47 days, resulting in a rapid shift between 9 and 47 days. Since the ground microclimate may buffer some of the fluctuations in air temperature and frogs can endure short cold spells without dormancy, we conservatively considered the 25 species with ≤27 days below their experimental temperature threshold as unlikely to show any sustained brumation.

In phylogenetic regressions (Freckleton et al. 2002), the brumation period increased with both latitude and elevation of the study sites, as well as with the variation in temperature ($r \geq 0.42$, $t_{114} \geq 5.00$, $P < 0.001$; Table S2). By contrast, the brumation period was inversely related with the annual mean temperature and precipitation, and the duration of the dry season ($r < -0.29$, $t_{114} < -3.24$, $P < 0.002$), but not significantly associated with longitude or the variation in precipitation ($r < 0.09$, $t_{114} < 0.95$, $P > 0.34$; Table S2). Except for the period of the dry season, all these results were qualitatively identical when focusing only on the 91 species with some expected brumation period (see above; Table S3). Among these 91 species, those inhabiting cooler and more seasonal climates entered and emerged from their inactive state at lower temperatures (Table S4), suggesting an increased cold tolerance to maximize their active period.

Effect of brumation on individual tissue investments. Based on this variation in the periods of active resource acquisition or metabolizing stored resources, respectively, we explored potential consequences for resource allocation in the same set of males across the 116 species. In phylogenetic regressions (Freckleton et al. 2002), neither snout-vent length (SVL) nor body mass covaried with the duration of brumation or any other environmental variable ($|r| \leq 0.15$, $|t_{114}| \leq 4.03$, $P \geq 0.11$; Table S5), with only a weak, non-statistically significant trend toward reduced body mass at higher elevation ($r = -0.18$, $t_{114} = -1.91$, $P = 0.06$, phylogenetic scaling parameter $\lambda = 0.94$ [95%CI: 0.85, 1.00]). The different organs and tissues themselves all increased with body size, but with different allometric slopes. Whereas the mass of body fat and of both the forelimb and hindlimb muscles showed a disproportionately steep increase relative to body size across the same 116 species (all $\beta \geq 1.10$ [1.03, 1.18]), the allometric slope was substantially shallower for brain size ($\beta = 0.49$ [0.44, 0.54]; Table S6; Fig. 1A). All remaining

tissues did not deviate from proportionate scaling (i.e., 95%CI including 1.00; Table S6; Fig. 1A). Hence, the evolution of brain size appears to be more constrained than that of other organs when selection favors larger body size.

With these different investments in brain size compared to other expensive tissues, we next tested if the evolution of brain size is constrained by extended brumation, similar to what has been suggested for mammals based on the mere presence or absence of any hibernation (Heldstab et al. 2018). Absolute brain size was independent of the brumation period (r = −0.10, t_{115} = −1.08, P = 0.28, λ = 0.89 [0.73, 0.97]), but controlling for SVL as a proxy of body size, males of species with protracted hibernation had relatively smaller brains, whether quantified during the breeding season (partial r, r_p = −0.31, t_{113} = −3.50, P < 0.001, λ = 0.35 [0.00, 0.61]; Fig. 2A, Table S7) or shortly before entering hibernation (N = 50 species means based on 2.64 ± 0.94 males each: r_p = −0.51, t_{47} = −4.03, P < 0.001, λ = 0.00 [0.00, 0.56]; Table S7). This trend was not a mere result of changing body size in response to brumation, as the SVL itself was independent of the hibernation period (see above; Table S5). On average, brumating species tended to have relatively smaller brains than those that are unlikely to overwinter for an extended period (Table S8), similar to the study of mammals that used the presence/absence of hibernation as a predictor (Heldstab et al. 2018). However, the above trend also applied to those 91 species that we classified as brumating for some period (Table S9), thus providing stronger evidence for a link between brumation and brain evolution than the coarse binary classification. This pattern further remained robust when we recalculated the predicted brumation periods such that frogs were only considered to enter their inactive state at temperatures that were 2°C or 4°C below their experimentally derived thresholds. These two more conservative thresholds for brumation simulated a potential buffering effect of underground burrows and other shelters relative to the seasonal variation in air temperatures

that was available in meteorological databases (details and validation in *Material and Methods*; Table S10).

Species with a prolonged brumation period further had relatively more body fat ($r_p \geq 0.25$, $t_{113} \geq 2.72$, $P \leq 0.008$; Fig. 2B) and, at least in breeding condition, relatively larger testes ($r_p = 0.36$, $t_{113} = 4.06$, $P < 0.001$, $\lambda = 0.77$ [0.40, 0.90]; Fig. 2C) and relatively smaller hindleg muscles ($r_p = -0.22$, $t_{113} = -2.37$, $P = 0.02$, $\lambda = 0.22$ [0.00, 0.51]; Tables S7). These patterns were generally consistent when using the presence/absence of brumation (Table S8), buffered temperature fluctuations (Tables S10), or excluding the 25 species that are unlikely to overwinter (except for the non-significant effect on body fat; Table S9). The size of the remaining tissues was independent of brumation (Tables S7 to S10) and did not change significantly between the two sampling periods (Fig. 2).

Comparisons between pre- and post-brumation males, on average across all species, revealed a ca. 50% decline in fat tissue and 100% increase in testis size, indicating resource depletion and testicular recrudescence while brumating, respectively (Fig. 1B). Of the remaining expensive tissues, only the 95%CI of brain size excluded zero, but this change was small compared to the two traits above and within the range of the many tissues with no significant change. Thus, the biological significance of this putative increase in brain size during brumation is questionable as it could simply be attributable to general differences between the few individuals per species that were sampled during each period. We will thus refrain from further interpretation. Among the two tissues with considerable change, the extent of fat depletion increased significantly with the brumation period ($r = -0.36$, $t_{48} = -2.68$, $P = 0.01$, $\lambda = 0.00$ [0.00, 0.55]), as did that of testis regrowth ($r = 0.39$, $t_{48} = 2.89$, $P = 0.006$, $\lambda = 0.82$ [0.16, 0.98]; Fig. 1C).

Since the patterns for relative testis size might be a response to a shorter, more synchronized mating season when brumation is long (Wells 2007), we tested for links between brumation duration and different breeding parameters. Here, prolonged brumation shortened the breeding season (t_{41} = −4.47, P < 0.001, λ = 0.00 [0.00, 0.38]; Fig. S2A), with a much stronger effect than any of the climatic variables (Table S11), particularly when considered jointly (Table S12). Hence, the effect of these climatic variables may be indirect, mediated by brumation. A shorter breeding season further increased the probability of dense breeding aggregations (phylogenetic logistic regression: N = 42, z = −3.03, P = 0.002, α = 0.02; Fig. S2B). That brumation might mediate the effects of climatic variables on breeding aggregations via the duration of the breeding season was also supported by a phylogenetic confirmatory path analysis (von Hardenberg and Gonzalez-Voyer 2013; Gonzalez-Voyer and von Hardenberg 2014) (Fig. S3; Table S13). Finally, when combining these data with previously published data on the density of breeding populations (Lüpold et al. 2017) (N = 8 species overlapping), this effect was also supported by a trend toward higher mean population densities in species with a shorter breeding season (r = −0.69, t_6 = −2.37, P = 0.06, λ = 0.00 [0.00, 1.00]), albeit based on a small sample size (Fig. S2C).

To explore possible causal links between these breeding parameters and variation in relative testis size, we conducted directional tests of trait evolution (Pagel 1994; Revell 2012), which test if changes in two traits are unilaterally dependent, mutually dependent, or independent (Pagel 1994). Although independent evolution was the best-supported scenario based on the Akaike Information Criterion (AIC), the model with changes in relative testis size dependent on those in the breeding season was not significantly different (ΔAIC = 0.83, w_{AIC} = 0.33 compared to independent model with w_{AIC} = 0.50) and suggested that testis size was most likely to increase in response to a shortened breeding season (Fig. S4A). A model predicting increases in relative

testis size in response to aggregation formation pointed in a similar direction (Fig. S4B), albeit not below a ΔAIC cut-off of 2 (ΔAIC= 2.38, w_{AIC} = 0.20 compared to independent model with w_{AIC} = 0.66). Hence, it is at least possible that breeding conditions could mediate the positive relationship between the brumation period and relative testis size in our relatively small sample of species.

Covariation between different tissues. Since all tissues depend on the same finite resources, their evolution in response to brumation is unlikely to be independent. In pairwise partial correlations controlling for SVL and phylogeny, all tissue masses covaried either positively or were not significantly associated with one another (Fig. S5). As such, our data do not support the expensive tissue (Aiello and Wheeler 1995) or the more general energy trade-off hypotheses (Isler and van Schaik 2006), which predict trade-offs of brain size with the size of the digestive tract or other costly organs, respectively. However, considering that brain size differed from fat tissue, hindlimb muscles and testes in the direction of allometric relationships or their responses to brumation duration, respectively, pairwise correlations may not reveal more complex allocation patterns. To examine the relative investments in these four most informative tissues simultaneously, we partitioned the total body mass into the proportional representation of each of these tissues (and the remaining mass combined as a control) to generate a five-variable compositional dataset (van den Boogaart and Tolosana-Delgado 2013). The combined mass of all four focal tissues scaled proportionately with body size (allometric β = 1.02 [0.96, 1.07], λ = 0.01 [0.00, 0.11]), confirming that each species allocated a size-independent share of its total resources to the four focal tissues combined. How these investments were distributed across these tissues, however, varied considerably between species. In pairwise correlations between the four focal tissues, transformed to centered log ratios (van den Boogaart and

Tolosana-Delgado 2008) and controlling for phylogeny, brain mass covaried negatively with both body fat and testis mass, whilst testis mass was negatively correlated with hindlimb muscle mass but not associated with body fat (Fig. 3A-D; Table S14). To further examine the effect of brumation duration on all five variables simultaneously, we conducted a phylogenetic multivariate regression analysis (Clavel et al. 2015) on the same compositional data, but now transformed to isometric log ratios as recommended for multivariate models (van den Boogaart and Tolosana-Delgado 2008, 2013). Brumation duration had a significant effect on the body composition of frogs (Pillai's trace = 0.33, effect size $\xi^2 = 0.30$, $P = 0.001$). The back-transformed coefficients of body fat (0.23) and testis mass (0.26) were greater, and those of brain mass, hindlimb muscles and the rest of the body were lower (0.16, 0.17, and 0.18, respectively), than the expected coefficient of 0.20 for all five variables if increasing brumation duration were to cause no compositional change (also see Fig. 3).

We further confirmed the negative associations of brain size with both body fat and testis mass, and that between testis mass and hindlimb muscles, in a phylogenetically informed principal component analysis (Revell 2012) (Fig. 3E,F, Table S15). Here, the first three principal components (PC1 to PC3) explained 84.7%, 8.0% and 5.1% of the total variance, respectively. Even though both PC2 and PC3 explained a relatively small proportion of the total variance, they separated the different tissues. PC2 was predominantly loaded by brain size (0.30) and testis mass (−0.44), and PC3 by brain size (0.26) and body fat (−0.31). Both cases thus indicated a negative association between the pairs of traits within the multivariate trait space (Fig. 3F). Further, brumation duration covaried negatively with PC2 ($r = -0.53$, $t_{114} = -6.59$, $P < 0.0001$, $\lambda = 0.74$ [0.49, 0.92]), consistent with a decrease in brain size and increase in testis mass towards longer brumation, but it was not significantly associated with PC3 ($r = -0.12$, $t_{114} = -1.30$, $P = 0.20$, $\lambda = 0.41$ [0.00, 0.70]).

Direct and indirect effects revealed by path analysis. To disentangle the evolutionary links between the relative sizes of costly tissues, and to test the most prominent hypotheses of brain evolution simultaneously in the direct context of brumation, we finally integrated these patterns in a phylogenetic confirmatory path analysis (von Hardenberg and Gonzalez-Voyer 2013; Gonzalez-Voyer and von Hardenberg 2014) based on 28 pre-specified candidate path models (Fig. S6; Table S16). The averaged model (Fig. 4) confirmed the negative effect of prolonged brumation on relative brain size (β = −0.15 [−0.22, −0.07]), which paralleled direct (β = 0.16 [0.07, 0.26]) and indirect positive effects on relative testis size (Fig. S7). These indirect effects on testis evolution were mediated by the relative amount of adipose tissue, which increased with both the brumation period (β = 0.13 [0.06, 0.20]) and the relative size of the digestive tract (β = 0.51 [0.37, 0.65]), and in turn had a positive effect on relative testis size (β = 0.39 [0.18, 0.59]).

Discussion

Our experimentally validated brumation periods and direct measures of the relative sizes of, and changes in, expensive tissues revealed novel insight into the complex and non-independent processes of brain and reproductive evolution in anurans whose ‘cold-blooded’ metabolism is particularly susceptible to environmental seasonality. Species in highly seasonal environments, which go through prolonged inactive periods, had relatively smaller brains than those in more stable climates. By reducing the brain tissue and its associated maintenance costs, brumating species redirected their additional fat reserves to reproduction, possibly due to the shorter breeding season with its socio-ecological consequences.

We demonstrated that anurans inhabiting cooler and more seasonal climates entered and emerged from their inactive state at lower temperatures, indicating increased cold tolerance to

maximize their active period. Yet, with the metabolic rate depending on its thermal environment in ectotherms, low temperatures might render foraging and digestion too inefficient to extend activity beyond certain thresholds (e.g., Riddle, 1909; Fontaine et al., 2018). In response to the varying periods of a positive *versus* negative net energy balance, we found that species with protracted brumation have relatively smaller brains. These results confirm that, unlike birds (Sol, 2009) and some mammals (van Woerden et al., 2010), challenging and unpredictable environmental conditions select for physiological rather than cognitive buffering in anurans (Luo et al., 2017). Supporting a relatively large brain may not be sustainable in the absence of continued resource intake, or large brains could be less tolerant to hypoxic conditions (Sukhum et al., 2016) during brumation. However, selection for relatively larger brains may also simply be stronger in species with a long active (and so short brumation) period owing to extended cognitive benefits such as predator evasion (Kotrschal et al., 2015) or exploitation of better and more diverse food sources (Lefebvre et al., 1997; Jiang et al., 2022).

In pairwise comparisons, the relative sizes of the tissues examined here, including the brain, were generally positively correlated. These results reject both the expensive tissue (Aiello and Wheeler, 1995) and the more general energy trade-off hypotheses (Isler and van Schaik, 2006), which predict trade-offs of brain size with the size of the digestive tract or other costly organs, respectively. This lack of support in anurans aligns with a previous report in mammals (Navarrete et al., 2011) despite their substantially smaller brains and vastly different ecology and physiology, including a lower metabolic rate and largely lacking physiological thermoregulation. When focusing jointly on the four tissues (brain, body fat, testes, hindlimb muscles) that covaried with brumation duration, however, relative brain size covaried negatively with the relative mass of both the fat tissue and testes as predicted by the fat−brain trade-off (Navarrete et al., 2011) or expensive sexual tissue hypotheses (Pitnick et al., 2006),

respectively. Species with an extended brumation period exhibited both a relatively larger amount of body fat in total and a higher degree of its depletion, providing direct evidence to the hypothesis that anurans buffer lean periods by metabolizing stored fat (Luo et al. 2017; Huang et al., 2020). Although adipose tissue may not itself be metabolically expensive, transporting it adds costs to locomotion, particularly when jumping away from predators (Moreno-Rueda et al., 2020) or climbing trees compared to moving horizontally on land or in water (Alexander, 2003; Hanna et al., 2008). Consistent with this notion, arboreal species tended to be leaner compared to (semi)aquatic or terrestrial species (Table S17), controlling for brumation duration and relative brain size, both of which we had shown to covary with body fat (Figs. 2 and 3).

In addition to a relatively smaller brain and relatively larger fat reserves, species with a prolonged brumation period also had relatively smaller hindleg muscles and larger testes. Since anurans move primarily using their hindlegs, the negative relationship between hindleg muscle mass and brumation duration may be linked to more movement during a longer active period, including predator evasion (Marchisin and Anderson, 1978; Liao et al., 2022). Further, one explanation for the negative association between brumation and relatively testis size could be that a shortened active period compacts the breeding season, resulting in denser and more aggregated breeding populations and likely more synchronous mating activity (Wells, 2007), as indicated by our path analysis. These aggregations heighten the likelihood that multiple males attempt to mount the same female simultaneously (Lüpold et al., 2017 p. 20), resulting in more intense male-male competition over fertilization and thus enhanced investments in sperm production (Liao et al., 2018; Lüpold et al., 2020). Our results thus reveal how the environmental variation and physical constraints that determine the species-specific brumation pattern might play a pivotal, albeit previously overlooked, role in shaping the socio-ecological context of

breeding, the mode and degree of sexual selection, and ultimately the evolution of mating systems, broadening Emlen & Oring's (Emlen and Oring, 1977) general predictions.

Species with protracted brumation not only exhibited relative larger testes, but also a greater change in testis size from pre- to post-brumation (i.e., breeding) condition. The testes of seasonally breeding anurans regress after the mating season and regrow before the next (Ogielska and Bartmańska, 2009). Whereas non-brumating species can compensate for the resources invested in testicular recrudescence by energy uptake, those with a short breeding season following a prolonged inactive period depend on the stored fat to regrow their testes before or immediately after emergence from their hibernaculum. Hence, resources are diverted away from the brain and other organs, which may be the case especially in species such as *Brachytarsophrys* spp., in which the fully developed testes combined weigh 12−14 times more than the brain (Data S1).

To integrate all these different patterns and test the most prominent hypotheses of brain evolution simultaneously in the direct context of brumation, we also performed a phylogenetic path analysis. This analysis corroborated the negative effect of brumation duration on relative brain size and revealed both its direct and indirect positive effects on relative testis size. The indirect effects on testis evolution were mediated by the amount of adipose tissue, which itself responded to variation in the inactive period (energetic demand) and the size of the digestive tract (energy uptake). That variation in body fat did not contribute to brain size evolution in this more comprehensive analysis compared to pairwise correlations suggests that the fat−brain trade-off may not be a direct one. Rather, prolonged brumation, and thus short active period, may enhance selection on fat storage for testicular investments in addition to starvation avoidance, while simultaneously selecting for smaller brains (or weakening selection for larger

brains) due to a shifted balance between the cognition-derived fitness benefits and the energetic costs related to brain size (Fig. 3).

A trade-off between brain and testis sizes has been reported for bats (Pitnick et al., 2006), albeit unsupported by a later study (Dechmann and Safi, 2009) or in other mammalian groups (Lemaître et al., 2009). In anurans, the apparent brain−testes trade-off may not be a direct functional one but result indirectly from opposing selection on brain and testis sizes via environmental seasonality and relative durations of the active and inactive periods. The testes may primarily evolve in response to the heightened levels of sperm competition and sperm depletion during the shortened and more synchronized breeding season. The brain, while also responding to sexual selection (Mai et al., 2020), is central to various activities other than mating, including feeding (Lefebvre et al., 1997) or predator avoidance (Kotrschal et al., 2015; Liao et al., 2022) that are themselves subject to climatic conditions and may independently influence brain evolution. In addition, whereas testes can regress to save energy when inactive (Ogielska and Bartmańska, 2009), brain metabolism may be less reducible (Mink et al., 1981), resulting in a different balance between fitness costs and benefits between these organs in relation to seasonality.

In conclusion, our analyses resolve how brumation in anurans, resulting from high environmental seasonality, may constrain the evolution of brain size and affect, directly or through its environmental correlates, resource allocation between costly tissues. These results reveal novel insight into the complex context of brain size evolution in far smaller-brained organisms than those typically studied, and whose ‘cold-blooded’ metabolism is particularly susceptible to environmental fluctuations. Our data also draw attention to the impact that varying brumation periods are likely to have on the operation of sexual selection and mating system evolution by modifying the timing and socio-ecological context of breeding during the

active period. In turn, these factors determine reproductive investments and, via differential resource allocation, may also affect brain evolution. These non-independent selective processes promoting diversification in different traits highlight the need to study the evolutionary trajectory of a given trait such as brain size in the immediate context of both simultaneous investments to other tissues and the species-specific ecology.

Materials and Methods

Sample collection and preparation. Between 2010 and 2020 and as part of concurrent studies, we collected a total of 396 sexually mature males from 116 anuran species (3.41 ± 0.95 males each) in post-brumation breeding condition and an additional 132 adult males from 50 of these species (2.64 ± 0.94 males each) shortly before entering their hibernacula (Data S1 and S2). For each species, we sampled all males at a single location in southern and western China with known longitude, latitude, and elevation (Data S3). Upon transfer to the laboratory, we sacrificed the individuals by single-pithing, measured their snout-vent length (SVL) to the nearest 0.01 mm with calipers and then preserved them in 4% phosphate-buffered formalin for tissue fixation.

After two months of preservation, we weighed each complete specimen to the nearest 0.1 mg using an electronic balance to obtain body mass before dissecting them following a strict protocol. We separately extracted the brain, heart, liver, lungs, kidneys, spleen, digestive tract, testes, limb muscles, and fat stores, cleaned these tissues and immediately weighed them to the nearest 0.1 mg with an electronic balance. We additionally measured the length of the digestive tract to the nearest 0.01 mm using calipers. We excluded emaciated individuals or those exhibiting visible organ pathologies from our analyses.

Environmental seasonality. For each collection site, we retrieved from the 30-year climate history of https://www.meteoblue.com the monthly mean temperature (in °C) and total precipitation (in mm) (Data S3) and used these values to calculate location-specific annual means and coefficients of variation. We also determined the duration of the dry season, P2T, as the number of months, for which the total precipitation was less than twice the mean temperature (Walter, 1971).

Brumation period. One way that anurans can physiologically respond to seasonality is by adjusting their thermal sensitivity and thus brumation period (Wells, 2007), which in turn could directly or indirectly affect the evolution of brain size (Heldstab et al. 2018). Hence, we estimated the brumation period for all 116 species. To this end, we visited the field sites for 30 of our species daily around the expected start and end times of brumation (based on prior experience). For each species, we recorded the dates and temperatures (using a Kobold HND-T105 high-precision thermometer to the nearest 0.1°C) when the last frogs of a given species were seen at the end of their active period (with no further activity detected for at least seven days) and when the first individuals were detected in the spring. For the same 30 species (and using the same individuals as for morphological measurements), we then experimentally simulated brumation using a Q18 temperature-controlled refrigerator in Shenzhen Pioneer (SAST). We gradually lowered and raised the temperature at a rate of 0.5°C/hour and recorded the temperature at which test subjects entered and left the typical brumation posture (i.e., motion-less four-point stance with the nictitating membranes drawn across the eyes). These threshold temperatures were tightly associated with the corresponding field measurements both for the start (r = 0.97, t_{28} = 22.26, P < 0.0001, λ = 0.04 [0.00, 0.47]) and end of the inactive state

(r = 0.98, t_{28} = 28.05, P < 0.0001, λ = 0.04 [0.00, 0.43]). Hence, we assessed the corresponding temperatures for all remaining species in the laboratory and estimated the brumation period based on the daily mean temperatures at the corresponding collection sites as retrieved from Chinese Meteorological Stations (http://www.lishi.tianqi.com) between 2012 and 2016. We defined the brumation period as the number of consecutive days in each year that remained below this threshold. For simplicity we determined the active rather than brumation period, starting with the first day that the mean daily temperature rose above the activity threshold and remained there for at least five consecutive days, and ending with the last day before the temperature dropped below the activity threshold and remained there until the end of the calendar year. The brumation period then represented the difference between the activity period and the total number of days in each calendar year. Across these five years, the measured temperature thresholds yielded highly repeatable species-specific estimates of the number of days below the activity range (R = 0.95 [95%CI: 0.93−0.96]), as determined by the *rpt* function in the *rptR* package (Stoffel et al. 2017) across all 116 species (Fig. S1; Table S1). Further, across the 30 species that were examined both in the lab and the field (see above), these predicted brumation periods were also correlated with the observed brumation periods in the field (r = 0.96, t_{28} = 18.03, P < 0.0001, λ = 0.05 [0.00, 0.48]; Fig. S8A), which themselves were highly repeatable between years within species (R = 0.98 [0.96−0.99]; Table S1).

Based on this data validation, we used for each species the mean brumation period predicted from our experimentally simulated temperature thresholds. However, to test for potential buffering effects of burrowing in the soil relative to the air temperatures reported by the meteorological stations, we also repeated these estimates by using more conservative thermal thresholds. Here, we restricted the putative brumation days to those with a reported air temperature of either 2°C or 4°C below the experimentally derived inactivity thresholds,

simulating prolonged activity by seeking shelter in burrows. The 2°C threshold was based on a pilot study comparing direct measurements of air and burrow temperatures for four different burrows in each of five of our study species (burrow depths: 32.0 ± 3.2 to 121.0 ± 17.8 cm; Fig. S9). Across these species, the burrow-to-air temperature difference reached 1.03 ± 0.35°C to 2.45 ± 0.60°C in measurements around the peak of the brumation period (i.e., early January; Fig. S9). However, since these temporal snapshots were based on sites at relatively low elevation (≤320 m a.s.l.) due to accessibility of burrows during winter, we also used a second, more conservative buffer (4°C below activity range) for comparison. These temperature buffers shortened the predicted brumation periods to a varying degree between species (Fig. S8B); yet the predicted periods covaried strongly between the different temperature thresholds (all $r > 0.90$, $t_{114} > 21.96$, $P < 0.0001$, all $\lambda < 0.01$).

Phylogeny reconstruction. To reconstruct the phylogeny, we obtained the sequences of three nuclear and six mitochondrial genes from GenBank (for accession numbers and sequence coverage see Data S4). The three nuclear genes included the recombination-activating gene 1 (RAG1), rhodopsin (RHOD) and tyrosinase (TYR). The six mitochondrial genes were cytochrome *b* (CYTB), cytochrome oxidase subunit I (COI), NADH dehydrogenase subunits 2 and 4 (ND2 and ND4), and the large and small subunits of the mitochondrial ribosome genes (12S/16S; omitting the adjacent tRNAs as they were difficult to align and represented only a small amount of data). We aligned the sequences by multi-sequence alignment (MUSCLE) in MEGA v.10.2.2 (Tamura et al., 2013) before comparing possible nucleotide substitution models. The best substitution model, as determined by the function *modelTest()* in the R (R Core Team, 2022) package *phangorn* (Schliep, 2011) based on the corrected Akaike Information Criterion, AICc, was GTR+Γ+I for all genes except RHOD, for which HKY+Γ had stronger support.

Using BEAUTi and BEAST v.1.10.4 (Suchard et al., 2018), we then constructed the phylogeny with unlinked substitution models, a relaxed uncorrelated log-normal clock, a Yule speciation process, and the best-supported nucleotide substitution models. We omitted time calibration due to a lack of fossil dates. We ran the Markov Chain Monte Carlo (MCMC) simulation for 55 million generations while sampling every 5,000th tree with a 10% burn-in. Most effective sample size (ESS) values by far exceeded 375 (i.e. all well above the recommended threshold of 200) for all but two tree statistics in the program Tracer v.1.7.2 (Rambaut et al., 2018), thus indicating satisfying convergence of the Bayesian chain and adequate model mixing. Finally, we generated a maximum clade credibility tree with mean node heights and a 10% burn-in using TreeAnnotator v.1.10.4 (Suchard et al., 2018), presented in Fig. S10.

Breeding conditions. To test if a prolonged brumation period reduces the time available for reproduction, thereby changing the level of competition over mates and fertilizations (Lüpold et al., 2017), we extracted the start and end dates of the breeding season from our field notes of concurrent studies on species-specific life histories. These data were available for 43 of our species (Data S3). We used recorded dates when the first and last clutches were observed in focal ponds as a proxy of mating activity, given that males release their sperm during oviposition in these external fertilizers. For each species, dates from at least two years were combined and averaged to obtain the mean duration of the breeding season.

We further recorded whether dense mating aggregations are typically observed in these species. We have previously shown that larger mating clusters, with multiple males clasping the same females, have a significant effect on the evolution of testis size due to the resulting competition among sperm for fertilization (Lüpold et al., 2017). Here, we had no detailed data on

the sizes of aggregations and so were only able to code the typical presence or absence of aggregations as a binary variable (Data S3).

Finally, we used our direct estimates of species-specific population densities from our previous study (Lüpold et al., 2017) to test whether a shorter breeding season results in denser breeding populations. Although population density is a more direct measure than the occurrence of aggregations, such data were available for only eight of our species, each based on multiple populations per species (Lüpold et al., 2017). All these data were not necessarily derived from the same years or populations of our main dataset, but given the within-species repeatability in breeding populations (Lüpold et al., 2017) and in the duration of the breeding season (R = 0.88 [0.79−0.93]; Table S1), these differences should be relatively small compared to the interspecific variation and mostly introduce random noise.

Data analyses. We conducted all statistical analyses in R v.4.2.0 (R Core Team, 2022), using log-transformed data for all phenotypic traits, and for the CV in temperature among the ecological variables. To account for non-independence of data due to common ancestry (Pagel, 1999; Freckleton et al., 2002), we conducted phylogenetic generalized least-squares (PGLS) or phylogenetic logistic regressions (e.g., for occurrence of breeding aggregations), using the R package *phylolm* (Ho and Ané, 2014) and our reconstructed phylogeny. To account for variation around the species means, we bootstrapped for each model (at 100 fitted replicates) the standardized regression coefficients along with the phylogenetic scaling parameter λ and calculated their corresponding 95% confidence intervals. The λ values indicate phylogenetic independence near zero and strong phylogenetic dependence near one (Freckleton et al., 2002).

Unless stated otherwise, all PGLS models focusing on the relative mass of tissues as the response included snout-vent length (SVL) as a covariate in addition to the focal predictor variable(s). We chose SVL instead of body mass because it is the commonly used measure of body size in anurans and independent of seasonal fluctuations in tissues such as body fat, testes, or limb muscles. One exception, however, was the analysis of phylogenetically informed allometric relationships, for which we cubed SVL such that a slope of 1 equaled unity (isometry). For these allometric relationships we calculated ordinary (generalized) least-squares rather than reduced major-axis regressions, because their greater sensitivity to changes in the steepness, but lower sensitivity to changes in scatter, capture allometric slopes more adequately (Kilmer and Rodríguez, 2017).

To examine the covariation between different tissues across species, we first calculated pairwise partial correlations controlling for SVL and phylogeny. To this end, we calculated the phylogenetic trait variance-covariance matrix between the pairs of focal variable and SVL using the function *phyl.vcv()* in *phytools* (Revell, 2012) with λ = 1 (i.e. Brownian motion), which we then scaled into a correlation matrix using *cov2cor()* in the *stats* package (R Core Team, 2022). Using the resulting correlation coefficients r_{xy}, r_{xz}, and r_{yz}, respectively, we then calculated the partial correlation coefficient $r_{xy.z}$ between the *x* and *y* variables of interest while accounting for SVL (*z*) following Crawley's (Crawley 2007) equation: $r_{xy.z} = \frac{r_{xy} - r_{xz} r_{yz}}{\sqrt{(1-r_{xz}^2)(1-r_{yz}^2)}}$, with the associated *t*-statistics and 95% confidence intervals converted using standard conversion ($t = r\sqrt{\frac{df}{1-r^2}}$) and the package *effectsize* (Ben-Shachar et al., 2020), respectively.

Since pairwise correlations do not necessarily capture more complex, multivariate allocation patterns, we used two additional approaches to explore how tissue sizes varied relative to others. Here, we focused on those four tissues that covaried with brumation duration or deviated

from proportionate scaling with body size: brain, body fat, testes, and hindlimb muscles. Using the function *acomp()* in the R package *compositions* (van den Boogaart and Tolosana-Delgado 2008), we partitioned total body mass of each species into a five-variable Aitchison composition in a logistic geometry (van den Boogaart and Tolosana-Delgado, 2013), consisting of the proportional representation of the four focal tissues and the combined remaining body mass. Since the focal tissues constituted a size-independent fraction of the total body, the closed composition of this combined mass should be unbiased relative to body size but can instead reveal differential contributions of these tissues to their total in a multivariate context (Aitchison, 1982; Muldowney et al., 2001; van den Boogaart and Tolosana-Delgado, 2013). For phylogenetic correlations between these variables following the description above, we used centered log ratios obtained by the function *clr()* in the same package, which maintains the original variable structure. However, owing to the reliance on a full rank of the covariance in multivariate analyses, we used the *ilr()* function to project the D-part composition isometrically to a $D-1$ dimensional simplex (Aitchison, 1982), essentially representing the log ratios between the D parts. This multivariate object we subjected to a phylogenetic multivariate regression against brumation duration using the functions *mvgls()* and *manova.gls()* in the package *mvMORPH* (Clavel et al., 2015). For interpretation in the context of the original variable space, we back-transformed the coefficients using the *ilrInv()* function in *compositions*. In addition to this compositional data analysis, we also performed a phylogenetically informed principal component analysis on the same focal tissues as log-transformed species means, using the *phyl.pca()* function of the package *phytools* (Revell, 2012). Here, we primarily focused on the directions of the loading vectors relative to the principal components and one another to glean information on the correlations between the original variables in the principal component space.

Finally, we performed phylogenetic confirmatory path analyses (von Hardenberg and Gonzalez-Voyer 2013; Gonzalez-Voyer and von Hardenberg, 2014) based on pre-specified candidate structural equation models, either to explore the direct and indirect effects of climatic variables on the duration of the breeding season or formation of breeding aggregations, or to better disentangle the different interrelationships between traits that could ultimately mediate the effect of brumation duration on brain and reproductive evolution. Using the R package *phylopath* (van der Bijl 2018), we examined the conditional independencies of each model, ranked all candidate models based on their *C*-statistic Information Criterion (CICc), and then averaged those with ΔCICc ≤ 2 from the top model (von Hardenberg and Gonzalez-Voyer, 2013). To avoid overparameterization of the path analysis on breeding parameters, which was based on only 43 species, we did not include testis mass (and so necessarily also SVL to control for body size) as additional variables in the same path models. However, we separately tested for correlated evolution using directional tests of trait evolution (Pagel, 1994; Revell, 2012). Based on (the weight of) the Akaike Information Criterion (AIC), we tested if changes in relative testis size and breeding parameters, respectively, were unilaterally dependent, mutually dependent, or independent (Pagel, 1994), using the *fitPagel()* function in the *phytools* package (Revell, 2012) with “fitDiscrete” as the optimization method and allowing all rates to differ (i.e., “ARD” model). Since these analyses are based on evolutionary transitions between binary states, we considered positive residuals of a log-log regression between testis mass and SVL as ‘relatively large testes’ and negative residuals as ‘relatively small testes.’ For the duration of the breeding season, we similarly split the distribution based on the mean duration, whereas aggregation formation was already coded as present or absent.

Acknowledgments

We thank K. Isler and C. van Schaik for helpful comments on an earlier draft of our manuscript, and J. Santin for insightful discussions. Financial support was provided by the National Natural Sciences Foundation of China (31970393 to WBL) and by the Swiss National Science Foundation (PP00P3_170669 and PP00P3_202662 to SL).

Data and code availability

All data and R scripts related to this article are deposited in the Figshare Data Repository (doi: 10.6084/m9.figshare.21078052).

Ethics statement

The specimens used in this study were collected with permission from the China West Normal University Ethical Committee for Animal Experiments (CWNU-202001), and the experimental protocols adhered to the current laws of China concerning animal experimentation.

References

Aiello LC, Wheeler P. 1995. The expensive-tissue hypothesis: the brain and the digestive system in human and primate evolution. Current Anthropology 36:199–221.

Aitchison J. 1982. The statistical analysis of compositional data. Journal of The Royal Statistical Society Series B-statistical Methodology 44:139–177.

Alexander RM. 2003. Principles of Animal Locomotion. Princeton University Press, Princeton, NJ.

Allman J, McLaughlin T, Hakeem A. 1993. Brain weight and life-span in primate species. Proceedings of the National Academy of Sciences of the United States of America 90:118–122.

Ben-Shachar MS, Lüdecke D, Makowski D. 2020. effectsize: Estimation of effect size indices and standardized parameters. Journal of Open Source Software 5:2815.

Benson-Amram, S., B. Dantzer, G. Stricker, E. M. Swanson, and K. E. Holekamp. 2016. Brain size predicts problem-solving ability in mammalian carnivores. Proceedings of the National Academy of Sciences of the United States of America 113:2532–2537.

Buzatto BA, Roberts JD, Simmons LW. 2015. Sperm competition and the evolution of precopulatory weapons: Increasing male density promotes sperm competition and reduces selection on arm strength in a chorusing frog. Evolution 69:2613–2624.

Cerri S, Bottiroli G, Bottone MG, Barni S, Bernocchi G. 2009. Cell proliferation and death in the brain of active and hibernating frogs. Journal of Anatomy 215:124–131.

Clavel, J., G. Escarguel, and G. Merceron. 2015. mvMORPH: an R package for fitting multivariate evolutionary models to morphometric data. Methods Ecology and Evolution 6:1311–1319.

Crawley MJ. 2007. The R Book. John Wiley & Sons Ltd, West Sussex.

Dechmann D K N, Safi K. 2009. Comparative studies of brain evolution: a critical insight from the Chiroptera. Biological Reviews 84:161–172.

Emlen ST, Oring L W. 1977. Ecology, sexual selection, and evolution of mating systems. Science 197:215–223.

Fei L, Ye CY. 2001. The Colour Handbook of the Amphibians of Sichuan. China Forestry Publishing House, Chengdu.

Fontaine S S, Novarro A J, Kohl K D. 2018. Environmental temperature alters the digestive performance and gut microbiota of a terrestrial amphibian. Journal of Experimental Biology 221:jeb187559.

Freckleton R P, Harvey P H, Pagel M. 2002. Phylogenetic analysis and comparative data: a test and review of evidence. American Naturalist 160:712–726.

Gonzalez-Voyer A, von Hardenberg A. 2014. An introduction to phylogenetic path analysis. Pp. 201–229 *in* L Z Garamszegi, ed. Modern Phylogenetic Comparative Methods and their Application in Evolutionary Biology. Springer-Verlag, Berlin.

Hanna J B, Schmitt D, and Griffin T M. 2008. The energetic cost of climbing in primates. Science 320:898–898.

Heldstab S A, Isler K, van Schaik C P. 2018. Hibernation constrains brain size evolution in mammals. Journal of Evolutionary Biology 31:1582–1588.

Heldstab S A, van Schaik C P, Isler K. 2016. Being fat and smart: A comparative analysis of the fat-brain trade-off in mammals. Journal Of Human Evolution 100:25–34.

Ho L S T, Ané C. 2014. A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. Systematic Biology 63:397–408.

Holenweg A-K, and Reyer H-U. 2000. Hibernation behavior of *Rana lessonae* and *R. esculenta* in their natural habitat. Oecologia 123:41–47.

Huang Y, Mai C L, Liao W B, Kotrschal A. 2020. Body mass variation is negatively associated with brain size: Evidence for the fat-brain trade-off in anurans. Evolution 74:1551–1557.

Isler K. 2011. Energetic trade-offs between brain size and offspring production: Marsupials confirm a general mammalian pattern. BioEssays 33:173–179.

Isler K, van Schaik C P. 2006. Costs of encephalization: the energy trade-off hypothesis tested on birds. Journal of Human Evolution 51:228–243.

Isler K, van Schaik C P. 2009. The expensive brain: A framework for explaining evolutionary changes in brain size. Journal of Human Evolution 57:392–400.

Jerison H J. 1973. Evolution of the Brain and Intelligence. Academic Press, New York.

Jiang Y, Luan X, Liao WB 2023. Anuran brain size predicts food availability-driven population density. Science China Life Sciences 66：413-415.

Kardong K V. 2019. Vertebrates Comparative Anatomy, Function, Evolution. 8th edition. McGraw-Hill, New York.

Kilmer J T, Rodríguez R L. 2017. Ordinary least squares regression is indicated for studies of allometry. Journal of Evolutionary Biology 30:4–12.

Kotrschal A, Buechel S D, Zala S M, Corral-Lopez A, Penn D J, Kolm N. 2015. Brain size affects female but not male survival under predation threat. Ecology Letters 18:646–652.

Lefebvre L, Reader S M, Sol D. 2004. Brains, innovations and evolution in birds and primates. Brain Behavior and Evolution 63:233–246.

Lefebvre L, Whittle P, Lascaris E, Finkelstein A. 1997. Feeding innovations and forebrain size in birds. Animal Behaviour 53:549–560.

Lemaître J-F, Ramm S A, Barton R A, Stockley P. 2009. Sperm competition and brain size evolution in mammals. Journal of Evolutionary Biology 22:2215–2221.

Liao W B, Huang Y, Zeng Y, Zhong M J, Luo Y, Lüpold S. 2018. Ejaculate evolution in external fertilizers: Influenced by sperm competition or sperm limitation? Evolution 72:4–17.

Liao W B, Jiang Y, Li D Y, Jin L, Zhong M J, Qi Y, Lüpold S, Kotrschal A. 2022. Cognition contra camouflage: how the brain mediates predator-driven crypsis evolution. Science Advances 8:eabq1878.

Lukas W D, Campbell B C. 2000. Evolutionary and ecological aspects of early brain malnutrition in humans. Human Nature 11:1–26.

Luo Y, Zhong M J, Huang Y, Li F, Liao W B, Kotrschal A. 2017. Seasonality and brain size are negatively associated in frogs: Evidence for the expensive brain framework. Scientific Reports 7:1–9.

Lüpold S, de Boer R A, Evans J P, Tomkins J L, Fitzpatrick J L. 2020. How sperm competition shapes the evolution of testes and sperm: A meta-analysis. Philosophical Transactions of the Royal Society B: Biological Sciences 375:20200064.

Lüpold S, Jin L, Liao W B. 2017. Population density and structure drive differential investment in pre- and postmating sexual traits in frogs. Evolution 71:1686–1699.

Lüpold S, Tomkins J L, Simmons L W, Fitzpatrick J L. 2014. Female monopolization mediates the relationship between pre- and postcopulatory sexual traits. Nature Communications 5:3184.

Mai C L, Liao W B, Lüpold S, Kotrschal A. 2020. Relative brain size is predicted by the intensity of intrasexual competition in frogs. American Naturalist 196:169–179.

Marchisin A, Anderson J D. 1978. Strategies employed by frogs and toads (Amphibia, Anura) to avoid predation by snakes (Reptilia, Serpentes). Journal of Herpetology 12:151–155.

Mink J.W, Blumenschine R J, Adams D B. 1981. Ratio of central nervous system to body metabolism in vertebrates: its constancy and functional basis. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology 241:R203–R212.

Moreno-Rueda G, Requena-Blanco A, Zamora-Camacho F J, Comas M, Pascual G. 2020. Morphological determinants of jumping performance in the Iberian green frog. Current Zoology 66:417–424.

Muldowney D, Connolly J, Keane M G. 2001. Compositional data analysis in the study of carcass composition of beef cattle. Livestock Production Science 67:241–251.

Navarrete A, van Schaik C P, Isler K. 2011. Energetics and the evolution of human brain size. Nature 480:91–93.

Niu Y, Chen Q, Storey K B, Teng L, Li X, Xu T, Zhang H. 2022. Physiological ecology of winter hibernation by the high-altitude frog *Nanorana parkeri*. Physiol. Biochem. Zool. 95:201–211. The University of Chicago Press.

Ogielska M, Bartmańska J. 2009. Spermatogenesis and male reproductive system in Amphibia—Anura. Pp. 34–99 *in* M. Ogielska, ed. Reproduction of Amphibians. Science Publishers, Enfield, NH.

Pagel M. 1994. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. Proceedings of the Royal Society B: Biological Sciences 255:37–45.

Pagel M. 1999. Inferring the historical patterns of biological evolution. Nature 401:877–884.

Pinder A W, Storey K B, Ultsch G R. 1992. Estivation and hibernation. Pp. 250–274 *in* M. E. Feder and W. W. Burggren, eds. Environmental Physiology of the Amphibians. University of Chicago Press, Chicago, IL.

Pitnick S, Jones K E, Wilkinson G S. 2006. Mating system and brain size in bats. Proceedings of the Royal Society B: Biological Sciences 273:719–724.

Place N J, Veloso C, Visser G H, Kenagy G J. 2002. Energy expenditure and testosterone in free-living male yellow-pine chipmunks. Journal of Experimental Zoology 292:460–467.

Psenner H. 1957. Neues vom Murmeltier *Marmota marmota*. Säugetierkd. Mitteilungen 5:4–10.

R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rambaut A, Drummond A J, Xie D, Baele G, Suchard M A. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Systematic Biology 67:901–904.

Reader S M, Laland K N. 2002. Social intelligence, innovation, and enhanced brain size in primates. Proceedings of the National Academy of Sciences of the United States of America 99:4436–4441.

Revell L J. 2012. phytools: An R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution 3:217–223.

Riddle O. 1909. The rate of digestion in cold-blooded vertebrates. The influence of season and temperature. American Journal of Physiology 24:447–458.

Ruf, T., and F. Geiser. 2015. Daily torpor and hibernation in birds and mammals. Biological Reviews 90:891–926.

Schliep K P. 2011. phangorn: Phylogenetic analysis in R. Bioinformatics 27:592–593.

Sibly R M. 1981. Strategies of digestion and defecation. Pp. 109–139 *in* C. R. Townsend and P. Calow, eds. Physiological Ecology: An Evolutionary Approach to Resource Use. Blackwell Scientific Publications, Oxford, UK.

Sol D. 2009. The cognitive-buffer hypothesis for the evolution of large brains. P. *in* R. Dukas and J. R. Ratcliff, eds. Cognitive Ecology II. University of Chicago Press, Chicago, IL.

Sol D, Duncan R P, Blackburn T M, Cassey P, Lefebvre L. 2005. Big brains, enhanced cognition, and response of birds to novel environments. Proceedings of the National Academy of Sciences of the United States of America 102:5460–5465.

Staples J F. 2016. Metabolic flexibility: hibernation, torpor, and estivation. Comprehensive Physiology 6:737–771.

Stinner J, Zarlinga N, Orcutt S. 1994. Overwintering behavior of adult bullfrogs, *Rana catesbeiana*, in northeastern Ohio. Ohio Journal of Science 94.

Stoffel M A, Nakagawa S, Schielzeth H. 2017. rptR: repeatability estimation and variance decomposition by generalized linear mixed-effects models. Methods Ecology and Evolution 8:1639–1644.

Suchard M A, Lemey P, Baele G, Ayres D L, Drummond A J, Rambaut A. 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evolution 4:vey016.

Sukhum K V, Freiler M K, Wang R, Carlson B A. 2016. The costs of a big brain: extreme encephalization results in higher energetic demand and reduced hypoxia tolerance in weakly electric African fishes. Proceedings of the Royal Society B: Biological Sciences 283:20162157.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution 30:2725–2729.

Tattersall G J, Ultsch G R. 2008. Physiological ecology of aquatic overwintering in ranid frogs. Biological Reviews 83:119–140.

Törnqvist L, Vartia P, Vartia Y O. 1985. How should relative changes be measured? The American Statistician 39:43–46.

Vahed K, Parker D J. 2012. The evolution of large testes: Sperm competition or male mating rate? Ethology 118:107–117.

van den Boogaart K G, Tolosana-Delgado R. 2013. Analyzing Compositional Data with R. Springer, Berlin.

van den Boogaart K G, Tolosana-Delgado R. 2008. “compositions”: A unified R package to analyze compositional data. Computers & Geosciences 34:320–338.

van der Bijl, W. 2018. phylopath: Easy phylogenetic path analysis in R. PeerJ 6:e4718.

van Gelder J J, Olders J H J, Bosch J W G, Starmans P W. 1986. Behaviour and body temperature of hibernating common toads *Bufo bufo*. Ecography 9:225–228.

van Woerden J T, van Schaik C P, Isler K. 2010. Effects of seasonality on brain size evolution: evidence from strepsirrhine primates. American Naturalist 176:758–767.

van Woerden J T, Willems E P, Van Schaik C P, Isler K. 2012. Large brains buffer energetic effects of seasonal habitats in catarrhine primates. Evolution 66:191–199.

von Hardenberg A, Gonzalez-Voyer A. 2013. Disentangling evolutionary cause-effect relationships with phylogenetic confirmatory path analysis. Evolution 67:378–387.

Walter H. 1971. Ecology of Tropical and Subtropical Vegetation. Oliver and Boyd, Edinburgh.

Wells K D. 2007. The Ecology and Behavior of Amphibians. University of Chicago Press, Chicago, IL.

Wells K D. 1977. The social behaviour of anuran amphibians. Animal Behavior 25:666–693.

Figures

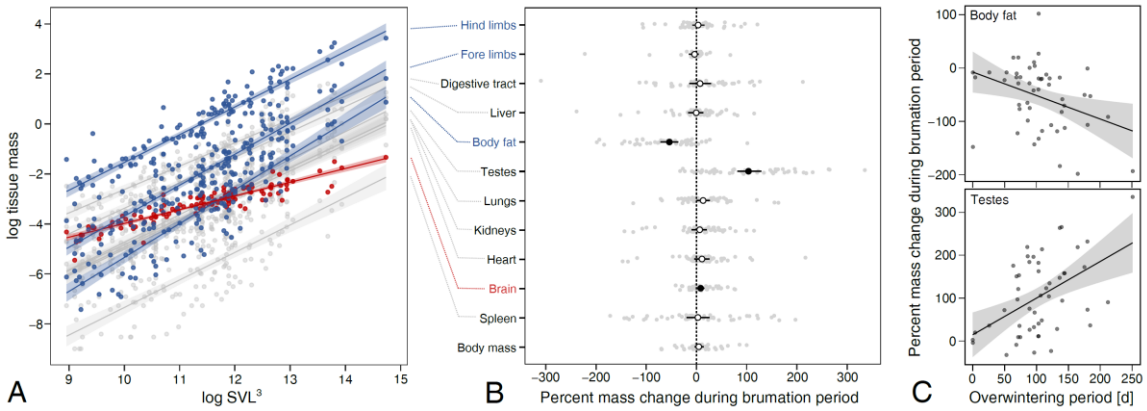

Figure 1. Allometric and seasonal variation in the mass of 11 organs and tissues. **A:** Allometric slopes between the mass of each tissue and cubed snout-vent length (SVL^3) so that proportionate scaling follows a slope of 1. Each point represents a species-specific mean value in breeding condition (*N* = 116). Relationships deviating from proportionate scaling (based on bootstrapped 95% confidence intervals) are highlighted in blue (steeper than unity) or red (shallower than unity). **B:** Mean percent change with 95% confidence interval for body mass and each individual tissue of 50 anuran species with data from both shortly before and after brumation (= breeding), based on absolute tissue masses between stages and log-transformed to maintain symmetry and additivity (Törnqvist et al. 1985): log(post-brumation / pre-brumation) × 100. The transparent grey dots depict the species-specific values. **C:** The effect of the brumation period on the percent mass change for the two tissues with a substantial seasonal variation (see panel b): body fat (top) and testes (bottom). Each point indicates a species.

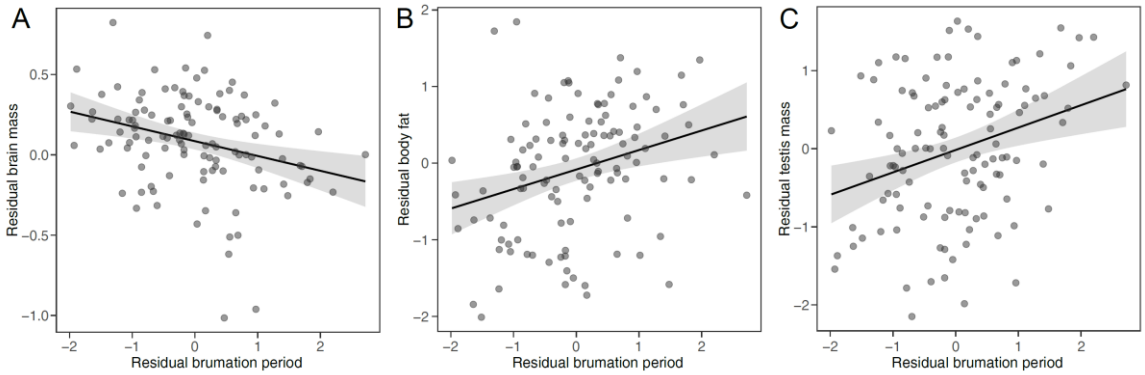

Figure 2. Effects of brumation duration on the relative tissue sizes. Relationships between brumation duration and the relative mass of the brain (**A**), body fat (**B**), and testes (**C**) across males of 116 anuran species in breeding (post-brumation) condition. All axes are controlled for the snout-vent length and phylogeny.

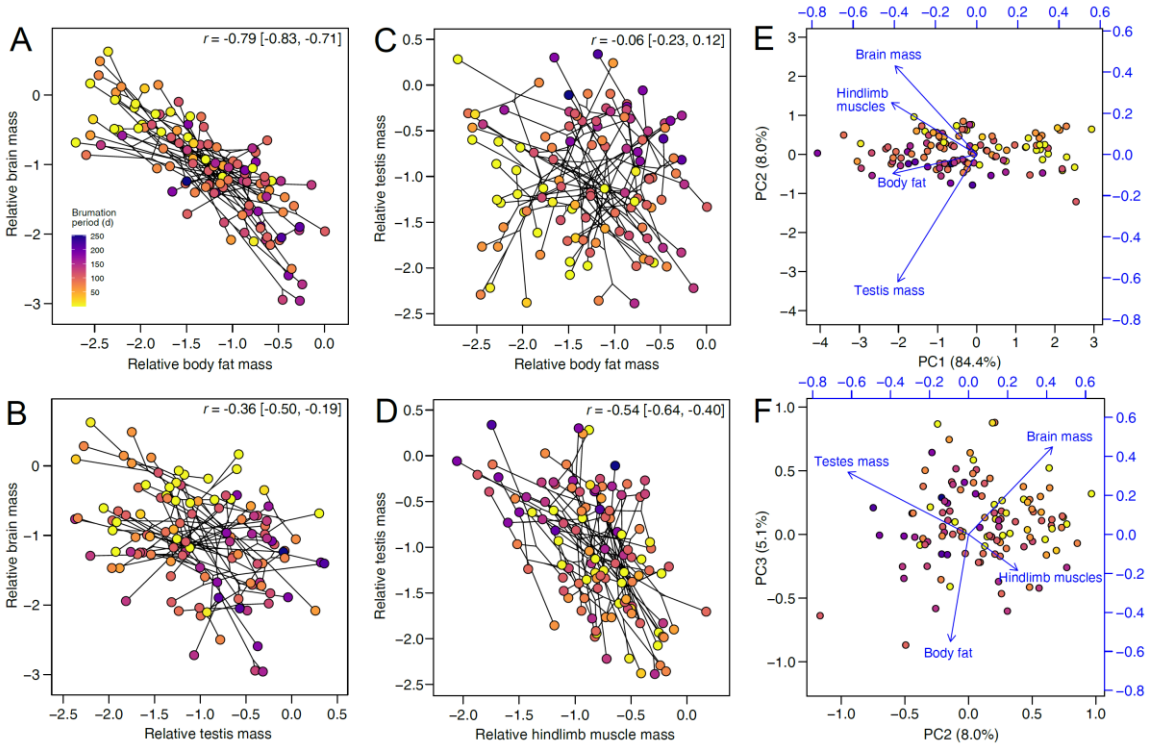

Figure 3. Effects of brumation duration on the relative tissue sizes. Panels **A-D** depict the phylogenetic correlations (shown as phylomorphospace plots (Revell 2012)) between the relative masses of (**A**) brain and body fat, (**B**) brain and testes, (**C**) testes and body fat, and (**D**) testes and hindlimb muscles, respectively, across the 116 species (results in Table S14). The relative tissue masses represent the centered log ratios of the compositional data, and the lines connect the nodes of the underlying phylogeny, indicating that phenotypic correlations are not simply the result of phylogenetic clustering. The correlation coefficients and 95% confidence intervals are indicated. The loadings from a phylogenetic principal component analysis (Revell 2012) on the same variables are also mapped as vectors onto biplots between (**E**) the first and second or (**F**) the second and third principal components. In all panels, the point colors reflect the species-specific brumation periods (see legend in panel **A**). Generally, where brumation was

relatively shorter or absent, species also tended to have relatively larger brains, less body fat and smaller testes, respectively, consistent with the univariate analyses (Fig. 2).

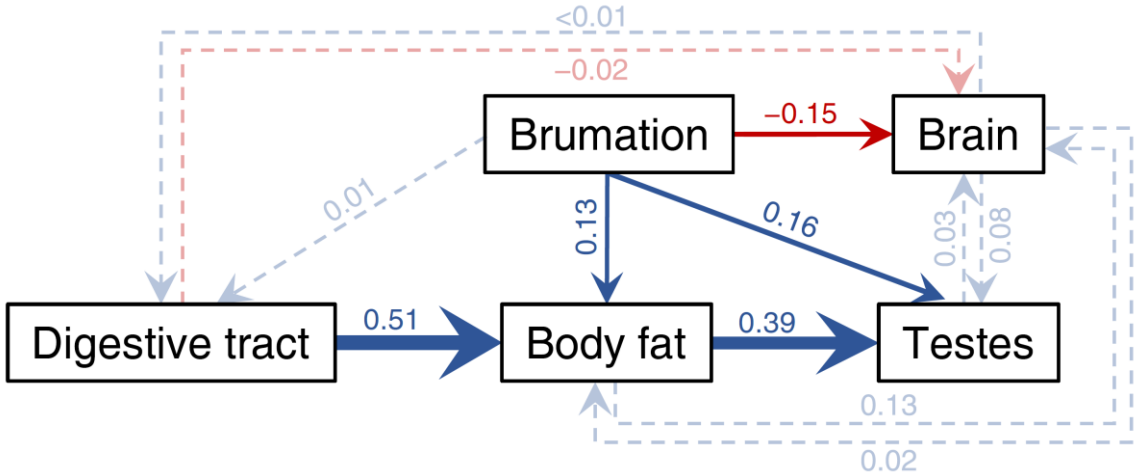

Figure 4. Results of the averaged phylogenetic path model. Visual representation of the average phylogenetic path model across 116 anuran species. Arrows reflect the direction of the path, with their widths being proportional to the standardized regression coefficients and colors indicating the sign (blue = positive, red = negative). Paths with 95% confidence intervals excluding 0 (i.e., arrows highly probable) are drawn as solid arrows, all others as dashed, semi-transparent arrows. For simplicity and to avoid over-parameterization, other organs were omitted in path models as they showed little covariation with brumation duration or brain size. All phenotypic traits were log-transformed, and all variables were controlled for body size via additional paths from log SVL. Although SVL had a strong effect on all variables (all $\beta > 0.37$), its thick blue arrows to each box are omitted in this figure only for visual clarity, but all path coefficients are presented with their 95%CI in Fig. S7, with further details in Fig. S6 and Table S16.