# Title: Individual variation in thermal plasticity and its impact on energetic scaling across different hierarchical levels

Fonti Kar1, Shinichi Nakagawa1,2, Christopher Friesen3, Daniel Noble1,2,4

1 *School of Biological Earth and Environmental Sciences, Ecology and Evolution Research Centre, University of New South Wales, Sydney, NSW, Australia*

2 Garvan address

3 UWollogong

4 *Ecology, Evolution and Genetics, Research School of Biology, The Australian National University, Canberra, ACT, Australia*

# Abstract

Rough outline:

* Metabolic rate, an intrinsic property of an individual, can limit many biological processes from individual to populations and ecosystems
* An individual’s energetic expenditure is sensitive to changes in the internal (e.g. body mass) and external environment (temperature), but whether this metabolic plasticity is consistently expressed or not is well understood.
* Variation in metabolic rate is organised hierarchically. If this phenomenon is not properly accounted for, individual processes can have overpouring effects on higher level energetic scaling relationships.
* We repeatedly measured metabolic reaction norms in an ectotherm model (*Lampropholis delicata* – the delicate skink) to characterise the repeatability of metabolic plasticity over a temperature gradient. We tested whether mass variation at the within-individual level can affect population level mass-scaling exponents and whether these scaling exponents change with temperature.
* Using function-valued and character state approaches, we found that the slope of the metabolic reaction norm is significant repeatable, and that repeatability of metabolic rate increased as a function of temperature. This change in repeatability was associated with individuals responding more predictably as temperatures got hotter.
* We also found that metabolic rate between temperatures that were more similar to be strongly correlated compared to temperatures that were more distinct under the character-state approach. This has evolutionary implications on how the shape of the reaction can evolve as it may depend on what statistical approaches researchers employ.
* Our results show that mass-scaling exponents at both among- and within- individual levels are temperature dependent and the sample population scaling exponents are underestimated when within-individual variation in mass is not properly accounted for.
* The proximate and ultimate causes of individual variation in metabolic plasticity are discussed

# Introduction

Animals live in a multifaceted world, where many aspects of the phenotype are highly responsive to the environment including life history ({Westneat:2009dz}), locomotor performance ({Careau:2014in}), behaviour (reviewed in{Dingemanse:2010bk}) and physiology ({Boratynski:2017jf}). Metabolic rate, in particular, is a fundamental measure in ecology and evolution as it determines an individual’s energetic capacity for competing processes such as growth, reproduction and somatic maintenance ({Brown:2004hp, DeJong:1992dr}). Metabolic rate is likely a target for selection because it has been shown to drive variation in ‘paces-of-life’ ({Reale:2010ef, Biro:2010ee, Malishev:2017ef}, {Robert:2010bs}). Furthermore, an individual’s metabolism contributes to the flow of energy, biomass and nutrients in the population which can have cascading impacts at the community and ecosystem level ({Allen:2005fs}{Barneche:2014jg}). Given the ecological and evolutionary significance of energy metabolism, there is a growing interest in the proximate and ultimate causes of variation in metabolic rate, as well as its sensitivity to biotic and abiotic factors ({Pettersen:2016fi, Norin:2016fo}). The degree to which metabolic rate changes with, for example, temperature and body mass, can be highly variable among vertebrate taxa ({Uyeda:2017jn}), populations of the same species (Wikelski et al., 2003, {Burton:2011fe}) as well as, individuals within the same population ({Norin:2018ba}). Understanding the link between the environment and metabolic rate across different hierarchical levels of biological variation can help elucidate eco-evolutionary dynamics such as species diversification {Glazier:2015fr, Brown:2004hp}.

Metabolic theories attempt to unify ecological processes across populations, communities and ecosystems based on physiological rates at the individual level ({Brown:2004hp, Gillooly:2001cg}). These theories rely on first principles of biology, chemistry and physics to explain the scaling of energetic demands with body mass. From individuals to ecosystems, metabolism is thought to exhibit a fixed scaling relationship with body mass (i.e., 3/4 or 2/3 power laws) and changes with temperature following a logarithmic function ({Brown:2004hp, Gillooly:2001cg}). This generalisation to higher levels of biological organisation may be overly simplistic because physiological mechanisms that govern metabolic scaling at the individual level are highly complex and variable ({Glazier:2005ei, White:2012ip}). For example, many explanations of a universal mass-scaling exponent are based on the assumption that supply of resources such as macronutrients to the metabolic machinery is constant but this is known to vary between individuals (Reviewed in {Glazier:2005ei}, {Metcalfe:2005tw}, {Speakman:2004fk, Steyermark:2005bx}).

Accumulating evidence challenging the ‘one-size-fits-all’ line of thinking has shown that mass scaling exponents are influenced by numerous factors. Extrinsic factors such as resource availability, in addition to intrinsic differences between species such as endo-/ ectothermy can interact to influence how individuals respond to the environment which impacts metabolic scaling ({Barneche:2016ke, White:2012ip},{Uyeda:2017jn}, {Clarke:2004fv}, {Killen:2010cw}, {Glanville:2006eo}). Interspecific variation in scaling exponents may have true biological meaning, but it may also be due to an ‘ecological fallacy’ where the energetic scaling relationships described at one level of variation is incorrectly attributed to higher levels (Van de pol, 2009). In other words, scaling relationships at the intra-individual level can impact population and species-specific estimates. If the goal is to establish an explanatory link across large scales of biological organisation, it is in our best interest to correctly account for hierarchical variation in metabolic rate. While the mechanisms driving interspecific variability in energetic scaling remains elusive, variation in metabolic plasticity across individuals may provide important insight to our understanding.

It is known that individuals consistently express different ‘metabolic norms of reaction’ and these differences are ascribed to individual variation in physiology. The physiological system that underpin the relationship between energetic demands and body mass are undoubtedly sensitive to environmental conditions. This mechanistic link may be a promising explanation of why metabolic scaling exponents are so diverse from populations to ecosystems. For example, membrane composition ({Hulbert:2007eh}), enzyme structure and function ({Somero:1978wh})

and mitochondrial capacity ({Salin:2012cn}) are all recognised to adjust to changes in the environment ({Seebacher:2005jt}). For example, in lizards, dietary fats can change the composition and fluidity of mitochondrial membranes, which in turn affect the oxidative capacities of the whole metabolic machinery {Simandle:2001eu}. Currently, it is unclear whether individual-by-environment interactions (I x E, {Nussey:2007bz}) could affect population-level scaling exponents, which could have carry-over effects to higher level scaling exponents. Despite the importance of individual variation in metabolic plasticity for explaining variation in metabolic scaling, repeatable metabolic reaction norms has only been reported in only a few species ({Briga:2017dr}, {Careau:2014bm}). In order to test these ideas we need to first characterise the repeatability of metabolic plasticity, which is tedious and require repeated measurements of individual responses across an environmental gradient over a relevant time-scale.

Part of the challenge in quantifying the repeatability of metabolic plasticity is that there are multiple ways in which plasticity itself can be modelled. Two approaches in modelling plasticity are prevalent and debated in evolutionary biology ({Via:1995hm}, {Hunt:2014wo}). The character-state approach models phenotypic change in a set of environments as discrete ‘states’. Under this method, selection pressures can vary between environments and the mean phenotypic values in each environment are subjected to selection. Whereas in the function-valued approach (also known as ‘polynomial’ approach, Via et al 1995), phenotypic changes in a trait across an environmental range is described by a mathematical function. In this circumstance, selection pressure is assumed to be equal across all environments and model parameters e.g. intercept and slope are the main targets of selection. Given that the same phenotypic trait measured in multiple environments are inherently correlated and may give rise to evolutionary constraints on the reaction norm ({FALCONER:1952uz}). Non-zero covariances between character states in different environments, and between the intercept and slope can dictate the extent to which these must evolve in tandem {Hunt:2014wo}. While the conceptual differences between the modelling approaches have sparked debates, both approaches can contribute to our understanding on how the shape of reaction norms can evolve. To the best of our knowledge, no study has assessed the merits of both approaches in understanding individual variation in reaction norms.

Here we examine how individuals vary in their metabolic rate in relation to body size and thermal plasticity using an ectotherm model, the delicate skink (*Lampropholis delicata*). We take advantage of both function-valued and character state approaches to answer the following questions about individual variation of reaction norms . (1) does mass scaling exponents differ at the among- and within- individual level? (2), do these mass scaling exponents change with temperature? (3), do individuals consistently differ in their plastic responses to temperature (i.e. thermal reaction norms) over time? (4), what are the cross-environment correlations of metabolic rate between different temperatures?

# Methods

## Lizard collection and husbandry

Forty-two male *L. delicata* were collected across two sites between 28 August and 8 September 2015, across the Sydney region. Lizards were caught by hand or by mealworm fishing and were transported individually in calico bags in an ice-cooler to Macquarie University. Lizards were housed in a temperature controlled room and was provided with a thermal gradient. Each lizard was kept individually in an opaque plastic enclosure measuring 35cm 25cm 15cm (L W H). Each enclosure was lined with newspaper and lizards were given access to a water bowl and tree bark as a refuge. Enclosures were placed under UV light (11L:13D). Lizards were fed three to four small crickets (*Acheta domestica*) dusted with calcium powder and multi-vitamin every two days when metabolism measurements were not taking place. An animal collection license was approved by the New South Wales National Parks and Wildelife Service (SL101549). All procedures were approved by the Macquarie University Ethics committee (ARA 2015/015) and University of New South Wales Animal Care and Ethics committee (ACEC 15/51A).

## Quantifying metabolic reaction norms

Metabolism assays were conducted between 26 December 2016 - 19 March 2017. We measured metabolic rate as CO2 production per unit time (VCO2 ml min-1) given that CO2 is often more accurate for small animals (RFE). Animals were measured at a random temperature in an inactive, post-absorptive state (REF). Overall, measurements were taken at temperatures between 22ºC and 32ºC at 2ºC increments over a three day period (measurements at two temperatures per day) and were repeated every 10 days (10 sampling sessions in total). Due to logistical constraints, lizards were randomly assigned to one of two blocks for metabolism measurements (block 1: n = 23, block 2: n = 22). We used two incubators (LabWit, ZXSD-R1090) to precisely control the ambient temperature at which measurements were taken (+/- 1ºC). In order to account for the carry over effects of extreme temperatures experienced by an individual on subsequent metabolic measurements at other temperature, the temperature order was randomly allocated to the incubators across the three days, within each sampling session.

Lizards were fasted for at least 24 hours prior to metabolic measurements, during which they had access to water in their enclosure between each sampling day. Within each sampling session, lizards were randomly assigned to opaque cylindrical metabolism chambers (volume = 146ml). On the day of measurements, body temperature of each individual was taken using an infrared laser gun) inside their enclosure (Stanley stht0-77365. Lizards were then gently encouraged into their assigned chambers and then weighed using a digital scale to the nearest 0.01g (Ohaus SP-202). After which, opened chambers were maintained in a dark environment inside the incubators for 30 minutes for animals to settle down. After 30 minutes, each chamber was flushed with fresh air and then sealed. 3 mL of air was then immediately removed from each chamber via a two-way valve to account for any residual CO2 that was not flushed out from the chambers. The chamber was then left in the incubator for an additional 90 minutes at the set temperature to measure CO2 production. At the end of the 90 min, two air samples were taken from each chamber and chambers were reopened and placed back into the incubator again for the next measurement temperature following the same procedure.

After air sampling, each 3 mL air sample was injected into the inlet line of an open-flow respirometer system (Field Metabolic System Sables System, Las Vegas NV, USA) to measure VCO2*.* The percentage of CO2 in our samples was measured using a flow rate set to 200 ml min-1. Water vapour was scrubbed from the inlet air with Drierite (Brand) prior to the measurement of CO2. The area of output peaks was integrated using LabAnalyst to calculate the percentage of CO2 (REF FOR LAB ANALYST). The total volume (ml) of CO2 produced by an individual was calculated as:

Eqn: 1

where %CO2 is the percentage of CO2 in air sample, which was corrected by subtracting any ‘residual’ CO2 from the initial flush from the larger of the two air samples, Vchamber is the volume of the chamber (146ml), Vlizard is the volume of the lizard, assuming that the mass of the lizard is the same as its volume and *t* is the duration of time in minutes after where the chamber has been sealed and the first air sample was taken (90 minutes).

## Statistical analysis

All statistical analyses were conducted using ‘R’ {TeamRAla:2013tx}. All temperature variables and VCO2 were log-transformed. Collinearity between our response (log VCO2 and predictor variables were checked using scatterplot matrices. Pearson correlation coefficients of these are presented in Table SXX. There were no differences in VCO2 between blocks of lizards or incubators in our initial models therefore these parameters were not included in our final models. We tested whether the previous measurement temperature or body temperature measured in home enclosure influenced VCO2 at subsequent temperatures due to acclimation. We found that a model containing ‘previous temperature experience’ as a covariate was better supported compared to a model with it excluded, we therefore included ‘previous temperature experience’ in all subsequent analyses.(See ESM for more details). For all our models we used Bayesian linear mixed models (LMM) from either the ‘brms’ or ‘MCMCglmm' R packages (REFS), ‘brms’ uses Stan, whereas ‘MCMCglmm’ uses Monte Carlo Markov Chain to sample from the target posterior distribution. For every model, we pooled the posterior estimates for multiple chains. For ‘brms’ models, we ran for 4 chains of 2000 iterations with a burn in of 1000 and a thinning interval of 1. For ‘MCMCglmm’ models, we ran 3 chains of 7510000 iterations with a burn in of 10000 and a thinning interval of 5000. We checked all chains were mixing well and converged by visually inspecting trace plots and also ensuring Rhat values are greater than 1.1. We also checked whether our chains were autocorrelated.

### Repeatability of thermal reaction norms

We used function-valued and character state approaches to quantify the repeatability of thermal reaction norms to determine whether they differed consistently among individuals. In all out repeatability models, body mass was first log-transformed and then z-transformed to account for the allometric scaling relationship between metabolic rate and body mass (ref Nakagawa et al 2018). For function-valued models, we coded ‘series ID’ - a categorical variable which denotes a unique combination of individual IDs and the sampling session IDs. This enabled us to account for variation in individual reaction norms between sampling sessions.

#### Reaction norm intercept – average response

Repeatable intercepts tell us whether individuals’ average metabolic rate vary consistently at a particular temperature. First, we used function-valued models to estimate the repeatability of the intercept of individual reaction norms. This class of models assumes that the intercept covaries with the slope of the linear reaction norm. The intercept represents the average trait value when log temperature = 0 (i.e. 1ºC), however this can be set at a biologically relevant temperature by mean-centering. Using MCMCglmm, we fitted six models to estimate the repeatability of the intercept at each measurement temperature where,

logVCO2 = logTempcen + zlogBodyMass + logPriorTemp + (ID+ logTempcen) + (Series+ logTempcen)

where logVCO2 log-transformed VCO2, logTempcen is the mean-centered log temperature in degrees Celsius at a given measurement temperature so that the intercept represents the average response at given log measurement temperature (i.e. log(22ºC) = 3.09, log(22ºC)centered at 22ºC = 0, log(24ºC)centered at 22ºC = 0.09, log(26ºC)centered at 22ºC = 0.17, etc), zlogBodyMass is log-transformed body mass that is then subsequently z-transformed (mean of 0 and sd of 1), logPriorTemp is log-transformed previous temperature. We fitted individual IDs as a random intercept and logTempcen as a random slope (i.e. (ID+ logTempcen)) and series ID and logTempcen as a random slope (i.e. (Series+ logTempcen)). For each of these models, we calculated adjusted repeatability (eqn 2) of the intercept VCO2 at each measurement temperature following {ArayaAjoy:2015ir} using the entire posterior distributions of the relevant variance components:

Eqn 2:

where is the repeatability estimate for logVCO2 at a particular temperature; *Vind0* is the individual intercept and *Vseries0* is the series intercept. To assess how repeatability of the intercept changed over the course of our study, we also calculated ‘short term’ and ‘long term’ adjusted repeatability according to {ArayaAjoy:2015ir}.

Eqn 4:

Eqn 5:

where *Ve0* is the residual variance. ‘Short-term’ repeatability can be interpreted as among-individual variation that includes both intrinsic differences between individuals as well as the effects of the sampling session on individuals. In contrast, long-term repeatability is a more conservative measure and represents repeatability in the classical sense, where phenotypic variation due sampling session is a part of the total pool of variation in the data (i.e. in the denominator of the calculation) (Nakagawa & Schielzeth, 2010).

We also used another method to derive ‘conditional’ repeatability (eqn 6) for intercept values at each temperature following Singer and Willet (2003) and Briffa (2013). We fitted one MCMCglmm function-valued model where logVCO2 = logTemp + zlogBodyMass + logPriorTemp + sampling session + (ID+ logTempcen). This allowed us to derive temperature-specific repeatability estimates using the covariance of the intercept and slope at each temperature.

Eqn 6:

where *Covind0,ind1* is the covariance of the individual intercept and individual slope and *T* is the measurement temperature at which repeatability is estimated. Note that the temperature variable must be on the same scale as the temperature variable with which the intercept and slopes variance components were estimated because the and are estimate at that scale (i.e. logVCO2, see {Brommer:2013gx}).

The character-state approach can also be used to calculate temperature-specific adjusted repeatability estimated. A character state model estimates variation of logVCO2 at each temperature which requires a substantial amount of data points ({Hunt:2014wo} and Houslay & Wilson (2017)). In other words, a character-state model estimates ‘intercepts’ at each temperature in a single run by fitting a multivariate response matrix and assumes that traits are “independent” but correlated with each other in a way that is estimated by the model:

Eqn 7:

where each row represents each individual in a particular sampling session and each column and each column represents logVCO2 measured at each temperature. For example, is the log metabolic rate measured at 22ºC for individual 1 in sampling session 1. Using brms, we fitted the same predictors as our function-valued models, note that log temperature is no longer a predictor as this is now part of the response matrix. We fitted individual ID and series ID as random intercepts. The character-state model estimates variance components and covariances for each temperature at the individual and series level and these can be used to calculate temperature specific adjusted repeatability following Nakagawa, S., & Schielzeth, H. (2010).

Eqn 8:

where refers to the adjust repeatability at a given temperature; *Vind0,T* , *Vseries0,T*  and *Ve0,T*

represents the individual intercept, series intercept and residual variance component at a given temperature.

#### Slope – plasticity

Function-valued approaches uses the slope of a linear thermal reaction norm to represent the plasticity of a trait over a temperature gradient. If slopes are repeatable, this would indicate that individuals differ consistently in how they respond to temperature over time. We used one of the previously fitted function-valued models to calculate the repeatability of the slope, (logTemperature centered at 22ºC).

Eqn 9:

where *Vind0* is the individual intercept and *Vseries0* is the series intercept. *Vind1* and *Vseries1* is the individual slope and series slope, respectively. While the character state approach may make less stringent assumptions on how the phenotype can change across an environment, identifying changes in repeatability across different environments using RT (eqn 8) is conceptually the same as quantifying whether slopes are repeatable or not.

### Cross-temperature correlations in metabolic rate

Metabolic rate measured at one temperature will be undoubtedly correlated with metabolic rate measured at another temperature. We estimated these cross-environment correlations using both statistical approaches. While correlations can be estimated directly from the character-state model at both the among and within individual level (by setting rescor = T in brms or rcov = ~us(trait):units in MCMCglmm), these are not directly estimated from function-valued models as they are part of the general slope describing the reaction norm. However, correlations can be calculated from function-valued models using matrix algebra and the variance-covariance matrix of the intercept and slope following Brommer (2013). We derived correlations from a function-valued model at the among and within-individual-among-sampling-session level (i.e. series) where we fitted logVCO2 as a response and same as previous models, logTemp, zlogBodyMass and logPriorTemp were included as predictors. Individual IDs and series were included as a random intercept and logTemp as random slope.. The variance-covariance (*K*) from this model is therefore denoted as:

Eqn 10:

where or is the covariance of the slope and intercept. The six measurement temperatures can be represented as a double column matrix with six rows,

Eqn 11:

where the first column contains ones and the second column is the six unique measurement temperatures on the same scale as the predictor used to estimate *K* . The among individual variance-covariance matrix, *P,* for the six temperatures can then be derived by multiplying K with and its transpose,

Eqn 12:

which results in a 6 x 6 variance-covariance matrix as follows,

where the diagonal , , are the among individual variances in log metabolic rate at all six temperatures and the off-diagonals represent the pairwise covariances of log metabolic between all six temperatures. These cross-temperature covariances can then be scaled to correlations by dividing the covariance between two temperatures by the square-root product of the variance in each of the two temperatures (i.e. the standard deviation at each temperature)

Eqn 13:

### Hierarchical mass scaling exponents at different temperatures

We used a ‘brms’ LMM to estimate within- and among-individual effects on mass-scaling exponents at all six measurement temperatures. This allowed us to test whether exponents were temperature dependent and whether scaling relationships at each temperature were affected by within individual variation in body mass. Individual mass changes throughout the course our experiment, the mean change in mass (mass at the first sampling session – the last sampling session) was 0.12g (n = 42, SD = 0.09, range = -0.04, 0.32).

We calculated the average mass across all measurement days and sampling sessions for each individual (among-individual effect). In order to obtain a within individual effect, we calculated for each individual, for each measurement day, deviations of mass on given day, in a given sampling session from the individuals mean (i.e. within-individual centering, see Van de Pol 2009). We fitted a model with logVCO2 as a response and included an interaction term between the among individual mass effect with temperature and another interaction term between the within individual mass effect with temperature. We subtracted the within individual estimates from the among individual estimates to get the difference to test whether they were significantly different from zero or not. We included individual ID as a random intercept and the within subject effect as a random effect since exploratory graphs show that individuals mass change at different rates through the study. We compared the within- and among individual scaling exponents with another model that doesn’t account for the hierarchal structure in the data. In this model, we fitted logVCO2 and included an interaction term between log body mass with temperature.

# Results

## Repeatability of thermal reaction norms

Overall, repeatability in log VCO2 increased with temperature (Fig. 2A). The intercept at all temperatures were repeatable over short and long temporal scales (Table XX). The FV- Singer & Willet approach showed highest repeatability at 22ºC, while the FV – Yimen approach found repeatability was highest at 30ºC. In contrast, the CS approach found repeatability highest at 30ºC. Upon closer inspection of the variance components at each temperature, within individual variation decreased over the temperature gradient, whereas among individual variation remained relatively consistent and only increased slightly with temperature (Fig 2B). In other words, individuals were responding more consistently as temperatures became hotter and there was a very slight increase in the between individual variation, explaining the higher repeatability. Congruent with the change in repeatability with temperature, individual slopes were significantly repeatable (Rslope = 0.48, 0.06 - 0.91) indicating a significant individual by environment interaction(I x E) that was consistent over time (Fig. 3).



**Figure 2** – A) Posterior mean of repeatability of log VCO2 at six measurement temperatures. Orange circles represent function-valued approaches in modelling thermal plasticity, orange open circles (○) represent the FV – Singer & Willet method, orange filled circles (•) present the FV – Yimen method. Blue filled circles (•) represent the character state approach in estimating repeatability. See Statistical analyses for more details. Error bars represent 95% credible intervals. B) Posterior mean of variance of log VCO2 at the among (right panel) and within (left panel) individual level across six measurement temperatures. Error bars represent 95% credible intervals.

**Figure 3** – A) Predicted individual reaction norms of log VCO2 across six measurement temperatures estimated using function-valued approaches (e.g., using FV – Yimen). B) Predicted individual reaction norms of log VCO2 across six measurement temperatures estimated using a character-state approach. Points represent predicted trait values. Each line represents a unique individual (n = 42) at sampling session one (left panel), five (middle panel) and ten (right panel).

## Cross-temperature correlations in metabolic rate

Metabolic rate across temperatures were positively correlated at the among individual level (Fig. 4). Certain individuals maintained a high metabolic rate, while others had a relatively low metabolic rate across all temperatures (Fig. S1). This creates a positive relationship between metabolic rate at the among individual-level across different temperatures. Metabolic rate measured at neighbouring temperatures (e.g. 22ºC and 24ºC) are strongly correlated, but the strength of this correlation decreased with increasing differences between the two temperatures (Fig. 4).

We detected a similar positive correlation pattern at the within individual level (Fig 4), although the correlations were weaker. Overall, cross-temperature correlations estimated using the FV approaches were a lot higher and the credible intervals were very narrow. The correlations estimated using FV approaches are congruent with the CS approach however, estimates differ in magnitude and credible intervals may not be estimated accurately.



**Figure 4** – Cross-temperature correlations of metabolic rate estimated using the function-valued approach (top) and the character-state approach (bottom) at the among-individual level (left) and at the within-individual level (right). Lower triangle represents posterior mean estimates, width and colour of the ellipse represents the strength of the correlation. Asterisks indicate that the correlation estimate is signficantly different from zero, dashes indicate that the credible intervals could not be estimated correctly i.e. exceeds r2 = 1

## Multilevel mass scaling exponents and temperature dependence

Among- and within- individual mass scaling exponents were significantly different from each other between 22ºC and 28ºC (Fig. 2A). Overall, within individual exponents were higher compared to among individual exponents. There was a trend for exponents to be underestimated when the within- and among-individual effects were not statistically accounted for. The within and among individuals exponents were the most different at 24ºC (Difference = XXXX, lower = XXXX, upper = XXXX, Fig. 2A).



**Figure 5** – A) Posterior mean estimates of three types of mass-scaling exponents across six measurement temperatures. Within individual scaling exponents (Black square ) describes the change in log VCO2 as an individual’s mass changes on the logarithmic scale. Between individual scaling exponents (Black triangle ) describes the change in log VCO2  across individuals as mass changes on the logarithmic scale after accounting for within individual effects. Similarly, the ‘average’ individual mass scaling exponents (White circles) represents the change log VCO2  across individuals as mass changes on the logarithmic scale, however it *does not account* for within individual variation in mass changes. The grey diamonds ♦ represent the difference between the between and within individual scaling exponents. The dashed line represents the exponent of 0.83 estimated for squamates from Uyeda (2017). The dotted line represents 0. Error bars represent 95% credible intervals. B) Raw log VCO2 plotted against log body mass for a random subset of 15 individuals across six measurement temperatures. Each uniquely coloured point represents one individual. Thick bold line represents the change in log VCO2  over log body mass across all individuals. Faint grey lines represent the change in log VCO2  over log body mass within an individual.

# Discussion

Our results show that metabolic rate, and how it responds to the environment, consistently differs across individuals over short and long time scales. We demonstrate that, by not accounting for the hierarchical structure of population-level data (i.e., disassociating within and between-individual variance), mass-scaling exponents can be underestimated and can change with temperature. More specifically, within-individual exponents can be significantly higher than among-individual exponents which suggests that an individual’s energy expenditure increases disproportionately as body mass increases. Metabolic rate was positively correlated across all temperatures. However, the strength of these cross-temperature correlations was not uniform across all temperatures and differed between the character-state and function-valued approach. Below we discuss the implications of our results on understanding how plasticity may evolve and how metabolic rate scales at different hierarchical levels.

## Consistent variation in thermal reaction norms

Consistent among-individual variation is a key prerequisite for any trait to evolve and sets the ‘upper limit of heritability’ because it is the raw material that natural selection acts on (Falcon and Mackay 1996, see Dohm 2002 for exceptions). Our findings show that metabolic plasticity (i.e., the slope of metabolic reaction norms) was significantly repeatable over time. Repeatability of metabolic rate increased as a function of temperature owing to the changes in the relative contributions of among- and within- individual variance components. Interestingly, variation among individuals moderately increased with temperature, but individual metabolic rate was also more predictable relative to their own responses. Individuals may have reached their physiological ‘ceiling’ at high temperatures and was therefore respiring more predictably. However, this is unlikely given that we were not measuring maximal metabolic rate (Biro et al 2018). Instead, catabolic and anabolic processes may be at equilibrium at warmer temperature which could promote consistency within individuals (Somero 1978). Indeed, 32ºC is well within the range of preferred temperatures of this species where biochemical activities are likely to be operating optimally (meta digitise white paper, {Goulet:2016dt, Merritt:2013cb}). The compounded effect of high among-individual and low within-individual variation in hot environments may mean that, not only is there a greater opportunity for selection in hot thermal environments, but selection can operate more effectively (Cleasby and Nakagawa, 2014, Janicke et al 2016). This may facilitate adaptive evolutionary change in the population metabolic reaction norm, particularly in thermal environments that are novel to the population {Ghalambor:2007bc}.

## Cross-temperature correlations of metabolic rate: Implications of different modelling approaches for understanding metabolic plasticity

Metabolic rate was positively correlated across all temperatures at both the within- and among-individual level. This suggests that while individuals differ in their plastic responses, their rank order in metabolic rate is maintained across different thermal environments. This result is contrary to the idea of individuals can trade-off between better functioning at one temperature at a cost of function at another temperature as seen in killifish (*Fundulus heteroclitus*) where hot and cold temperature specialists for swimming endurance exists within the same population ({Powers:1998fv}{AngillettaJr:2003cp}). Moreover, consistent individual differences in metabolic rate independent of the environment can drive different ‘paces-of-life’ that are hypothesised to lead to consistent differences in suites of traits (Biro & Stamps, 2010; Careau et al 2008). For example, trade-offs between energy availability, reproduction and longevity can favour ‘proactive’ individuals with a high metabolic rate, active and bold personalities, that reproduce earlier at the cost of a shorter lifespan ({Reale:2010ef}).

Assuming phenotypic correlations are congruent with underlying genetic correlations ({Roff:2017gu, Roff:1995kt}), correlations between reaction norm attributes can have important evolutionary implications in understanding constraints on the evolution of metabolic plasticity. The strength of cross-temperature correlations can dictate how strongly selection on one component of the reaction norm (e.g. the intercept) will result in indirect selection on another (e.g. the slope) ({Via:1995hm}). We found that cross-temperature correlations between neighbouring temperatures (e.g., 28 C vs. 32C) were strong and the strength of this correlation is weakened at more distinct temperatures (e.g., 22C vs. 32C) when modelling with the character-state approach. While the correlations from the function-valued approach are in agreement with the character-state approach, the magnitude of correlations across temperatures remained strong. This is due to the important limitation of function-valued approaches whereby phenotypic values are strongly dependent on the covariance between the intercept and slope. In contrast, when modelling under the assumptions of the character-state approach, the shape of thermal reaction norms can evolve with weaker constraints and greater malleability. Although both approaches are equivalent when modelling phenotypic change in two environments {Hunt:2014wo, Via:1995hm}, one advantage of function-valued approaches is the ability to feasibly describe more complex reaction norm shapes by fitting higher order polynomials. On the other hand, the character-state approach requires a lot more data points to estimate means and covariances for each environment to uncover non-linear patterns. Differences between approaches may be ameliorated when curvature is properly accounted for in non-linear reaction norms, however we were unable to test this because our measurement temperatures spanned the normal operative temperature of the species where the reaction norm is likely to be linear ({Doody:2009dz}).

## Hierarchical differences in metabolic scaling at different temperatures

Our results are consistent with the growing number of interspecific studies that show temperature effects on mass-scaling exponents ({Killen:2010cw, Barneche:2016ke, }{Glazier:2005ei, Glazier:2015fr, Price:2012eg}). Generally, these studies demonstrate the mass scaling exponents increased with temperature and vary among species of different ecology (e.g. such as benthic or pelagic). To the best of our knowledge, we are one of the few studies that show temperature dependence of mass-scaling at both the among- and within- individual level ({Norin:2018ba}). This supports the idea that there is no universal metabolic allometry at lower levels of biological organisation ({Glazier:2005ei}, {White:2006fw}, {Norin:2018ba}). We hypothesise that thermal acclimation may explain why both the among- individual mass exponent was lowest at 24ºC. Individuals of different masses can remodel their physiology in order to optimise their energy expenditure to common housing conditions (~25ºC) ({Chevin:2010cw}{Seebacher:2014gf}). These compensatory adjustments involve changes in membrane composition or enzyme efficiency, which can alter how individuals of different masses to respire at different rates across at different temperatures, thereby impacting mass scaling exponents (Reviewed in {Seebacher:2010cb}, {Glanville:2006eo}).

Body mass is a key driver of metabolic processes and varies within the life time of an individual. We found that when within-individual variation in mass are not accounted for, the population mass-scaling exponents were always underestimated. This has important implications for predictive models that make use of scaling relationships to extrapolate individual level processes to ecosystems. Without correctly taking into account of the hierarchical structure of metabolic data, inferences at the population level can be misinterpreted {vandePol:2009em}. Notably, our within-individual estimates were at least three times greater than the (1.63 – 2.67), which is substantially higher than the within-individual scaling exponent of 0.79 reported in green iguanas ({Maxwell:2003hh}). It is important to note that variation in body mass in our sample population is small, which could give rise to large values of within-individual mass-scaling exponents. Nonetheless, it is still intriguing to point out that our estimates were in line of studies of endotherms such as bats and birds ({McLean:2007tl}, {Kvist:2001wt}). These results may possibly indicate that even the slightest changes body composition within an individual’s life, can impact body mass and energy expenditure {Scott:1996en}. In support of this, changes in fat mass strongly predicted within-individual variation in basal metabolic rate in Redshanks, a species of migratory bird ({Scott:1996en}). Food limitation and metabolising different energy stores could help explain these within individual effects given animals were intermittently fasted prior to measurements and were measured over a long period of time (4 months) ({McCue:2010dg}). Animals are known to adjust their physiological systems by shifting from carbohydrate-based energy reserves to more lipid- or protein-based reserves during periods of intermittent fasting (McCue, 2010). Utilising different types of energy sources may require different amounts of ATP which can in turn could impact metabolic rate. Catabolism of different energy fuels may help explain the diversity of intra-individual scaling exponents observed in vertebrate empirical studies (Reviewed in Glazier 2005). For example, an impressive intra-individual scaling exponent of 1.82 was observed in long-distance migratory waders. These birds have specifically evolved to mobilise, transport and utilise a range of energy reserves in order to travel long distances with limited opportunities to feed {Kvist:2001wt}. Future work is needed to investigate the physiological mechanisms and interactions of food limitation, fuel supply on scaling of metabolic rate.

# Conclusion

Our study emphasises the importance of considering individual variation in thermal plasticity of metabolic rate. Moreover, we demonstrate that the hierarchical structure of individual data (among- and within- individual effects) can influence population level estimates of mass-scaling exponents. By using two conceptual frameworks on how we model phenotypically plastic traits, we show that metabolic reaction norms in response to temperature may not be strictly linear and may have the capacity to evolve more malleable forms. While we do not advocate the use of any single approach in modelling plasticity, our goal is to illustrate how differences in assumptions of the function-valued and character-state approaches can influence the inferences we draw from them on how reaction norms may evolve. We provide our dataset and code to show how all these techniques can be implemented, with the hope to encourage researchers to use both approaches in order to gain a holistic view of their reaction norm data.

# Acknowledgements

Martin at the Lizard Lab, Christine, Interns, Stephan, Martin Stevens at Math department, Tobias Uller discussions

# References