INDIVIDUAL VARIATION IN HEAT SUBSTITUTION

CAROLINE MALONEY

Thesis submitted to the University of Ottawa
partial fulfillment of the requirements of the
MSc degree at the Ottawa-Carleton Institute of Biology

Department of Biology
Faculty of Science
University of Ottawa

Abstract

Endotherms living in cold environments must pay the energetic cost of maintaining a high core body temperature. This cost can be potentially alleviated by an important yet often overlooked mechanism: “activity-thermoregulatory heat substitution” (i.e., the use of the heat generated by active skeletal muscles to replace heat that would have been generated by thermogenesis). While substitution has been documented numerous times, the extent of individual variation in substitution has never been quantified. I used a respirometry cage system to repeatedly measure substitution through the concomitant monitoring of metabolic rate (MR) and locomotor activity in 46 female white-footed mice (*Peromyscus leucopus)* in neutral and cold ambient temperatures. I took a total of 117 measures of substitution by quantifying the difference in the slope of the relationship between MR and locomotor activity speed at two different ambient temperatures. Consistency repeatability (±se) of substitution was 0.313±0.131 – hence, about a third of the variation in substitution occurs at the among-individual level. Including key morphological traits such as trunk surface area, tail mass, heart mass, and body length accounted for the majority of the among-individual variation, suggesting that I have successfully identified traits underlying individual differences in substitution. Overall, my results show that substitution is repeatable and hence might potentially be subject to selection. Future studies should test if substitution conveys fitness advantages directly (by providing energetically cheaper activity which in turn can be utilized for reproduction), or indirectly (i.e., driven by individual differences in morphology). Future studies should also test if there is a trade-off between substitution and dry heat transfer (a thermoregulatory mechanism essential for preventing hyperthermia).

Résumé

Les endothermes vivant dans des environnements froids doivent payer le coût énergétique du maintien d'une température corporelle élevée. Ce coût peut être allégé via un mécanisme potentiellement important mais souvent négligé : la « substitution de chaleur » (c.-à-d. l'utilisation de la chaleur générée par l'activation des muscles squelettiques pour remplacer la chaleur qui aurait été générée par la thermogenèse). Alors que le phénomène de substitution a été largement documenté, l'étendue de la variation individuelle en substitution n'a jamais été quantifiée. J'ai utilisé un système de respirométrie en cage pour mesurer à plusieurs reprises la substitution grâce au suivi simultané du taux métabolique (TM) et de l'activité locomotrice chez 46 souris femelles à pattes blanches (*Peromyscus leucopus*) à des températures ambiantes neutres et froides. J'ai pris un total de 117 mesures de substitution en quantifiant la différence de pente de la relation entre la TM et la vitesse d'activité locomotrice à deux températures ambiantes différentes. La répétabilité (±se) globale de la substitution était de 0.313 ± 0.131 - par conséquent, environ un tiers de la variation en substitution se produit au niveau interindividuel. L'inclusion de traits morphologiques tels que la surface du tronc, la masse de la queue, la masse du coeur et la longueur du corps expliquait la majorité des variations entre les individus, suggérant que j'ai réussi à identifier les traits sous-jacents aux différences individuelles en substitution. Dans l'ensemble, mes résultats montrent que la substitution est répétable et peut donc potentiellement faire l'objet d'une sélection. Les études futures devraient tester si la substitution apporte des avantages de fitness directement (en fournissant une activité énergétiquement moins chère qui à son tour peut être utilisée pour la reproduction), ou indirectement (c'est-à-dire, entraînée par des différences individuelles en morphologie). Les études futures devraient également vérifier s'il existe un compromis entre la substitution et le

transfert de chaleur via le rayonnement et la convection (un mécanisme de thermorégulation essentiel pour prévenir l'hyperthermie).

Acknowledgements

First, I would like to thank my supervisor, Dr. Vincent Careau, for accepting me into his lab and allowing me to work on this project. I greatly appreciate his support, especially considering the pandemic and the difficulties that emerged. I am grateful for what Vincent has taught me during the course of this thesis and hope to move ahead implementing and improving on what I have learned. As well, this project would not have been possible without our lab's funding sources.

I would also like to thank my committee members Dr. Heath MacMillan (Carleton University) and Dr. Emily Standen (University of Ottawa) for their guidance and feedback on this project. I appreciate the insights they had and I believe it has positively shaped my thesis.

I would like to thank my fiancé, Xiaoli, for sharing a home office with me, bringing me cups of coffee and dropping me off on campus on chilly days. I could not have done this without him. It would be remiss for me to not mention the support of the other members of my lab, whether they helped review my code, read my drafts, or even just listened to me during trying times. On campus, the veterinary staff of the Vanier animal care facility did an excellent job of maintaining our mouse population, training me in blood-drawing technique and were available at a moment's notice to help.

Finally, I acknowledge that my research at the University of Ottawa was conducted on the unceded territory of the Algonquin Anishinabe people. I extend my respect to them as the caretakers and keepers of this land.

Table of Contents

Abstract ii
Résumé iii
Acknowledgements v
Table of Contents vi
Introduction 1
Activity-Thermoregulatory Heat Substitution 1
Individual Variation & Repeatability 3
Substitution and Covariates 5
Objectives 7
Methods 7
Ethics 7
Study Animals 7
Experimental Design 8
Respirometry 9
Torpor 11
Tissue Sampling 11
Statistical Analysis 12
Caveat: error in raw MR measurements 15
Results 16
Repeatability 16
Repeatability of Substitution 16
Covariates 17
Discussion 17
Conductivity 18
Body Length 19
Hematocrit 21
Organ & Muscle Mass 21
Experimental Constraints 23
Future Directions 26
Conclusion 28
References 29
Tables 37
Figures 42

Introduction

Activity-Thermoregulatory Heat Substitution

Cold environments represent a unique challenge for endothermic animals. Endotherms can only maintain their high core body temperature without any homeostatic or behavioural responses within a small range of ambient temperatures (Hedrick and Hillman, 2016). When the ambient temperature (T_a) is below an endotherm's thermoneutral zone (TNZ), it must engage in thermogenesis to maintain its body temperature. This represents an additional energetic cost to animals, who either increase their energy intake (Arnold et al., 2006) or reduce the cost of thermogenesis (Hetem et al., 2016; Levesque et al., 2016) to avoid the consequences of energy shortfalls. Endotherms can reduce the energetic cost of thermoregulation through various physiological and behavioural mechanisms. Many birds and mammals save energy by using torpor—a state of decreased physiological activity accompanied by lower core body temperature—either in a daily or seasonal pattern (also known as hibernation) (Geiser, 2020). Endotherms can also avoid cold exposure, such as red squirrels in boreal forests who stay in their well-insulated nests instead of foraging on the coldest days (Menzies et al., 2020).

There are situations in which it is impossible for endotherms to use torpor and cold avoidance behavioural mechanisms (Maresh et al., 2015). In fact, it has been estimated that most endotherms routinely experience T_a that are lower than their TNZ (Humphries and Careau, 2011). In these circumstances, a potentially important yet understudied energy-saving mechanism is the activity-thermoregulatory heat substitution (hereafter, substitution). Substitution occurs when heat generated from physical activity (work) replaces the heat that a resting endotherm below the TNZ would have had to produce through thermogenesis to maintain

its body temperature (Lefèvre and Auget, 1931). The mechanism by which substitution occurs is thought to be the diversion of heat from active skeletal muscle to the body's core (Liwanag et al., 2009). The central benefit of substitution is to reduce or outright eliminate the cost of activity (COA) in the cold for an endotherm, as energy that would be typically lost as heat efflux is used in lieu of thermogenesis (Humphries and Careau, 2011). The saved energy can, in turn, be potentially invested in physiological processes that improve fitness. For example, locomotor activity is integral to resource acquisition through foraging effort, and mating success involves many processes (e.g., courtship, fighting and chasing) that elevate metabolic rate (MR) and require skeletal muscle involvement. By making activity energetically cheaper, substitution may ease energy constraints on an organism, allowing for higher energy allocation towards fecundity, somatic growth, or survival (Stearns, 2000).

Quantifying substitution requires extensive time and effort because it involves measuring MR in resting vs active animals at different temperatures. Perhaps the most straightforward method of estimating substitution is to compare the COA in the TNZ (COA_{TNZ}) and below the TNZ ($COA_{<TNZ}$). If no substitution occurs, COA_{TNZ} and $COA_{<TNZ}$ will be the same (Figure 1, solid lines). In other words, the slopes of the relationship between MR and activity are parallel within and below the TNZ, and MR increases by the same amount for each unit activity. If substitution occurs below the TNZ, however, some (or all) of the additional energy required for thermoregulation is "paid for" by the heat produced by activity, such that MR does not increase by the same amount for each unit activity (i.e., the slope becomes shallower; Figure 1, dotted line). Therefore, substitution can be quantified as the difference between COA_{TNZ} and $COA_{<TNZ}$ (Figure 1). Note that it is possible that the COA is greater below than within the TNZ (e.g., due to the disruption of the boundary layer; Pauls, 1981), in which case $COA_{<TNZ}$ will be greater than COA_{TNZ} and calculating substitution as $COA_{TNZ} - COA_{<TNZ}$ would return a negative value.

Research on substitution has been primarily focused on quantifying the extent of substitution at the population level and how substitution is affected or driven by the mode and intensity of activity and traits like thermal conductance. Substitution is reduced or absent in endotherms when they use an atypical mode of locomotion (e.g., terrestrial locomotion in birds) (Bryant et al., 1985; Pohl and West, 1973) and have an increased thermal conductance when active than when resting due to the state of the boundary layer (Hart, 1950; Hart and Heroux, 1955; Pauls, 1981). Morphology and behaviour both play important roles in substitution; for example, sea lions use their extensive peripheral fat stores to reduce conductance and avoid high swim speeds that would likewise reduce conductance (Liwanag et al., 2009). Moreover, substitution seems more likely to occur under voluntary activity (Chappell and Hammond, 2004; Chappell et al., 2004) than forced activity (Hart and Heroux, 1955; Pauls, 1981; Yousef et al., 1973). Humphries and Careau (2011) conducted a meta-analysis of substitution studies and found that substitution is negatively correlated with size in birds (but not in mammals) and positively correlated with intensity of activity in both birds and mammals. Overall, substitution is a common, potentially important energy-saving mechanism subject to extensive interspecific variation, yet we still do not know the extent and consistency of individual differences in substitution.

Individual Variation & Repeatability

Individual variation plays a key role in the adaptive evolution of any trait, as it represents the "raw material" upon which natural selection acts. Moreover, heritable individual differences are a prerequisite for the adaptive divergence among populations, species, and taxa (Hayes and

Jenkins, 1997). Given the potential importance of substitution for fitness in endotherms, it is surprising that no study has attempted to quantify whether individuals consistently differ in the degree to which they substitute. In other words, we still do not know if substitution is a potentially heritable trait and therefore whether selection (directly or indirectly) could act on this trait. Focussing on individual variation not only represents a first step to understanding the adaptive nature of substitution, but may also help identifying its underlying mechanisms and functional relationships with other traits.

For most ecologically and evolutionary relevant phenotypic traits (e.g., behaviour or physiological parameters), an individual's expression of the trait is labile and can change in response to environmental conditions, across time, or in response to changes in other traits (Araya-Ajoy et al., 2015). Therefore, singular measurements might not represent variation among individuals. Total phenotypic variance in a trait is the sum of among-individual variance (i.e., the difference in individuals mean trait values due to genetic and permanent environmental factors) and within-individual variance (i.e., the differences in trait value for a single individual due to measurement error, labile environmental effects and acclimation). Individual repeatability, calculated as the ratio of among-individual variance over phenotypic variance, is a useful metric because it gives an idea of how much consistent individual differences compare to variation that occurs within individuals. Repeatability also represents the upper limit to heritability (Boake, 1989; Wolak et al., 2012).

Estimating individual repeatability requires taking repeated measures on multiple individuals to partition the trait variance at the among-vs within-individual levels. Since a certain degree of variation in the measurements can be caused by non-biological effects such as different measuring devices or "session" effects that are often an unavoidable part of experimental design,

one form of repeatability that accounts for such "inflation of phenotypic variance" (Nakagawa and Schielzeth, 2010) is termed "consistency repeatability". Consistency repeatability provides an estimate as to how much variation in the trait is attributable to differences among individuals, but cannot account for whether that variation is attributable to individual variation in other traits (such as fixed morphological or physiological traits). In some cases, it is also desirable to calculate "adjusted repeatability" in a trait by accounting for biological traits as fixed effects (Nakagawa and Schielzeth, 2010). Estimating the adjusted repeatability alongside the consistency repeatability is important, as it gives an idea of how much of the inter-individual differences in a trait is due to among-individual variation in other, co-varying traits (Roche et al., 2016; Santos et al., 2015).

Substitution and Covariates

Among all possible covariates to substitution, the most likely important set of traits are those that influence heat dissipation. Insulation (pelage and fat) modulates thermal conductance, so variation in traits like fur density and skin (fat and tissue) density affect how much heat is lost from working muscles. Dry heat transfer through conduction and convection (the primary heat-loss mechanism used by rodents) occurs through a gradient between the surface and the environment. Therefore, the overall body shape of an animal that determines its surface/volume ratio might influence substitution (Mitchell et al., 2017b; Mitchell et al., 2018). As well, the size of morphological features— primarily tails in rodents—can enhance dry heat transfer (Škop et al., 2020).

Another set of possible covariates to substitution include traits that are energetically expensive by nature. Vital organs that are "functionally significant" such as the heart, liver, and kidneys incur a large ongoing metabolic cost (Konarzewski and Diamond, 1995), and their size is in individuals is directly tied to energy availability (Mitchell et al., 2017a). As allocating energy to somatic growth and maintenance can be costly, selection for energy saving mechanisms may occur alongside selection for organ size. Hence, individual variation in these vital organs may generate individual variation in substitution.

Skeletal muscle is also energetically expensive—although muscle tissue has a low metabolic rate per gram at rest, it constitutes the majority of lean body mass and is highly metabolically demanding when the animal is active (Raichlen et al., 2010). The heat used in substitution also is produced by skeletal muscles (González-Alonso, 2012), which could mean that the size of specific skeletal muscles may covary with substitution both because of potential energy savings and because of the role of heat production.

Finally, hematocrit (the proportion of red blood cells to whole blood) may also be a good target for investigation as an "expensive" trait. The process of creating red blood cells (erythropoiesis) is halted or attenuated in response to food deprivation, suggesting that maintaining red blood cells is energetically costly (Rios et al., 2005). Furthermore, hematocrit has been found to be a strong predictor of exercise intensity in mice (Rezende et al., 2006), increases with cold acclimatization (Barceló et al., 2017), and has been measured in previous studies alongside muscle morphology and vital organ size (Kelly et al., 2017; Swallow, 2005).

Objectives

The main objective of this thesis is to test whether activity-thermoregulatory heat substitution is repeatable in white-footed mouse (*Peromyscus leucopus*), a small North American rodent. *Peromyscus* mice have been used in previous studies of substitution (Chappell and Hammond, 2004; Chappell et al., 2004) and are ideal in that they occupy a large thermal habitat and are active year-round (Borniger and Nelson, 2017). To quantify substitution, I measured oxygen consumption (VO_2) as a proxy for MR and voluntary locomotion speed as a proxy for activity intensity at two different ambient temperatures below the TNZ. To estimate consistency repeatability, I took repeated measures of substitution and used mixed models to partition the phenotypic variance at the among- and within-individual levels. A second objective of this thesis was to test whether substitution is covarying with a suite of morphological traits such as fur density, surface area, organ size, and hematocrit. Calculating adjusted repeatability after having accounted for these covariates, and comparing it to the consistency repeatability, will provide an idea of how much the morphological traits explain inter-individual differences in substitution.

Methods

Ethics

All procedures were approved by the Animal Care Committee at the University of Ottawa and were in accordance with the Canadian Council on Animal Care's guidelines.

Study Animals

Individually marked (ear punches) adult female white-footed mice, originally purchased from the

Peromyscus Genetic Stock Center (University of South Carolina), were used for the experiment. Mice were 10-13 months old at the start of the study. Mice were housed in group of four in standard rat cages (42 × 21 × 20 cm) and rodent chow and water were provided *ad libitum* in addition to nesting material and enrichment. The light cycle in the room was maintained with 11 hours of light followed by 13 hours of dark, with a 30-min gradual switches starting at 7:00 (dark to light) and at 17:30 (light to dark). Room temperature was controlled at 22°C as per the standard of the facility.

Experimental Design

On September 24th, 2020, the experiment started with the objective to have 3 repeated measures of substitution in 48 individuals over 4 months. Since we could only measure 8 individuals at a time (see below), mice were assigned to 6 rotating groups of 8 mice each. Each week a different group was measured while the other mice remained in their home cage. Therefore, there was a period of 5-8 weeks in between repeated tests, and the experiment lasted until January 9th, 2021. Natural and accidental deaths in the mice population reduced the number of mice in the study, but all tests where the animal was not removed early or died during either temperature condition were included for analysis.

To measure substitution, mice were individually placed in 1 of 8 metabolic cages (randomly assigned) for a period of 96 hours (4 days), during which MR and locomotor activity were monitored. Metabolic cages were housed within an environmental cabinet (Sable Systems International, North Las Vegas, NV, USA) set at either 22°C (the temperature the animals were acclimatized to) or 10°C for the first and second 48 h of the test (order of temperature randomized). A thermometer placed on the cabinet wall adjacent to the second lowest shelf

confirmed that the temperature stayed within ±0.5°C of the set temperature. The cabinet's light source was set to cycle to 12 hours of light followed by 12 hours of dark. Mice were inspected through the glass door of the cabinet daily but remained undisturbed while tested. No nesting material was provided in these cages, and enrichment was limited to the running wheel and a body mass monitor (for which the primary function in this study was to allow mice to "hide"). This protocol resulted in one 48-hour set of metabolic and behavioural observations for each temperature condition for each test, which allowed for calculation of substitution (see below). Metabolic cages were cleaned every two weeks, between tests. This was repeated until all groups were tested three times.

At the end of each test week, the tested group of mice had blood drawn for hematocrit measures. Blood was taken from the saphenous vein, filling individual capillary tubes. Tubes were placed into a hematocrit centrifuge and spun at a speed of 10,000 rpm for 5 minutes (Wennecke, 2004). Hematocrit was calculated as a ratio of the length of the packed red blood cells to total length of the column.

Body mass was collected for each test using two different measurements (the mass before entering the metabolic cage and the mass after exiting the metabolic cage). The two mass measurements were highly positively correlated with each other ($r = 0.895$, $P < 0.001$), so initial mass was used as a predictor variable when needed in analyses.

Respirometry

Rates of oxygen consumption (VO_2), voluntary wheel running, and home-cage activity were all measured with an 8-cage Promethion multiplex system (Sable Systems International, North Las

Vegas, NV, USA), which uses pull-mode flow-through respirometry. A flow generator module pulled a constant air flow multiplexed from the 8 metabolic cages. The gas analysis module measured O_2 consumption while taking changes in water vapour dilution and barometric pressure into account.

Metabolic cages were supplied with a stainless-steel wheel (11.5 cm diameter). Mice had free access to the wheel and air flow around the wheel was unrestricted to integrate wheel activity with respirometry. Wheel revolutions were recorded by a reed switch placed parallel to the wheel, which responded to a magnet placed on the outer wheel rim. Locomotor activity in the cage was measured using the BXZ-1 beam break activity monitor. The activity monitor was designed to record movement across the cage floor and ignore fixed objects such as the hoppers and wheel. Any cage activity with a speed less than 0.01 m per s was considered as nonlocomotory cage activity (e.g., grooming).

The dwell time (i.e., a period of continuous metabolic and behavioural data monitoring for a given cage) was set at 30 seconds and the inter-leave ratio was set to 4, such that each cage was monitored for 30 sec every 5 minutes. For a period of 24 hours, a total of 288 30-sec VO_2 measurements were collected for each cage, along with the distance moved on the wheels and in the cage during the same 30-sec as the VO_2 measurement. Data from the Promethion system were processed and transformed using Expedata (Sable Systems International, North Las Vegas, NV, USA). For each 30-sec measurement, total locomotion speed (in m per s) was quantified by adding the distance run on the wheel to the distance moved in the cage, then dividing by 30. Time (in s) spent resting (during the dwell period) was defined as measurements in which the mouse was inactive (absence of locomotor and non-locomotor activity such as grooming, eating,

or drinking). Changes of temperature in the cabinet were simultaneous with initial baseline respirometry measurements and were excluded from final data analysis.

Torpor

For substitution to be accurately assessed, the animal must be normothermic, otherwise the metabolic depression induced by reduced body temperature (T_b) will return underestimated and/or negative substitution values (see black star in Figure 1). Despite the choice of the lower temperature condition being predicated on avoiding torpor, inspection of the data suggested that the mice engaged in torpor during the tests at 10°C. Torpor is typically identified by significant reductions in T_b, however it is also distinguishable by periods in which MR is much lower—down to 29% of the expected rate—and followed by a rewarming period in which MR increases by up to 11.6 times (Diedrich et al., 2015). Torpor occurred in bouts during the light cycle at 10°C. Torpor during these periods was defined by a resting (i.e., inactive in both temperature conditions) mouse displaying lower mean MR at 10°C than at 22°C (Figure 2), during the light cycle (the "daytime")—as *Peromyscus* only exhibit torpor diurnally (Lynch et al., 1978a). Observations where the mouse was engaged in torpor were removed from the data set before calculation of substitution (see below). In total, 3586 observations out of 134741 were removed from the data set.

Tissue Sampling

After the third set of tests, mice were sacrificed in groups by a two-step euthanasia process of CO_2 inhalation and cervical dislocation. Post-euthanasia, the body length of the mice was

measured from the tip of the snout to the base of the tail using a standard ruler. Tail length was measured separately, and tails were cut at the base and weighed. The skin of the mouse was sprayed with an 80% ethanol solution (Bagchi and Macdougald, 2019). A ventral incision from the mouth to the tail was made and the skin was removed in one piece, excluding the skin on the head, paws, and tail. The fascia and small fat deposits adhering to the skin were kept. The skin was stretched and pinned onto a corkboard and set to dry overnight. An electric razor was used to crop the hair from the skin. The blade was set at 0.4 mm and hair shorter than this length was not cropped. The cropped hair and shaved skin were weighed separately. The area of the skin was measured using a ruler and approximating a rectangle and used as a proxy for the trunk surface area of the mouse.

The heart, kidneys, and liver were removed from the body cavity. Fat deposits, mesentery, and connective tissue were trimmed from the organs, then the organs were rinsed in saline and blotted dry (Morawietz et al., 2004; Scudamore et al., 2014). The organs were all weighed separately. The right gastrocnemius muscle was removed from the mice using a method adapted from Kelly et al. (2017). Fascia was first removed from the muscle complex using a surgical probe (Wang et al., 2017). The calf muscle complex was separated from the tibia-fibula using a surgical probe. The Achilles tendon was then cut midway. The gastrocnemius muscle was separated from the soleus and plantaris muscles using forceps. The muscle was then cut from the condyles of the tibia and fibula, and then weighed.

Statistical Analysis

All models were fitted using ASReml-R 4.0 (Butler et al., 2018). A series of linear models were used to extract individual substitution values separately for each test (See Figure 3 and Table 1

for a representation of the data and the model output for one test on one mouse). The model was fitted with VO_2 (in mlO_2 per s) as the response variable and the predictors were total activity speed (in m per s), temperature (10 vs 22°C), and their interaction (Table 1). The reference level for the temperature variable was set at 10°C, such that the model estimate for the activity speed variable corresponded to the slope of the MR-speed relationship at 10°C, which is equivalent to the COA (in mlO_2 per m) at 10°C. More importantly, the estimate for the "temperature × speed" interaction corresponded to how the slope of the MR-speed relationship differed at 22 vs 10°C, which is equivalent to substitution (Table 1). As I did not use forced activity, measuring substitution relied on the voluntary activity of the mice. Therefore, tests where mice had less than 5 observations (out of approximately 576) where the speed exceeded 0.1 m per s in either of the two temperature conditions were removed from the dataset (Figure 4). The speed of 0.1 m per s was used as a threshold as slower locomotor speeds were not considered to be sufficiently "active". Twelve tests from 11 individual mice were removed. In total, 117 substitution measures were calculated and used for analysis from 46 mice.

A quick look at the raw data (Figure 3) suggests the linear models used to extract substitution violated key assumptions (namely homoscedasticity and the normality of residuals). However, we do not use the linear models to test for significance, and according to Lande and Arnold (1983), selection (or substitution) estimates derived from coefficients do not "depend on distribution assumptions". For completeness, we re-estimated substitution using a second method that did not involve model fitting, using the following equation:

$$(1)\ COA = \frac{\Delta VO_2}{\Delta\, Speed} = \frac{\mu_A - \mu_R}{\mu_S}$$

where μ_A represents the mean MR when the mouse was moving at speeds greater than or equal to 0.1 m per s during the test, μ_R represents the mean MR when the mouse was at rest during the test, and μ_S represents the mean speed of the mouse during the test. Substitution was calculated as the difference in COA between the two temperature conditions for the test. This means of deriving substitution will be referred to henceforth as the "second method".

Other metabolic parameters—namely daily energy expenditure (DEE) and COA at 10°C and 22°C—were extracted for each mouse for each test. DEE was calculated as the average metabolic rate (kcal per hour) for the test, using the Weir equation (Weir, 1949). COA was extracted from the model (Table 1). These parameters were extracted for comparison with substitution, along with the body mass and mean locomotor speed for each test for each mouse.

Once the measures of substitution (and other variables like body mass, DEE, and COA) were extracted, I standardized the variables to a mean of 0 and variance of 1 (to ensure that estimates were comparable across studies and improve the interpretability of the results, see Schielzeth, 2010) and ran linear mixed models (LMM) to estimate repeatability. First, I estimated consistency repeatability (Nakagawa and Schielzeth, 2010) by only including experimental variables such as test sequence (1 to 3, categorical) and metabolic cage (factor with 8 levels) as fixed effects. The mouse identity was fit as a random effect to estimate the among-individual variance (V_I). The within-individual variance was estimated as the residual variance (V_e). Statistical significance of V_I was tested with a likelihood ratio test (LRT).

To find possible covariates of substitution, I re-ran the LMM with substitution as the dependent variable after including several additional fixed effects. I included all of the morphological traits (body mass, body length, tail length, surface area, fur mass, skin mass, tail mass, and organ masses), torpor bouts (two-level factor, presence, or absence of torpor during a

test), hematocrit, and the age of the mouse in the model and tested for their statistical significance with a conditional Wald test. The variance inflation factor (VIF) for each fixed effect was calculated to assess whether the model exhibited multicollinearity. The VIFs for all added fixed effects to the model were well below the conservative upper limit of 5 (Akinwande et al., 2015). The V_I and V_e estimates from this second model were used to calculate adjusted repeatability (Nakagawa and Schielzeth, 2010). Morphological measurements were not collected for 10 individuals in the experiment—these mice were omitted from the LMM with the additional fixed effects. Tests where no hematocrit measurements were also excluded. The total number of mice in this data subset was 36, with 88 observations (See Table 4). For completeness, I also re-calculated consistency repeatability on this subset of the data.

Caveat: error in raw MR measurements

The temperature control cabinet distributed air through a central fan on the ceiling of the cabinet, but additional fans were placed on each shelf to homogenise temperature vertically across shelves. However, this had unforeseen consequences on the air flow within the cabinet. As the fans increased the speed of air flowing past the cages adjacent to the fans, air pressure around the cages decreased (the Bernoulli effect, personal communication with Sable Systems, January 18th, 2021). Lighton and Halsey (2011) recommend that pull-through systems should be maintained in environments with a slight positive pressure, as this ensures that all excurrent air is captured by the system. If not all excurrent air is measured, then the flow rate recorded by the system is underestimated, and the MR is in turn underestimated (Lighton, 2008). Overall, this created an experimental setup in which raw measurements of MR were different between cages, with underestimated MR measurements from cages closest to the fans. One month before this study,

the same system and the same mice were used in an experiment where metabolic cages were maintained outside the temperature control cabinet. In that experiment there was no reported effect of cage on MR, and DEE was highly repeatable at 0.786 (Abdeen et al., under review). Variation in air pressure around the cages was a major source of measurement error in our experiment, increasing the residual variance and resulting in low repeatability estimates for DEE. The extent to which measurement error in raw MR measurements affected repeatability of substitution is unknown, but is probably less problematic than for DEE because substitution was calculated as the difference in the COA at 10 vs 22°C, which were presumably equally underestimated (both measured with the same cage, see figure 5).

Results

Repeatability

Body mass had the highest repeatability (R = 0.90), followed by mean locomotor speed (R = 0.648 at 10°C and 0.491 at 22°C). Although COA was significantly repeatable at 22°C (R = 0.269), repeatability was lower and nonsignificant at 10°C (R = 0.126, P = 0.143). The repeatability of DEE in both temperatures was low and nonsignificant (Table 3).

Repeatability of Substitution

Consistency repeatability (±se) of substitution was 0.313±0.132 (P = 0.00282), indicating that approximately one third of variation in substitution was attributable to differences among individuals (Figure 6). The metabolic cage and test sequence did not significantly influence substitution. Overall, substitution was expressed by the mice (Figure 7) and mean substitution

was 0.0168 mLO_2 per m (Table 2). Results remained qualitatively similar when using the second method for deriving substitution (Figure 6) with a consistency repeatability (±se) of 0.278±0.132 (P = 0.0064). There was a very strong positive correlation between substitution estimates derived from the two methods (r = 0.952, P < 0.001).

Covariates

Adjusted repeatability (calculated using V_I and V_e from a LMM with multiple fixed effects, see Table 3) was 0.0957±0.177 (P = 0.286). Using the same subset of the data (i.e., substitution measures without a missing value for one of the covariates), consistency repeatability was similar to when using all the observations (R = 0.294±0.152, P = 0.015). The same significant covariates were found using substitution as derived from the second method.

Body length and heart mass were significantly and positively related to substitution (Table 4, Figure 8). The trunk surface area of the mouse and liver mass were significantly and negatively related to substitution (Table 4, Figure 8).

Discussion

Kemp (2006) considered endothermy as a paradigm for the evolution of complex traits. Indeed, endothermy has multifarious implications for many complex traits like body temperature, metabolic rate, and locomotor activity. Substitution is a complex trait that emerges from interplay between metabolic rate, locomotor activity, and heat dissipation. Complex traits are influenced by several genetic and environmental factors, and as such individual differences in complex traits should be repeatable over time. While it was already known that substitution is

influenced by the environment (McNamara et al., 2004; Travis et al., 1999), the individual repeatability of substitution remained unknown. This thesis established that around one third of the phenotypic variation in substitution is attributable to individual differences. Moreover, the expression of substitution is dependant on variation in several underlying traits—in this case, morphology, and anatomy of the organism. Given that substitution is complex and its expression is highly dependant on other traits, understanding how it evolves requires consideration of the physiological roles of surface area, body length, and organ size.

Conductivity

Conductivity of fur, skin, and fat determines how much metabolic heat lost to peripheral tissue is retained vs. lost to the outer environment (González-Alonso, 2012). Here, I could not directly measure individual differences in conductivity, which is best quantified as the heat transfer and temperature difference across the gradient of the tissue (Boyles and Bakken, 2007; Jacobsen, 1980; Knight, 1987). However, I measured the mass of the skin and pelage, which have been used as a validated proxy for the conductivity (Barnett, 1959). Despite the supposedly strong influence that skin and fur mass have on conductance, I found no relationship with substitution. Peripheral fat deposits are also known to influence conductivity, however as *Peromyscus* mice do not accumulate much subcutaneous fat and remain lean throughout their lifetime (personal observations, November 19th, 2020), fat deposits were too small and visually indistinct from skin to excise and measure accurately.

A strong and negative relationship was found between the trunk surface area of the mice and substitution. All else being equal, heat transfer is directly proportional to surface area through which the heat is being conducted. Indeed, it has been well established that the ratio of

surface area to mass changes the rate of heat exchange: greater surface area to mass ratios results in more dry heat loss in an animal (Mitchell et al., 2018). In mice, the bulk of dry heat loss is facilitated by trunk surface area—which probably explains why surface area strongly and negatively covaried with substitution.

Although statistically nonsignificant ($P = 0.0781$), a negative trend was found between tail mass and substitution. In rodents, tails are well vascularized and sparsely covered in hair (Škop et al., 2020), and the degree of heat loss contributed by the tail varies between species. While in laboratory rats the tail contributes 17% of total heat loss, in laboratory mice the tail contributes only 4.9% (Škop et al., 2020). Although the extent of heat loss occurring through the tail is unknown for *Peromyscus*, the trend detected here suggests it might play a role in substitution. Overall, a mouse with a small surface area (trunk and tail) should have a better ability to conserve heat generated by activity, and hence, potentially display higher levels of substitution than a mouse with a large surface area.

Body Length

Intuitively, a mouse with a longer body should experience higher heat loss than a shorter one, as the surface area would be greater for a longer body—but instead body length was positively related to substitution. This result is possibly attributable to variation in head size, which was not directly measured in this study but was included in overall body length. In *Peromyscus leucopus*, the length of the skull is 26% of the total body length (Clark, 1941). The coefficient of variation for body length and skull length was reported by Clark to be 4.19 and 2.27 respectively (1941). What this implies is that mice in this study might have varied in the overall dimensions of their

heads, and the effect of head size would have been picked up by the body length variable. Differences in cranial size produce differences in brain size. Brains are the most energetically expensive organ in the body—accounting for 20% of total oxygen consumption in humans (Wang et al., 2014) and are consistent in their energetic costs (Raichle and Gusnard, 2002). Many mammals (including members of *Peromyscus*) have much larger brains than would be expected given their BMR (Isler and Schaik, 2006). Animals with disproportionally large brains are thought to compensate for the incurred metabolic cost through extraneous mechanisms, and it is possible that substitution is a mechanism that can compensate for greater brain mass.

Furthermore, brains are more sensitive to temperature changes and remain cooler than the rest of the body (Matsuda-Nakamura and Nagashima, 2014). Due to this sensitivity the brain relies on heat loss mechanisms beyond dry heat transfer. Selective brain cooling works through several mechanisms that are often species dependant, but two mechanisms that are found in most mammals is the pre-cooling of arterial blood headed to the brain and the drainage of cooled venous blood to veins around the brain (Caputa, 2004). What this all means is that there is a smaller temperature gradient between the environment and the surface of the head, resulting in less heat loss in the cold. If variation in head size was picked up by the body length variable, mice with longer bodies had bigger heads than mice with shorter bodies. Mice with longer bodies would lose less heat than mice with shorter bodies, as although there is a greater total surface area for longer mice there is proportionally less area with a large temperature gradient. This reduced heat loss likely contributes to greater substitution. In order to evaluate the possibility that brain size is a key covariate of heat substitution, measuring brain size directly or by the proxy of endocranial volume in relation to substitution would be necessary.

Hematocrit

Hematocrit was chosen as a covariate for investigation due to both its potential metabolic cost and its relationship to aerobic performance, but no relationship was found. Higher numbers of red blood cells directly increase O_2 in the blood which can increase maximal oxygen consumption (Schuler et al., 2010), making it more energetically expensive. Hematocrit is also a plastic trait, known to generally increase in response to changing environmental conditions (Petit and Vézina, 2014; Tufts et al., 2013), or in response to activity and or wheel access (Kelly et al., 2017). Although wheel access was limited to the mice compared to studies where a "training effect" from aerobic exercise is a goal (Baum et al., 2018; Kelly et al., 2017; Lark et al., 2018), it is possible that hematocrit varied in relation to activity but not in relation to substitution.

Organ & Muscle Mass

The heart, liver, and kidneys are all considered to be part of the "metabolic machinery," with larger organs sizes needed for the maintenance of a high aerobic scope (Gçbczyński and Konarzewski, 2009; Kelly et al., 2017). Larger "metabolic machinery" organs are also correlated with higher BMRs, and indeed in laboratory mice variation in organ size accounts for 52% of variation in BMR (Konarzewski and Diamond, 1995). This high cost of organ maintenance may be offset by the use of substitution, which is why the heart, liver, and kidneys were measured as possible covariates. However, the relationships between substitution and the mass of two organs (liver and heart) were contradictory. Furthermore, no relationship was found between kidney mass and substitution. This suggests that there are other functional explanations for the relationship between organ size and substitution.

Liver mass is both correlated with maximal VO_2 consumption (Kelly et al., 2017) and increases in response to cold acclimatization (Kgwatalala and Nielsen, 2004) and because of these previous findings it was expected that liver mass would positively covary with substitution. However, the opposite relationship was found. An explanation for this is that the liver is implicated in non-shivering active thermogenesis, which is suppressed by exercise (Arnold et al., 1986). The liver has a key role in the activation of brown adipose (Simcox et al., 2018) which produces heat through increased proton leak in the mitochondria. The liver also facilitates lipolysis of white adipose tissue, which is used to produce heat (Himms-Hagen, 1984). It is possible that mice that substitute heat poorly use active non-shivering thermogenesis more extensively and therefore maintain larger livers to facilitate this form of thermogenesis.

Heart mass positively covaried with substitution. Use of substitution by a mouse could compensate for the energy costs incurred by a larger heart. There may be an additional function reason for the relationship—a larger heart may be an advantage to exercising mice experiencing cold stress. Cold stress induces cardiovascular changes—not only does vasoconstriction reduce blood flow to peripheral tissue, but arteries supplying the extremities decrease their flow and conductance (up to 40% for the femoral artery, González-Alonso, 2012). Reducing blood flow directly causes less heat loss in peripheral tissue and in the extremities. However, there are consequences to reducing blood flow. As blood volume is concentrated in core tissues, blood pressure increases accordingly, which in turn overloads the heart (Choo et al., 2018; Halonen et al., 2011). This pressure overload results in cardiac hypertrophy and contractile abnormalities, which can result in heart failure (Lu and Xu, 2013). A naturally larger heart can accommodate increases in blood pressure better than a smaller heart (as it has the capacity to safely hold greater blood volume). Therefore, mice with larger hearts likely have less heat loss from active skeletal

muscles than mice with smaller hearts, as they have the potential for greater reductions in peripheral blood flow and arterial conductance.

Given that larger muscles are more metabolically costly than smaller muscles (Raichlen et al., 2010) it may appear surprising that gastrocnemius muscle mass had no relation to substitution. A possible reason for the lack of relationship is that gross anatomical size —while used as a proxy in other experiments—does not totally explain the energy consumption, heat production, or heat dissipation capacities of the muscle (Dlugosz et al., 2009; Gavini et al., 2014; Mukherjee et al., 2020).

Experimental Constraints

I used female adult mice, but a mixed population of both males and females would have allowed investigation into sex differences in substitution, although other experiments relating to performance and metabolic rate also used all-female populations (Kelly et al., 2017; Konarzewski and Diamond, 1995; Swallow, 2005). It is known that female mice run much farther than male mice (Lark et al., 2018), but are more variable in their day-to-day running speed and total time spent running (Manzanares et al., 2019), which could affect the degree of substitution used.

The age of the mice was not a significant covariate with substitution in this experiment, but there was very little variance in the age-range of the mice. Juveniles run longer distances and at a greater speed than older mice (Manzanares et al., 2019), and mammalian juveniles generally incur greater metabolic costs to thermoregulation compared to adults (Liwanag et al., 2009). Determining whether substitution is used differentially across life stages will require a greater range of age than used in the current study.

We housed and experimented upon the mice at 22°C and did not provide nesting material during experimentation. However, rodents in the laboratory are typically housed at temperatures that are thermoneutral for humans (between 20 to 22°C), but well below their TNZ (Seeley and MacDougald, 2021). Mice housed at these temperatures and not between 29-32°C do exhibit signs of cold stress (namely impaired immune function and increased thermogenesis), although this is mitigated by providing food and nesting material. However, the choice of a temperature below the TNZ was validated by numerous studies (Arnold et al., 1986; Chappell and Hammond, 2004; Chappell et al., 2004; Sears et al., 2009). The persistence of using a temperature below the TNZ in research (including in the aforementioned studies) is likely attributable to facilities themselves being maintained in the human TNZ in accordance with health and safety guidelines (Seely and MacDougald, 2021). Experiments on mice housed at temperatures within the human TNZ would have to account for the acclimatory effect of a higher experimental temperature. The results here cannot represent the full extent of substitution, but a future research endeavor could focus on including a temperature in the TNZ. On the other hand, is also possible that mice would be averse to activity at a temperature in the TNZ, as even at these temperatures rodents risk hyperthermia (Rezende and Bacigalupe, 2015; Speakman and Król, 2010; van Klinken et al., 2013). This would complicate measurements of substitution based on voluntary exercise more difficult because mice may not express the full range of activity intensity in the TNZ.

It was necessary to remove torpor bouts from the dataset, as substitution can only be correctly estimated when the animal is normothermic. In this study mice used both substitution and torpor in conjunction as temporally separated energy-saving mechanisms—which Geiser (2020) suggested is common for *Peromyscus.* White-footed mice have been well-established to engage in torpor when faced with cold ambient temperature (Lynch et al., 1978a; Lynch et al.,

1978b; Rhodes, 1980) or food deprivation (Diedrich et al., 2015). Torpor has been shown to be induced at temperatures as high as 15°C, but only in laboratory settings where movement was severely restricted (Hill, 1975). In wild populations, torpor is more typically observed at lower temperatures (Lynch et al., 1978a; Lynch et al., 1978b). Shorter photoperiods (less than 12 hours light) have been positively correlated with torpor bouts, but members of *Peromyscus* have been described as using daily torpor year-round whenever the ambient temperature falls (Geiser, 2020). No relationship was found between the use of torpor (as identified by torpor bouts present during tests in individuals) and substitution. However, it would be prudent to further investigate how mice use torpor in conjunction with substitution, and how these mechanisms covary. Torpor is not an "all-or-nothing" phenomenon—it has been demonstrated that there is individual variation in the use of torpor in different conditions, with some mice being "torpor sensitive" and some being "torpor resistant" (Sheafor and Snyder, 1996). Mice (regardless of sensitivity) will use torpor if the ambient temperature is low and/or food is also withheld (Sheafor and Snyder, 1996). Use of a lower sub-thermoneutral temperature condition would induce spontaneous torpor in a higher proportion of the population, facilitating a more comprehensive study as to how mice use different energy saving mechanisms (Lynch et al., 1978b). Furthermore, torpor as a trait is expressed less in rodent species with a higher BMR, and "may reflect the presence of selective forces that acted on the metabolic machinery" (Careau, 2013). Better quantifying the relationship between substitution and torpor would also clarify the relationship between covariates such as heart mass and liver mass and substitution.

Future Directions

Using lines created from an experimental evolution approach (i.e., a laboratory selection experiment) would be a promising future avenue of research. Selection experiments have been described as a "powerful tool for investigating linkages" between quantifiable morphological, physiological, and biochemical traits and "physiological performance" of the animal (Swallow et al., 2009). To better understand how substitution covaries with other important organismal traits, one could measure the degree of substitution utilized by existing mouse lines artificially selected for high voluntary wheel running, low heat loss, or high BMR alongside their control lines (Brzęk et al., 2007; Carter and Swallow, 2000; Kelly et al., 2017; Kgwatalala and Nielsen, 2004; Swallow et al., 2009). Although I have not specifically connected wheel running to the use of substitution, Carter and Swallow (2000) discovered that mice selected for high levels of voluntary wheel running ("high-runners", HR) created smaller nests compared to control mice. Carter and Swallow (2000) also found that HR mice did not consume greater amounts of food nor spent more time running on wheels than control mice. This suggests that HR mice incur greater costs of thermoregulation (as they are less avoidant of cold), but do not do so by increasing overall energy expenditure. An interesting possibility is that HR mice used substitution more than cold avoidance (nest building) as a thermoregulatory mechanism.

Linking substitution and fecundity directly would demonstrate fitness advantages to substitution as a phenotype. Many rodents (including *Peromyscus*) have two reproductive phenotypes: photoperiod responsive reproduction and non-responsive reproduction (Heideman et al., 2010). Rodents that are responsive to photoperiod reduce reproductive capacity in response to shortening daylength, with a complete cessation when the daily photoperiod is typical of that during the winter. Rodents that are not responsive maintain reproduction all year, regardless of

temperature and light conditions. As the need for thermoregulation typically increases when day lengths are shorter (i.e., in the winter), year-long reproduction imposes a possible unfeasible energy demand on a rodent (Borniger and Nelson, 2017; Kaseloo et al., 2012). Rodents that are non-responsive to photoperiod do have a fitness advantage over mice that are responsive if they can adequately meet the increased energy demands of reproduction all year long. Lines of white-footed mice exist that have been selected for year-long reproduction. Experiments with these lines have demonstrated that the selected mice are larger, have higher BMRs, consume more food, and are more active than control mice (Heideman et al., 2010; Kaseloo et al., 2012; McDonnell et al., 2019). It would be fruitful to investigate whether these selected lines of mice also use heat substitution more extensively than control line mice.

In addition to more closely examining the selection for substitution and direct fitness benefits to substitution, another important direction is to more closely look at trade-offs. The most apparent trade-off to substitution is an impaired response to hyperthermia. Endotherm survival is more threatened by T_a above the TNZ than below it. There is an upper limit to how much heat can be lost via dry heat transfer and evaporative cooling, beyond which the T_b will rise to lethal levels (Rezende and Bacigalupe, 2015). There are also energetic costs attributable to high temperatures, namely those incurred by cooling and the Q_{10} effect (Rezende and Bacigalupe, 2015). We have demonstrated that substitution has a negative relationship with both tail mass and trunk surface area—both of which represent the main sources of heat loss in rodents. Traits that decrease risks of hyperthermia may reduce the ability to use substitution. If such a trade-off between heat dissipation and substitution exists, then the selection gradient for substitution would depend on the thermal niche of the organism (Gvoždík, 2018). For

Peromyscus in particular—a genus that is broadly generalist—selection for substitution likely occurs in higher latitudes, but may decrease with increased climate change.

Conclusion

Our study provides the first repeatability estimate for substitution as a trait and has demonstrated that there is individual variation in the trait attributable to phenotypic variation in morphology. However, given the constraints of the experimental set-up and the unexpected use of another energy saving mechanisms by the mice, it is likely that this estimate is conservative and I have not identified the full extent of individual variation in substitution. I suggest future research should focus on mitigating these constraints and expand upon this work by measuring substitution in conjunction with other energy-saving mechanisms. Subsequent work should also be done to illuminate whether this trait conveys direct fitness advantages or constitutes a trade-off with necessary heat-loss—both of which are important to fully defining the phenotypic expression of substitution.

References

Abdeen, A., Agnani, P. and Careau, V. (under review). The active mouse rests within: energy management among and within individuals. Functional Ecology (MS # FE-2021-00627.R1, resubmitted on 2021-08-13).

Akinwande, M. O., Dikko, H. G. and Samson, A. (2015). Variance Inflation Factor: As a Condition for the Inclusion of Suppressor Variable(s) in Regression Analysis. *Open Journal of Statistics* **05**, 754–767.

Araya-Ajoy, Y. G., Mathot, K. J. and Dingemanse, N. J. (2015). An approach to estimate short-term, long-term and reaction norm repeatability. *Methods in Ecology and Evolution* **6**, 1462–1473.

Arnold, J., LeBlanc, J., Cote, J., Lalonde, J. and Richard, D. (1986). Exercise suppression of thermoregulatory thermogenesis in warm- and cold-acclimated rats. *Canadian Journal of Physiology and Pharmacology* **64**, 922–926.

Arnold, W., Ruf, T. and Kuntz, R. (2006). Seasonal adjustment of energy budget in a large wild mammal, the Przewalski horse (Equus ferus przewalskii). II. Energy expenditure. *Journal of Experimental Biology* **209**, 4566–4573.

Bagchi, D. P. and Macdougald, O. A. (2019). Identification and dissection of diverse mouse adipose depots. *Journal of Visualized Experiments* **2019**, 1–9.

Barceló, G., Love, O. P. and Vézina, F. (2017). Uncoupling basal and summit metabolic rates in white-throated Sparrows: Digestive demand drives maintenance costs, but changes in muscle mass are not needed to improve thermogenic capacity. *Physiological and Biochemical Zoology* **90**, 153–165.

Barnett, S. A. (1959). The skin and hair of mice living at a low environmental temperature. *Quarterly Journal of Experimental Physiology and Cognate Medical Sciences* **44**, 35–42.

Baum, O., Sollberger, C., Raaflaub, A., Odriozola, A., Spohr, G., Frese, S. and Tschanz, S. A. (2018). Increased capillary tortuosity and pericapillary basement membrane thinning in skeletal muscle of mice undergoing running wheel training. *Journal of Experimental Biology* **221**,.

Boake, C. R. B. (1989). Repeatability: Its role in evolutionary studies of mating behavior. *Evol Ecol* **3**, 173–182.

Borniger, J. C. and Nelson, R. J. (2017). Photoperiodic regulation of behavior: Peromyscus as a model system. *Seminars in Cell and Developmental Biology* **61**, 82–91.

Boyles, J. G. and Bakken, G. S. (2007). Seasonal changes and wind dependence of thermal conductance in dorsal fur from two small mammal species (Peromyscus leucopus and Microtus pennsylvanicus). *Journal of Thermal Biology* **32**, 383–387.

Bryant, D. M., Hails, C. J. and Prys-Jones, R. (1985). Energy Expenditure by Free-Living Dippers (Cinclus cinclus) in Winter. *The Condor* **87**, 177.

Brzęk, P., Bielawska, K., Książek, A. and Konarzewski, M. (2007). Anatomic and molecular correlates of divergent selection for basal metabolic rate in laboratory mice. *Physiological and Biochemical Zoology* **80**, 491–499.

Butler, D. G., Cullis, B. R., Gilmour, A. R., Gogel, B. J. and Thompson, R. (2018). ASRemlR Reference Manual Version 4. *ASReml-R Reference Manual* 176.

Caputa, M. (2004). Selective brain cooling: a multiple regulatory mechanism. *Journal of Thermal Biology* **29**, 691–702.

Careau, V. (2013). Basal metabolic rate, maximum thermogenic capacity and aerobic scope in rodents: Interaction between environmental temperature and torpor use. *Biology Letters* **9**,.

Carter, P. A. and Swallow, J. G. (2000). Nesting behavior of house mice (Mus domesticus) selected for increased wheel-running activity. *Behavior Genetics* **30**, 85–94.

Chappell, M. A. and Hammond, K. A. (2004). Maximal aerobic performance of deer mice in combined cold and exercise challenges. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*.

Chappell, M. A., Garland, T., Rezende, E. L. and Gomes, F. R. (2004). Voluntary running in deer mice: Speed, distance, energy costs and temperature effects. *Journal of Experimental Biology* **207**, 3839–3854.

Choo, H. C., Nosaka, K., Peiffer, J. J., Ihsan, M., Yeo, C. C. and Abbiss, C. R. (2018). Peripheral blood flow changes in response to postexercise cold water immersion. *Clin Physiol Funct Imaging* **38**, 46–55.

Clark, F. H. (1941). Correlation and Body Proportions in Mature Mice of the Genus Peromyscus. *Genetics* **26**, 283.

Diedrich, V., Kumstel, S. and Steinlechner, S. (2015). Spontaneous daily torpor and fastinginduced torpor in Djungarian hamsters are characterized by distinct patterns of metabolic rate.
Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology.

Dlugosz, E. M., Chappell, M. A., McGillivray, D. G., Syme, D. A. and Garland, T. (2009). Locomotor trade-offs in mice selectively bred for high voluntary wheel running. *Journal of Experimental Biology* **212**, 2612–2618.

Gavini, C. K., Mukherjee, S., Shukla, C., Britton, S. L., Koch, L. G., Shi, H. and Novak, C. M. (2014). Leanness and heightened nonresting energy expenditure: Role of skeletal muscle activity thermogenesis. *American Journal of Physiology - Endocrinology and Metabolism* **306**,.

Gçbczyński, A. K. and Konarzewski, M. (2009). Metabolic correlates of selection on aerobic capacity in laboratory mice: a test of the model for the evolution of endothermy. *Journal of Experimental Biology* **212**, 2872–2878.

Geiser, F. (2020). Seasonal Expression of Avian and Mammalian Daily Torpor and Hibernation: Not a Simple Summer-Winter Affair†. *Frontiers in Physiology* **11**, 1–19.

González-Alonso, J. (2012). Human thermoregulation and the cardiovascular system. *Experimental Physiology* **97**, 340–346.

Grueber, C. E., Nakagawa, S., Laws, R. J. and Jamieson, I. G. (2011). Multimodel inference in ecology and evolution: Challenges and solutions. *Journal of Evolutionary Biology* **24**, 699–711.

Gvoždík, L. (2018). Just what is the thermal niche? *Oikos* **127**, 1701–1710.

Halonen, J. I., Zanobetti, A., Sparrow, D., Vokonas, P. S. and Schwartz, J. (2011). Relationship between outdoor temperature and blood pressure. *Occupational and Environmental Medicine* **68**, 296–301.

Hart, J. S. (1950). Interrelations of Daily Metabolic Cycle, Activity, and Environmental Temperature of Mice. *Canadian Journal of Research* **28d**, 293–307.

Hart, J. S. and Heroux, O. (1955). Exercise and temperature regulation in lemmings and rabbits. *Canadian Journal of Biochemistry and Physiology*.

Hayes, J. P. and Jenkins, S. H. (1997). Individual Variation in Mammals. *Journal of Mammalogy* **78**, 274–293.

Hedrick, M. S. and Hillman, S. S. (2016). What drove the evolution of endothermy? *Journal of Experimental Biology* **219**, 300–301.

Heideman, P. D., Pittman, J. T., Schubert, K. A., Dubois, C. M. R., Bowles, J., Lowe, S. M., Price, M. R., Pd, H., Jt, P., Ka, S., et al. (2010). Variation in levels of luteinizing hormone and reproductive photoresponsiveness in a population of white-footed mice (Peromyscus leucopus). *Am J Physiol Regul Integr Comp Physiol* **298**, 1543–1548.

Hetem, R. S., Maloney, S. K., Fuller, A. and Mitchell, D. (2016). Heterothermy in large mammals: Inevitable or implemented? *Biological Reviews* **91**, 187–205.

Hill, R. W. (1975). Daily torpor in Peromyscus leucopus on an adequate diet. *Comparative Biochemistry and Physiology – Part A: Physiology* **51**, 413–423.

Himms-Hagen, J. (1984). Nonshivering thermogenesis. *Brain Research Bulletin* **12**, 151–160.

Humphries, M. M. and Careau, V. (2011). Heat for nothing or activity for free? Evidence and implications of activity-thermoregulatory heat substitution. In *Integrative and Comparative Biology*, pp. 419–431.

Isler, K. and Schaik, C. P. V. (2006). Metabolic costs of brain size evolution. *Biology Letters* **2**, 557–560.

Jacobsen, N. K. (1980). Differences of thermal properties of white-tailed deer pelage between seasons and body regions. *Journal of Thermal Biology* **5**, 151–158.

Kaseloo, P. A., Crowell, M. G., Jones, J. J. and Heideman, P. D. (2012). Variation in basal metabolic rate and activity in relation to reproductive condition and photoperiod in white-footed mice (*Peromyscus leucopus*). *Can. J. Zool.* **90**, 602–615.

Kelly, S. A., Gomes, F. R., Kolb, E. M., Malisch, J. L. and Garland, T. (2017). Effects of activity, genetic selection and their interaction on muscle metabolic capacities and organ masses in mice. *Journal of Experimental Biology* **220**, 1038–1047.

Kemp, T. S. (2006). The origin of mammalian endothermy: a paradigm for the evolution of complex biological structure. *Zoological Journal of the Linnean Society* **147**, 473–488.

Kgwatalala, P. M. and Nielsen, M. K. (2004). Performance of mouse lines divergently selected for heat loss when exposed to different environmental temperatures. II. Feed intake, growth, fatness, and body organs. *Journal of Animal Science* **82**, 2884–2891.

Knight, F. M. (1987). The Development of Pelage Insulation and Its Relationship to Homeothermic Ability in an Altricial Rodent, Peromyscus leucopus. *Physiological Zoology* **60**, 181–190.

Konarzewski, M. and Diamond, J. (1995). Evolution of basal metabolic rate and organ masses in laboratory mice. *Evolution* **49**, 1239–1248.

Lande, R. and Arnold, S. J. (1983). The measurement of selection on correlated characters. *Evolution* **37**, 1210–1226.

Lark, D. S., Kwan, J. R., McClatchey, P. M., James, M. N., James, F. D., Lighton, J. R. B., Lantier, L. and Wasserman, D. H. (2018). Reduced nonexercise activity attenuates negative energy balance in mice engaged in voluntary Exercise. *Diabetes* **67**, 831–840.

Lefèvre, J. and Auget, A. (1931). Le problème des relations entre les chaleurs du travail et du repos. Solution et lois. Pourquoi le travail est plus économique aux basses températures. *Comptes rendus hebdomadaires des séances de l'Académie des sciences* **192**, 818–820.

Levesque, D. L., Nowack, J. and Stawski, C. (2016). Modelling mammalian energetics: the heterothermy problem. *Climate Change Responses* **3**,.

Lighton, J. R. B. (2008). *Measuring Metabolic Rates: A Manual for Scientists* (pp. 200).

Lighton, J. R. B. and Halsey, L. G. (2011). Flow-through respirometry applied to chamber systems: Pros and cons, hints and tips. *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology* **158**, 265–275.

Liwanag, H. E. M., Williams, T. M., Costa, D. P., Kanatous, S. B., Davis, R. W. and Boyd,

I. L. (2009). The effects of water temperature on the energetic costs of juvenile and adult California sea lions (Zalophus californianus): The importance of skeletal muscle thermogenesis for thermal balance. *Journal of Experimental Biology* **212**, 3977–3984.

Lu, S. and Xu, D. (2013). Cold stress accentuates pressure overload-induced cardiac hypertrophy and contractile dysfunction: Role of TRPV1/AMPK-mediated autophagy. *Biochemical and Biophysical Research Communications* **442**, 8–15.

Lynch, G. R., White, S. E., Grundel, R. and Berger, M. S. (1978a). Effects of photoperiod, melatonin administration and thyroid block on spontaneous daily torpor and temperature regulation in the white-footed mouse, Peromyscus leucopus. *Journal of Comparative Physiology B* **125**, 157–163.

Lynch, G. R., Vogt, F. D. and Smith, H. (1978b). Seasonal study of spontaneous daily torpor in the white-footed mouse, Peromyscus leucopus. *Physiological and Biochemical Zoology* **51**,.

Manzanares, G., Brito-Da-Silva, G. and Gandra, P. G. (2019). Voluntary wheel running: Patterns and physiological effects in mice. *Brazilian Journal of Medical and Biological Research* **52**, 1–9.

Maresh, J. L., Adachi, T., Takahashi, A., Naito, Y., Crocker, D. E., Horning, M., Williams, T. M. and Costa, D. P. (2015). Summing the strokes: Energy economy in northern elephant seals during large-scale foraging migrations. *Movement Ecology* **3**,.

Matsuda-Nakamura, M. and Nagashima, K. (2014). Protection of the brain against heat damage. *The Journal of Physical Fitness and Sports Medicine* **3**, 217–221.

McDonnell, S. P., Kaseloo, P. A., Wran, V. E. and Heideman, P. D. (2019). Testosterone does not mediate variation in basal metabolic rate and activity in relation to reproductive condition and photoperiod in white-footed mice (Peromyscus leucopus). *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology* **331**, 456–462.

McNamara, J. M., Ekman, J. and Houston, A. I. (2004). The effect of thermoregulatory substitution on optimal energy reserves of small birds in winter. *Oikos* **105**, 192–196.

Menzies, A. K., Studd, E. K., Majchrzak, Y. N., Peers, M. J. L. and Boutin, S. (2020). Body temperature, heart rate, and activity patterns of two boreal homeotherms in winter: homeostasis, allostasis, and ecological coexistence. *Functional Ecology* **1**,.

Mitchell, S. E., Tang, Z. H., Kerbois, C., Delville, C., Derous, D., Green, C. L., Wang, Y., Han, J. J. D., Chen, L., Douglas, A., et al. (2017a). The effects of graded levels of calorie restriction: VIII. Impact of short term calorie and protein restriction on basal metabolic rate in the C57BL/6 mouse. *Oncotarget* **8**, 17453–17474.

Mitchell, G., Sittert, S. van, Roberts, D. and Mitchell, D. (2017b). Body surface area and thermoregulation in giraffes. *Journal of Arid Environments* **145**, 35–42.

Mitchell, D., Snelling, E. P., Hetem, R. S., Maloney, S. K., Strauss, W. M. and Fuller, A. (2018). Revisiting concepts of thermal physiology: Predicting responses of mammals to climate change. *J Anim Ecol* **87**, 956–973.

Morawietz, G., Ruehl-Fehlert, C., Kittel, B., Bube, A., Keane, K., Halm, S., Heuser, A. and Hellmann, J. (2004). Revised guides for organ sampling and trimming in rats and mice – Part 3. *Experimental and Toxicologic Pathology* **55**, 433–449.

Mukherjee, S. D., Koch, L. G., Britton, S. L. and Novak, C. M. (2020). Aerobic capacity modulates adaptive thermogenesis: Contribution of non-resting energy expenditure. *Physiology & Behavior*.

Nakagawa, S. and Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: A practical guide for biologists. *Biological Reviews* **85**, 935–956.

Pauls, R. W. (1981). Energetics of the red squirrel: A laboratory study of the effects of temperature, seasonal acclimatization, use of the nest and exercise. *Journal of Thermal Biology* **6**, 79–86.

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

Pohl, H. and West, G. C. (1973). Daily and seasonal variation in metabolic response to cold during rest and forced exercise in the common redpoll. *Comparative Biochemistry and Physiology – Part A: Physiology* **45**, 851–867.

Raichle, M. E. and Gusnard, D. A. (2002). Appraising the Brain's Energy Budget. **99**, 10237–10239.

Raichlen, D. A., Gordon, A. D., Muchlinski, M. N. and Snodgrass, J. J. (2010). Causes and significance of variation in mammalian basal metabolism. *J Comp Physiol B* **180**, 301–311.

Rezende, E. L. and Bacigalupe, L. D. (2015). Thermoregulation in endotherms: physiological principles and ecological consequences. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology* **185**, 709–727.

Rezende, E. L., Gomes, F. R., Malisch, J. L., Chappell, M. A. and Garland, T. (2006). Maximal oxygen consumption in relation to subordinate traits in lines of house mice selectively bred for high voluntary wheel running. *Journal of Applied Physiology* **101**, 477–485.

Rhodes, D. H. (1980). Effects of temperature, photoperiod, and daily torpor on thyroid function in the white-footed mouse, Peromyscus leucopus. *General and Comparative Endocrinology* **42**, 134–140.

Rios, F. S., Oba, E. T., Fernandes, M. N., Kalinin, A. L. and Rantin, F. T. (2005).

Erythrocyte senescence and haematological changes induced by starvation in the neotropical fish traíra, Hoplias malabaricus (Characiformes, Erythrinidae). *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* **140**, 281–287.

Roche, D. G., Careau, V. and Binning, S. A. (2016). Demystifying animal 'personality' (or not): why individual variation matters to experimental biologists. *Journal of Experimental Biology* jeb.146712.

Santos, C. D., Cramer, J. F., Pârâu, L. G., Miranda, A. C., Wikelski, M. and Dechmann, D. K. N. (2015). Personality and morphological traits affect pigeon survival from raptor attacks. *Sci Rep* **5**, 15490.

Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution* **1**, 103–113.

Schuler, B., Arras, M., Keller, S., Rettich, A., Lundby, C., Vogel, J. and Gassmann, M. (2010). Optimal hematocrit for maximal exercise performance in acute and chronic erythropoietin-treated mice. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 419–423.

Scudamore, C. L., Busk, N. and Vowell, K. (2014). A simplified necropsy technique for mice: Making the most of unscheduled deaths. *Laboratory Animals* **48**, 342–344.

Sears, M. W., Hayes, J. P., Banta, M. R. and McCormick, D. (2009). Out in the cold: Physiological capacity influences behaviour in deer mice. *Functional Ecology* **23**, 774–783.

Seeley, R. J. and MacDougald, O. A. (2021). Mice as experimental models for human physiology: when several degrees in housing temperature matter. *Nature Metabolism* **3**, 443–445.

Sheafor, B. A. and Snyder, G. K. (1996). Energy partitioning in torpor-sensitive and torpor resistant deer mice (*Peromyscus maniculatus*). *Can. J. Zool.* **74**, 1179–1186.

Simcox, J., Geoghegan, G., Maschek, J. A., Bensard, C. L., Miao, R., Lee, S., Jiang, L., Huck, I., Kershaw, E. E., Donato, J., et al. (2018). Acylcarnitines As a Fuel Source for Brown Fat Thermogenesis. **26**, 509–522.

Škop, V., Liu, N., Guo, J., Gavrilova, O. and Reitman, M. L. (2020). The contribution of the mouse tail to thermoregulation is modest. *American Journal of Physiology - Endocrinology and Metabolism* **319**, F438–F446.

Speakman, J. R. and Król, E. (2010). Maximal heat dissipation capacity and hyperthermia risk: neglected key factors in the ecology of endotherms: Heat dissipation limit theory. *Journal of Animal Ecology* no-no.

Stearns, S. C. (2000). Life history evolution: successes, limitations, and prospects. *Naturwissenschaften* **87**, 476–486.

Swallow, J. G. (2005). Phenotypic and Evolutionary Plasticity of Organ Masses in Response to Voluntary Exercise in House Mice. *Integrative and Comparative Biology* **45**, 426–437.

Swallow, J. G., Hayes, J. P., Koteja, P. and Garland, T. (2009). Selection experiments and experimental evolution of performance and physiology. In *Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments* (ed. Garland, T.), pp. 301–351. University of California Press.

Travis, J., McManus, M. G., Baer, C. F., Fisheries, F. and Collins, F. (1999). Sources of variation in physiological phenotypes and their evolutionary significance. *American Zoologist* **433**, 422–433.

Tufts, D. M., Revsbech, I. G., Cheviron, Z. A., Weber, R. E., Fago, A. and Storz, J. F. (2013). Phenotypic plasticity in blood-oxygen transport in highland and lowland deer mice. *Journal of Experimental Biology* **216**, 1167–1173.

van Klinken, J. B., Berg, S. A. A. van den and Dijkl, K. W. van (2013). Practical aspects of estimating energy components in rodents. *Frontiers in Physiology* **4 MAY**,.

Wang, H., Wang, B., Normoyle, K. P., Jackson, K., Spitler, K., Sharrock, M. F., Miller, C. M., Best, C., Llano, D. and Du, R. (2014). Brain temperature and its fundamental properties: a review for clinical neuroscientists. *Front. Neurosci.* **8**,.

Weir, J. B. de V. (1949). New methods for calculating metabolic rate with special reference to protein metabolism. *The Journal of Physiology* **109**, 1–9.

Wennecke, G. (2004). Hematocrit - a review of different analytical methods. *Acutecaretesting.Org* 1–9.

Wolak, M. E., Fairbairn, D. J. and Paulsen, Y. R. (2012). Guidelines for estimating repeatability. *Methods in Ecology and Evolution* **3**, 129–137.

Yousef, M. K., Robertson, W. D., Dill, D. B. and Johnson, H. D. (1973). Energetic cost of running in the antelope ground squirrel Ammospermophilus leucurus. *Physiological Zoology* **46**, 139–147.

Tables

Table 1 Parameter estimates from two representative multiple linear regression models used to quantify substitution, with metabolic rate (MR; oxygen consumption in mL per s) as a function of voluntary locomotor speed (m per s), temperature (reference level 10°C), and their interaction for **A** single mouse exhibiting considerable heat substitution (mouse #23517, test #2) and **B** another mouse showing no substitution (mouse #23501, test # 1). The "speed × temperature" interaction term represent the difference between the slope of the MR-speed relationship at 22 °C vs 10°C, thus representing substitution. The "speed" term represents the slope (cost of activity) at 10°C. The slope at 22°C can be calculated by adding the "speed x temperature" interaction term to the "speed" term. See Figure 3 for a visual representation of the regression lines fitted through these two sets of data.

	A) mouse #23517, test #2		B) mouse #23501, test # 1	
Source	Estimate	± se	Estimate	± se
Intercept	0.0624	± 0.000529	0.0495	± 0.00035
Speed	0.0864	± 0.00602	0.116	± 0.00631
Temperature [22]	-0.0235	± 0.000749	-0.00165	± 0.000491
Speed × Temperature [22]	0.0312	± 0.0093	0.0000682	± 0.0104

Table 2 Descriptive statistics for body mass (g), mean locomotor activity (m per s), daily energy expenditure (DEE, kcal· per h), cost of activity (COA, mLO_2 per m), and substitution (mLO_2 per m) in 46 female white-footed mice measured a total of 117 times. Mean, standard deviation, and range (minimum to maximum) are reported. Statistics for activity, DEE, and COA are presented separately for observations taken at 22°C and 10°C. Substitution was derived by using model coefficients that compared COA at 22°C vs 10°C (see Table 1) and by the second method described in text.

	Mean	SD	Min	Max
Body Mass	19.5	3.45	13.8	33.5
Locomotor Activity (10°C)	0.0539	0.0348	0.00302	0.17
Locomotor Activity (22°C)	0.0294	0.024	0.00278	0.106
DEE (10°C)	0.667	0.328	0.207	1.49
DEE (22°C)	0.473	0.219	0.151	1.12
COA (10°C)	0.0371	0.0367	-0.0165	0.124
COA (22°C)	0.0539	0.0284	-0.0131	0.164
Substitution	0.0169	0.0174	-0.0261	0.0672
Substitution (Second Method)	0.0140	0.0179	-0.0354	0.0579

Table 3 Table 3 Among-individual variance (V_I), residual variance (V_e), and repeatability (R) in substitution **A** estimated from a linear mixed model (LMM) with all observations and only nuisance variables (i.e., metabolic cage and test sequence) included as fixed effects (consistency R), **B** restricted to observations without any missing value for the covariates (consistency R), **C** that included all the covariates listed in Table 4 (adjusted R) **D** estimated from a LMM using substitution measures derived from the second method (described in text) as the dependant variable (consistency R).

	V_I	SE	χ^2	P	V_e	SE	R	SE
A) All observations	0.306	0.134	7.65	**0.0028**	0.672	0.119	0.313	0.132
B) Complete cases analysis	0.22	0.119	4.7	**0.0151**	0.529	0.111	0.294	0.152
C) With covariates	0.0594	0.11	0.32	0.286	0.562	0.122	0.0957	0.177
D) Second Method	0.279	0.133	6.18	**0.0064**	0.723	0.128	0.278	0.132

Substitution is standardized to a mean of 0 and variance of 1.
Significance of random intercept estimate tested with LRT
Non-significant fixed effects (test sequence and cage) not shown
Significant P values (P<0.05) bolded

Table 4: Among-individual variance (V_I), residual variance (V_e), and consistency repeatability (R) in body mass, daily energy expenditure (DEE), cost of activity (COA), and mean locomotor speed in female white-footed mice estimated from separate linear mixed models with metabolic cage and test sequence included as fixed effects. Variance and repeatability for activity, DEE, and COA are presented separately for observations taken at 22°C and 10°C.

	V_I	SE	χ^2	P	V_e	SE	R	SE
Body Mass	0.913	0.203	106.81	**<0.001**	0.0969	0.0174	0.904	0.0264
Locomotor Speed (10°C)	0.609	0.162	38.9	**<0.001**	0.33	0.0591	0.649	0.0874
Locomotor Speed (22°C)	0.422	0.133	20.29	**<0.001**	0.437	0.0777	0.491	0.112
DEE (10°C)	0.0319	0.0569	0.33	0.283	0.469	0.0824	0.0636	0.116
DEE (22°C)	0.0741	0.0651	1.53	0.108	0.473	0.0831	0.136	0.121
COA (10°C)	0.0874	0.0850	1.14	0.143	0.609	0.108	0.126	0.126
COA (22°C)	0.203	0.1	5.58	**0.00907**	0.55	0.098	0.269	0.129

All traits are standardized to a mean of 0 and variance of 1.
Significance of random intercept estimate tested with LRT
Significance of fixed effects (test sequence and cage) not shown
Significant P values (P<0.05) bolded

Table 4: Summary of the linear mixed model that includes all possible covariate traits as fixed effects. Coefficient estimates and standard errors from the model are reported, as well as the *F*-test statistic, degrees of freedom (df), and significance.

Model term	N_{obs}	Units	Coefficient	SE	F (df)	*P*
Intercept			0.108	0.338	0.0778(1,18.4)	0.783
Total mass	117	g	0.219	0.141	2.09(1,31.5)	0.134
Age	117	Days	0.338	0.189	3.25(1,23.9)	0.0838
Torpor bouts	117	N/A	-0.135	0.294	0.21(1,63.5)	0.658
Hematocrit	88	Ratio	0.0639	0.109	0.345(1,62.2)	0.559
Body Length	37	cm	0.352	0.147	5.75 (1,22)	**0.0255**
Tail Length	37	cm	-0.201	0.119	2.86(1,19.9)	0.106
Surface Area	37	cm^2	-0.409	0.139	8.15(1,23.7)	**0.00879**
Tail mass	37	g	-0.286	0.156	3.35(1,26.9)	0.0781
Fur mass	37	g	0.107	0.154	0.483(1,17.7)	0.496
Skin mass	37	g	0.0387	0.127	0.0923(1,21.7)	0.764
Heart mass	36	g	0.342	0.148	5.31(1,27.7)	**0.029**
Liver mass	37	g	-0.259	0.114	5.18(1,19.4)	**0.0344**
Kidneys mass	37	g	-0.189	0.136	1.94(1,20.3)	0.178
Right gastrocnemius mass	37	g	0.0233	0.103	0.0512 (1,17.8)	0.824

All effects are standardized to a mean of 0 and variance of 1.
Torpor bouts: two-level factor (torpor bout present during test y/n)
Significance of fixed effects tested with a Wald test.
Fixed effects of Test & Cage not shown.
Significant *P* values ($P < 0.05$) are bolded, all significant values have a corresponding significant F-test statistic.
Hematocrit: ratio red blood cells to total plasma collected

Figures

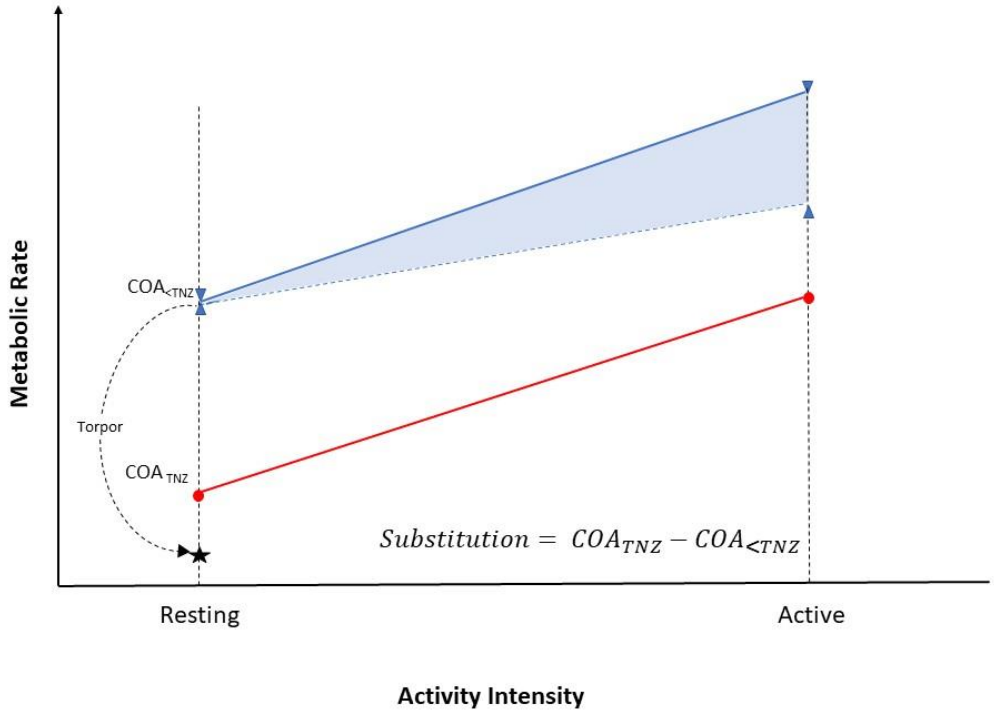

Figure 1 Hypothetical relationship between metabolic rate (MR) as function of activity intensity in an endotherm within and below the thermoneutral zone (TNZ). Substitution can be quantified as the difference between cost of activity (COA) within the TNZ (COA_{TNZ}, red line connecting dots) and below the TNZ ($COA_{<TNZ}$, blue lines connecting triangles). When COA_{TNZ} and $COA_{<TNZ}$ are parallel (solid lines), there is no substitution. When the $COA_{<TNZ}$ is shallower than COA_{TNZ} (solid red line vs dotted blue line), there is substitution, and the shaded area represents net energy savings from substitution. In some circumstances, a resting animal below the TNZ might enter torpor and therefore reduce MR below the cost of resting within the TNZ (black star), overestimating $COA_{<TNZ}$ and showing why substitution must be calculated using measurements on normothermic animals.

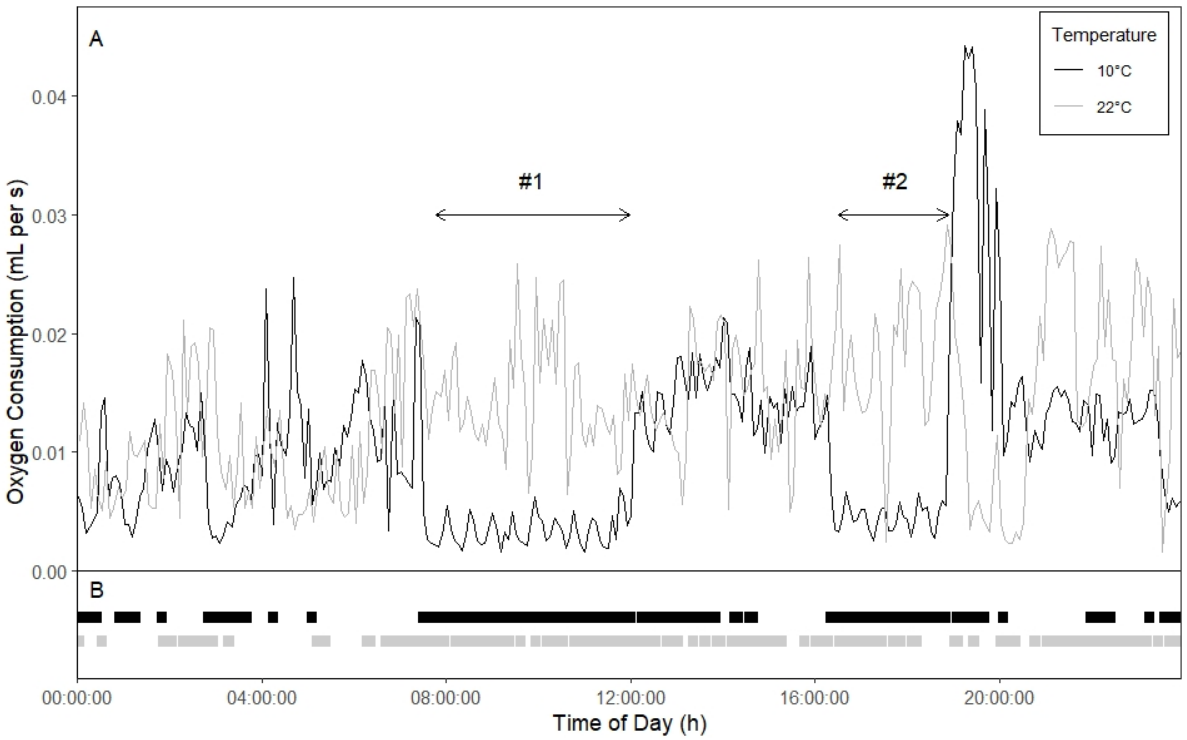

Figure 2 A Oxygen consumption (mL per s) and **B** inactivity (solid bars) as function of time of day (hours) for a single individual (mouse #23512, test #1), at 10°C (black line and bars) and 22°C (grey line and bars). Time has been limited to one 24-hour period within the test for both temperature conditions to better illustrate torpor bouts #1 (8:00 to 12:00) and #2 (16:00 to 19:00 hours, where oxygen consumption for the same animal at rest was lower at 10°C than 22°C for the entire duration of the torpor bout. Also note how torpor bout #2 is immediately followed by a marked increased in oxygen consumption despite inactivity, which is presumably reflects the re-warming phase.

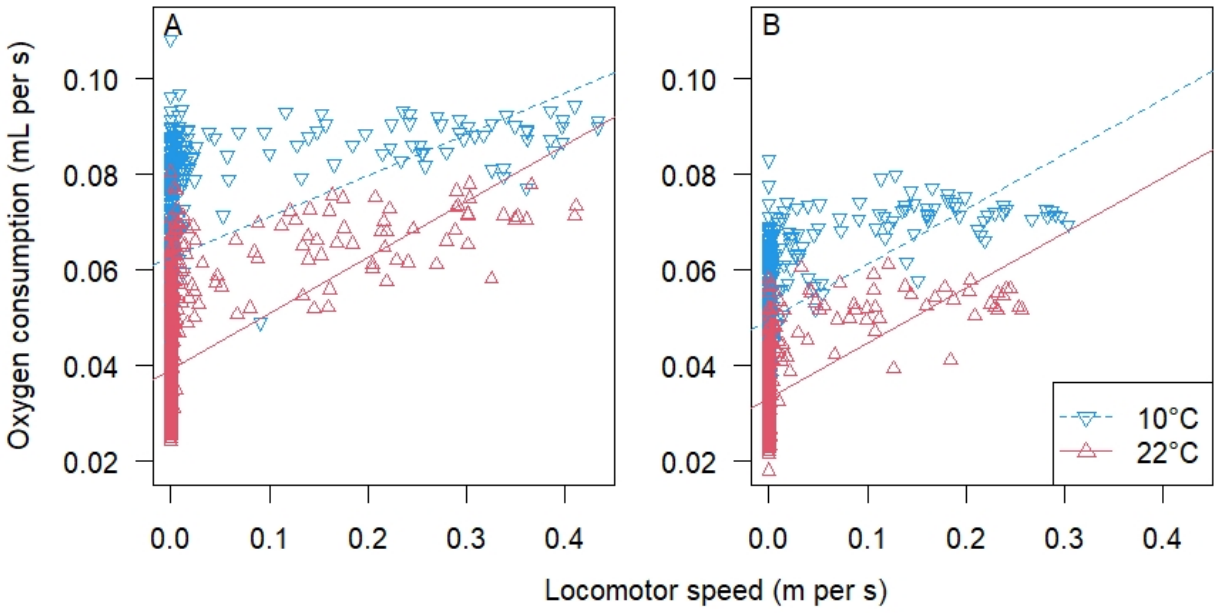

Figure 3 Oxygen consumption (mL per s) as a function of voluntary wheel-running and home-cage locomotory speed (m per s) at 22°C (red upper triangles and solid line) and 10°C (blue lower triangles and dashed line) for **A** a single mouse exhibiting heat substitution (mouse #23517, test #2) and **B** another mouse showing no substitution (mouse #23501, test # 1).

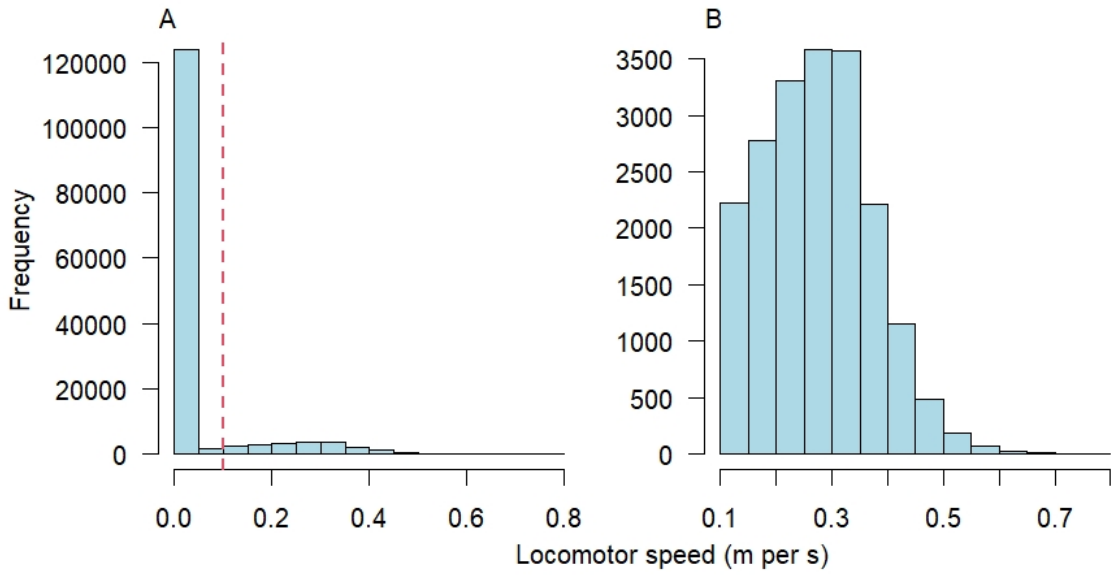

Figure 4 Frequency distribution of locomotory speed (m per s; as the sum of voluntary wheel-running and home-cage locomotion) observations for **A** the whole population across all tests and **B** speed observations greater than or equal to 0.1 m per s (i.e., above the dashed line in **A**, the threshold above which a mouse was considered "active" in this study). Mice that had fewer than 5 observations above the threshold for each temperature condition for each test were excluded from analysis.

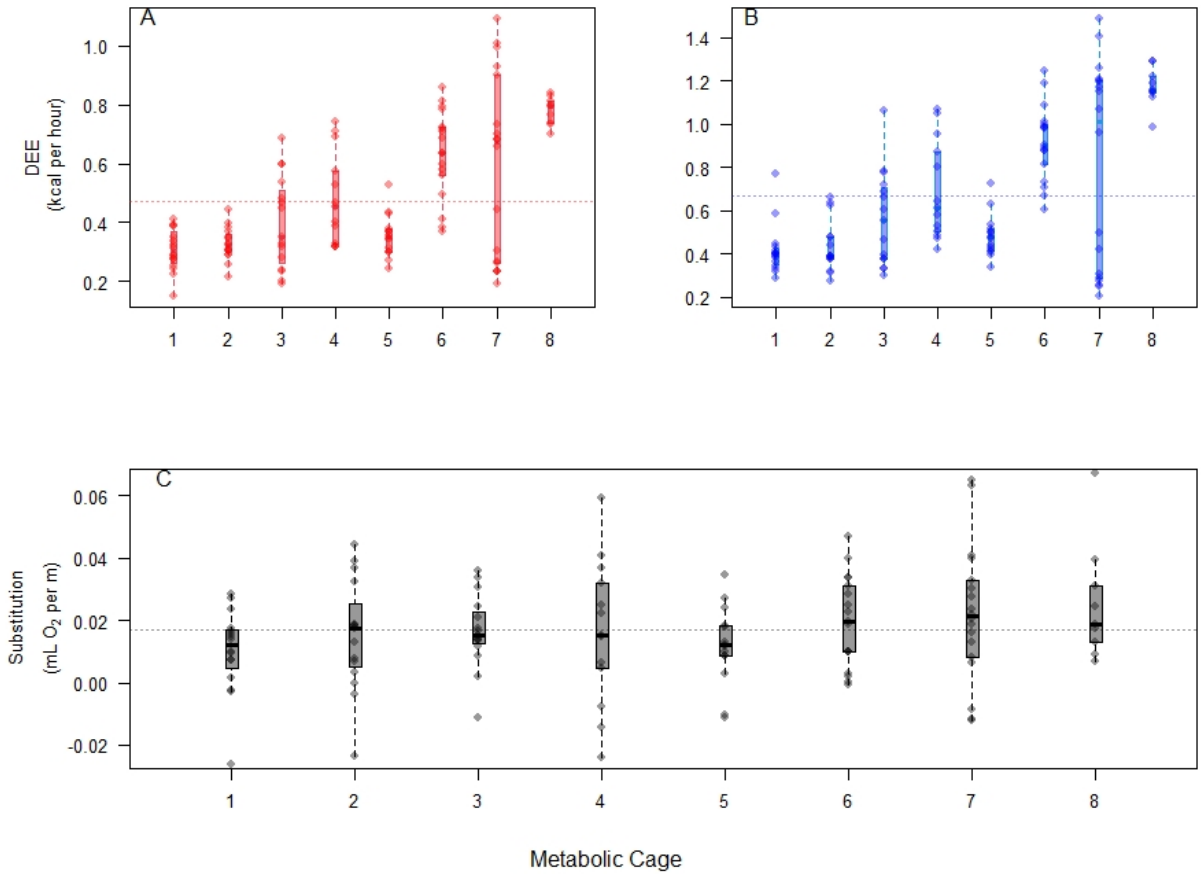

Figure 5 A Daily energy expenditure (DEE, kcal per h) at 10°C, **B** DEE at 22°C, and **C** substitution (mLO_2 per m) as a function of metabolic cage (1 through 8). Data points are used to indicate individual tests and the dotted line indicate the population average in each graph. Note that for DEE, cage effects are similar between A and B, but nonexistent in C.

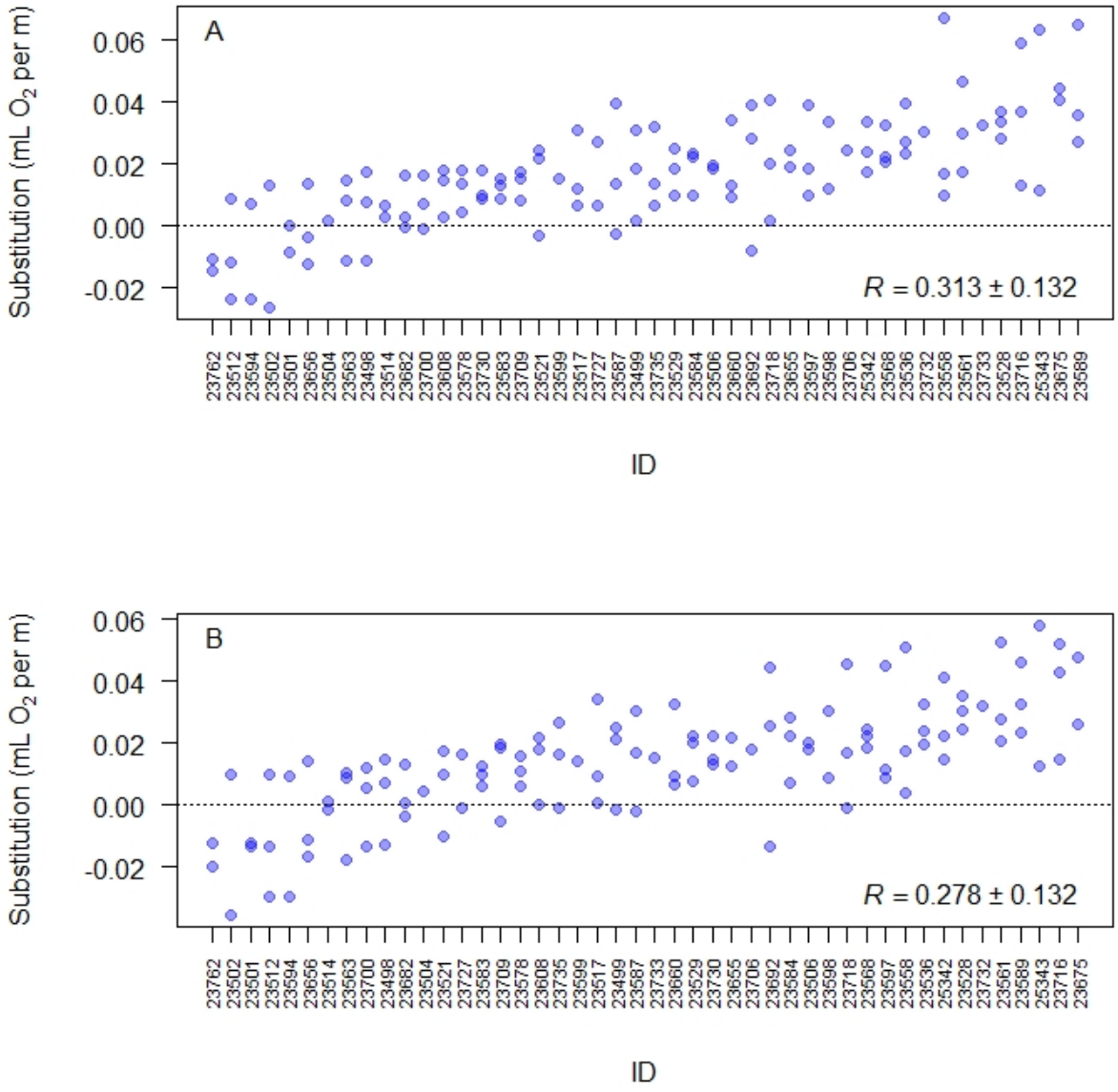

Figure 6 Among- and within-individual variance in substitution in 46 white-footed mice, ordered on the x-axis from lowest to highest average substitution **A** derived from model coefficients (see Table 1) and **B** derived from the second method (see text) . Consistency repeatability estimate is indicated for substitution ($R \pm$ SE)

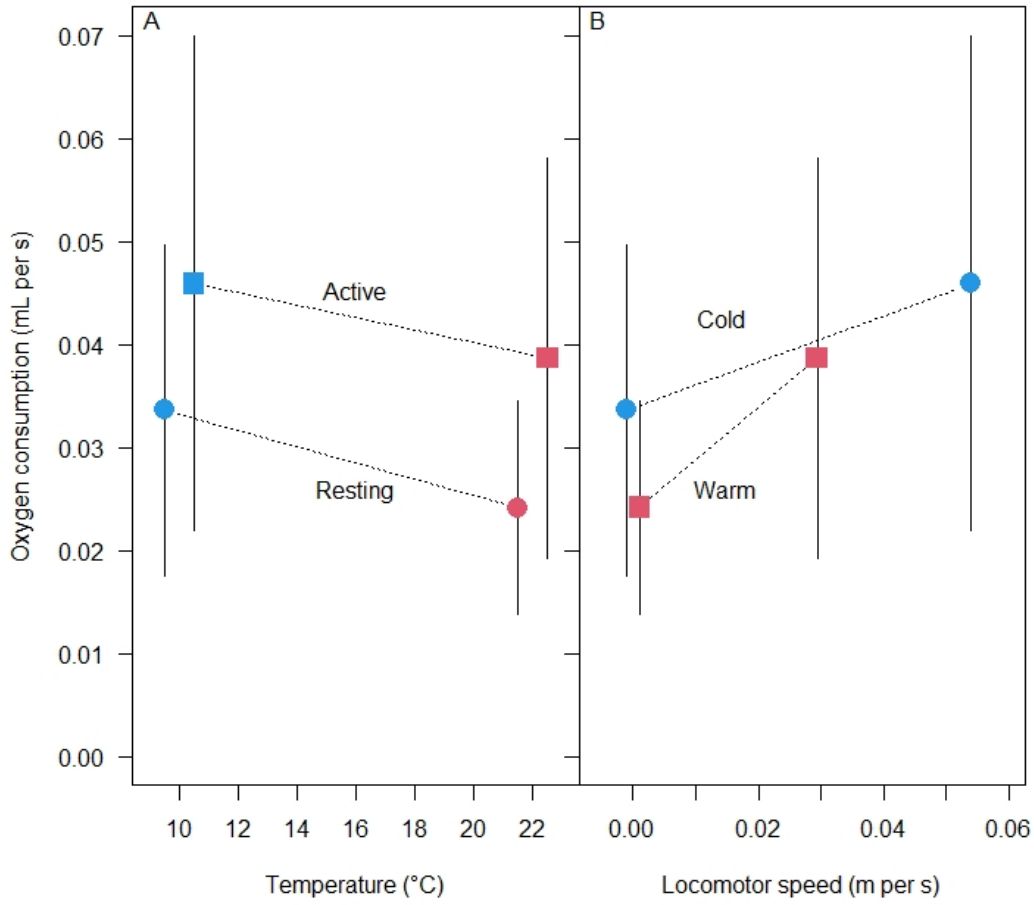

Figure 7 Average oxygen consumption (mL per s) as function of **A** ambient temperature and **B** mean locomotor in resting (speed = 0 m per s, solid circles) and active (speed < 0.1 m per s, solid squares) white-footed mice at 10°C (blue symbols) and 22°C (red symbols). Values are the population means ± sd for each condition across all tests. Substitution is not apparent in **A** when comparing the change in oxygen consumption between the two temperature conditions in active vs resting mice, but becomes obvious when considering that average locomotor speed was much higher at 10°C than 22°C.

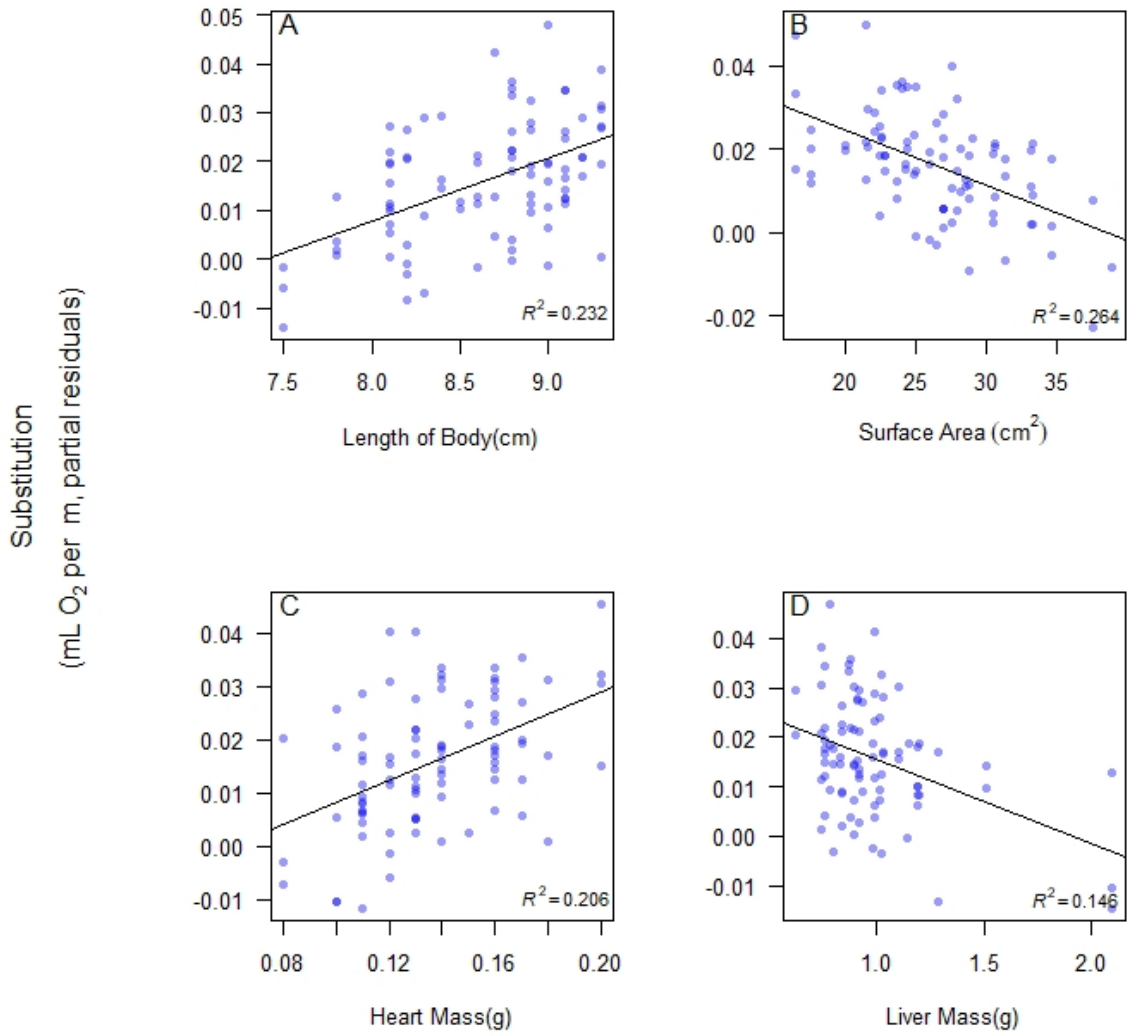

Figure 8 Partial residuals of substitution (mLO_2 per m) derived from the full linear mixed model as a function of statistically significant covariates (See Table 4). Solid lines indicate the line of best fit. Adjusted R^2 for the covariate and the partial residuals is indicated.