# Title: Individual variation in thermal plasticity and its impact on metabolic scaling

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

# Abstract

# Introduction

The environment is comprised of numerous factors that can simultaneously affect multiple traits. Phenotypic plasticity allows individuals to modify their phenotype when the environment changes. Phenotypic changes in response to the environment can be conceptualised as a reaction norm, where a trait is measured as a function of an environmental variable (e.g. activity rate at increasing levels of predator density (REF?)). Many aspects of a phenotype, such as life history ({Westneat:2009dz}), locomotor performance ({Careau:2014in}), behaviour (reviewed in{Dingemanse:2010bk}) and physiology ({Boratynski:2017jf}) are responsive to environmental conditions. Metabolic rate, undoubtedly a very labile trait, determines an energy turnover of an individual and is repeatedly expressed within a lifetime of any organism. Metabolic rate is sensitive to fluctuations in temperature and body condition and is likely to be target for selection due to its strong links to behaviour, growth and reproductive strategies (Reale et al, Stamps TREE).

Myriad of studies have shown that metabolic rate differs consistently among individuals of a population (White et al, Nespolo) but whether individuals consistently differ in the thermal plasticity of their metabolic rate has received less attention.

While there is interest in understanding among individual variation in metabolism and their response to the environment, metabolic theories attempt to explain patterns at broad taxonomic scales using fundamental relationships between body size, temperature and metabolic rate (REF?). This body of work uses mechanistic processes such as thermodynamics of biochemical reactions to estimate a fixed scaling exponent that describes how metabolic rate proportionally changes with body mass and temperature (3/4 or 2/3 power laws, Brown, and Gilloly). However, there is accumulating evidence that challenges this ‘one-size-fits-all’ line of thinking (Glazier). Numerous studies have found that scaling exponents vary with temperature and among closely related species which suggests animals may adaptively adjust how their metabolic rate scales with body size in different thermal environments (Clark et al). Ecological differences between species such as foraging mode, endo-/ectothermy and genome size can, at least in part, explain some of the variation in scaling exponents at higher taxonomic levels, however much of variability remains elusive (reviewed in Glazier 2005, Uyeda et al 2017). While interspecific variation in scaling exponents may have true biological meaning, it may also be due to an ‘ecological fallacy’ where the scaling relationship 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 inflate inter-individual estimates, which in turn can inflate intraspecific and interspecific estimates. It is therefore important to correctly account for the hierarchical nature of biological data if the goal is to use these estimates for predicting ecological processes. Examining how metabolic rate is influenced by body mass and temperature from the bottom-up may prove useful for understanding the processes that shape metabolic reaction norms.

Consistent individual variation in reaction norms provide useful insight about the evolutionary potential of plasticity because individual differences reaction norm slope and intercept are the raw material that selection acts on ({Nakagawa:2010hv}). While the determinants of individual differences in plastic responses are not fully understood (reviewed in {Dingemanse:2013kf}), empirical studies show genetic variation, maternal effects {Anonymous:QHEHH5Cn}), early-life developmental conditions ({Beaman:2016cn}), can all contribute to individual variation. Consistent individual variation is commonly represented as repeatability (intra-class coefficient, ( Nakagawa, Wilson). Repeatability is often interpreted as the ‘upper bound of heritability’ because phenotypic differences between individuals is partly due to genetic differences, ({Wilson:2018iy, MR:2002bn}but see Dohm,). The effectiveness of selection acting on repeatable individual differences also depends on variation at the within-individual, i.e. how predictable is an individual relative to its own mean (Fig. 1), therefore it is crucial to distinguish variability attributed to among individual and within individual effects to understand the capacity for selection to act on consistent individual variation in reaction norms. However, how we model phenotypically plastic traits, influences how we quantify variance components of reaction norms and understand the evolution of reaction norms.

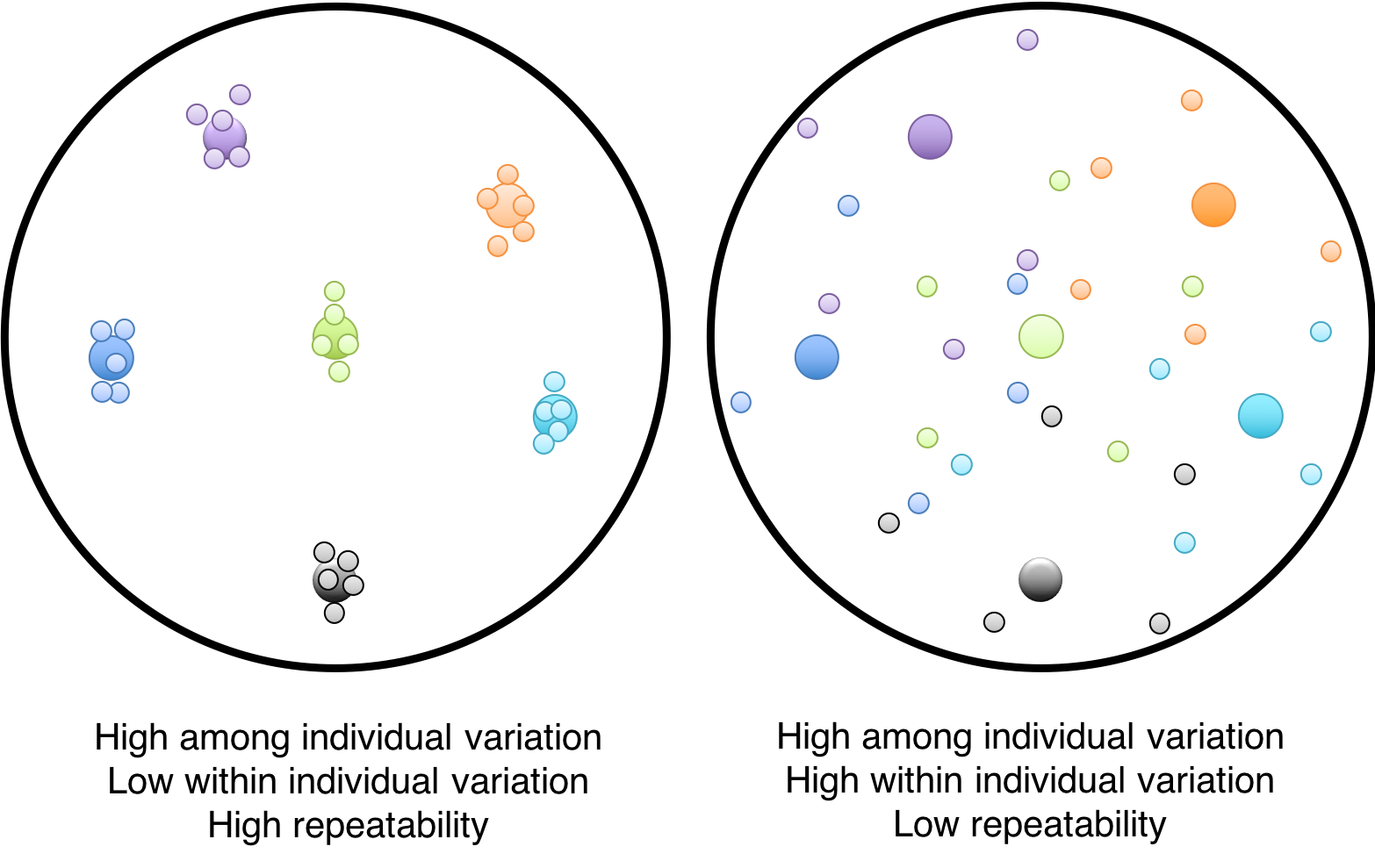


Figure 1: Graphical representation of among- and within-individual variance components of a trait. Each colour represents a unique individual. The large point represents each individual’s mean, smaller points represent within-individual deviations from the individual mean. Left: high, among-individual variation and low within-individual variation gives rise to high repeatability. Assuming among-individual variation is determined by genetic variation, the capacity of selection is high in this scenario. In contrast, when there is high among individual variation and high within individual variation, repeatability is low and selection at a given time is acting on within-individual noise.

Two approaches 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 line of thinking, evolutionary processes are environment-specific and act on individual variation in phenotypic values in each environment and the covariance of phenotypic values between different environments. 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. Under this approach, selection pressure is assumed to be equal across all environments and operates on individual variation in model parameters as well as their covariances e.g. intercept and slope term. Given that the same phenotypic trait measured in multiple environments are inherently correlated, evolutionary constraints can occur on reaction norm components. Non-zero covariances dictate how strongly character states in other environments is indirectly selected upon, it also means that the elevation and angle of reaction norm must evolve in tandem. While the conceptual differences between the modelling approaches have sparked debates on which approach is more appropriate under what circumstances, both approaches imply that an optimal reaction norm shape may not ever be reached and provide valuable insight in the evolution of plasticity.

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?

While plasticity is pervasive in nature, the evolutionary processes that shape reaction norms are not well understood.

# 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 while 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 all our models we used Bayesian linear mixed models (LMM) from either the ‘brms’ or ‘MCMCglmm' R packages (REFS)

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 can potentially influence 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 model details).

### Repeatability of thermal reaction norms

We used function-valued and character state approaches to quantify the repeatability of metabolic thermal reaction norms to determine they differed consistently among individuals. 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 approaches to estimate the repeatability of the intercept of individual reaction norms. This class of models assumes that the intercept is 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 function-valued 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 + (ID+ logTempcen) + (Series+ 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. Character state models variation of logVCO2 at each temperature which requires a substantial amount of data points for proper estimation ({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 level where we fitted logVCO2 as a response and same as previous models, logTemp, zlogBodyMass and logPriorTemp were included as predictors. Individual IDs and logTemp were included as a random intercept and slope, respectively. Sampling session ID was included as a covariate instead of using ‘series’ as a random effect because to the best of our knowledge no method has been proposed on deriving correlations from more than one random effect. 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 vector 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 over a temperature gradient

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. Exploratory plots of individual mass across the entire course of the experiment (nobs = XXX) shows that XXX. they were fasted for at least 24 hours prior to the first day of measurements.

We first 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 from the individuals mean across sampling sessions (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

## Multilevel mass scaling exponents and temperature dependence



**Figure 2** – A) Posterior mean estimates of three types of mass-scaling exponents across six measurement temperatures. Within individual scaling exponents (Black inverted triangle ▼) describes the change in log VCO2 as an individual’s mass changes on the logarithmic scale. Between individual scaling exponents (Black up-pointed 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 (Year). 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.

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

## Repeatability of thermal reaction norms

**Figure 3** – 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.

Overall, repeatability in log VCO2 increased with temperature. 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. 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 I x E interaction that was consistent over time (Fig. 4).

**Figure 4** – 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 between individual level (Fig. 5). Certain individuals maintained a high metabolic rate, while others had a relatively low metabolic rate across all temperatures (Fig. S1) creating a positive relationship between metabolic rate at the between 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. In short, the correlations estimated using FV approaches are congruent with the CS approach however, estimates differ in magnitude and credible intervals were inaccurate.

**Figure 5** – 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.

# Discussion

Our study highlights the importance of considering how individual variation in metabolic rate respond to temperature. We demonstrate that, by not accounting for the hierarchical structure of population-level data (i.e., disentangling within and between-individual variance), mass- scaling exponents can be overestimated . Additionally, we show that within-individual mass-scaling exponents can be significantly higher than among-individual exponents and that these exponents can change with measurement temperature. We found that repeatable differences between individuals in metabolic rate increased as temperatures became hotter and this was due to a reduction in within-individual variation and an increase in among-individual variation. Metabolic rate was also positively correlated across all temperatures, however the strength of these cross-temperature correlations were not uniform across all temperatures. Correlations were strongest at higher temperatures and between neighbouring temperatures and these weakened with increasing differences between measurement temperatures.

## Multilevel metabolic scaling exponents change with temperature

Our results are consistent with the growing number of studies that show temperature effects on mass-scaling exponents and this occurred at both the among- and within- individual level. Thermal acclimation could explain why the among-individual exponent in our study was lowest at 24ºC. Individuals of different masses can remodel their physiology to common housing conditions (~25ºC) in order to optimise their energy expenditure. This may allow lighter individuals to respire at a disproportionally higher rate to enhance growth, while heavier individuals can maintain their larger size while conserving energy. In contrast, the within individual exponent was highest at 24ºC. This suggests that when individuals are acclimated to a specific temperature, their metabolic rate responds more drastically to the slight changes in their mass compared to other temperatures. Although it may seem counter-intuitive, biochemical rates within an individual may operate more efficiently at an acclimated temperature as the physiological system is more primed to operate in response to changes in body mass. Mechanistic studies will be necessary to explain the interactive effect of temperature and body mass on physiological rates, particularly at the within individual level.

Body mass is a key driver of metabolic processes and varies within the life time of an individual. We found that within-individual scaling exponents were significantly higher than among-individual exponents. Most notably, our within-individual estimates were three times greater than the ¾ power law (1.63 - 2.67). Food limitation and metabolising different energy stores could help explain these within individual effects (REFS). Our study animals were fasted for at least four days during each sampling session (24 hours prior to three days of measurements). During periods of intermittent fasting, individuals’ mass and ultimately the types of resources being metabolised (e.g., fat or protein) will inevitably fluctuate which may affect their metabolic rate (REFS). Animals are known to adjust their physiological systems by shifting from carbohydrate-based energy reserves to more lipid- or protein-based reserves which can influence metabolic rate (McCue, 2010). Catabolism of different energy fuels may help explain the diversity of intra-individual scaling exponents observed in vertebrate empirical studies (Reviewed in Glazier 2005). In support of this, 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. Utilising different energy sources require different amounts of ATP which can in turn impact metabolic rate (refs). However, future work is needed to investigate the physiological mechanisms and interactions of food limitation, fuel supply on energy expenditure.

## Consistent variation in thermal reaction norms

### Temperature-specific repeatability and variance components

Numerous studies have shown that metabolic rate varies consistently between individuals (reviewed in Nespolo and White reviews), but does this individual variation change as a function of the environment? We show that repeatability increased with temperature, which suggests that individual reaction norms are not parallel (Nussey 2007, Neils paper). Temperature-specific repeatability can have important evolutionary consequences on the evolution of responses to changing temperature, assuming there is a genetic basis in variation of metabolic rate. More consistent individual variation in certain environments provides greater opportunity for selection (Janicke et al 2016). However, the effectiveness of selection also depends on the relative contributions of both among- and within-individual variation (ref). One striking result we found was that among individual variation increased with temperature, while within individual variation decreased. This indicates that differences among individuals were more pronounced as temperatures become hotter, but individuals’ metabolic rate were more predictable relative to their own responses (Cleasby and Nakagawa, 2014). Taken together, our results suggest that natural selection in a specific temperature environment may not be effective because contributions of difference variance components are not uniform across different thermal environments. This could potentially give rise to non-linear reaction norm shapes at the population level, simulation studies would be very insightful to test this hypothesis.

### Repeatability of thermal plasticity

Repeatable differences in thermal reaction norms are predicted to capture the upper limit of the evolutionary potential of plasticity (refs?). From an evolutionary perspective, repeatability sets the upper limit of heritability because consistent among-individual differences incorporate additive genetic variance as well as irreversible environmental effects (Falcon and Mackay 1996, see Dohm 2002 for exceptions). While previous studies have shown metabolic rate is a repeatable and heritable, the capacity for an individual to respond to temperature has received considerably less attention. We found consistent individual differences in plasticity of metabolic rate in response to temperature. Certain individuals maintained high metabolic rates, while other individuals had consistently low metabolic rate across all temperatures. Individual variation in thermal plasticity has also been previously reported for performance traits in clawed frogs (Careau et al 2014a) and for metabolic rate in salamanders (Careau et al 2014b) and zebra finches (Briga and Verhulst, 2018). Consistent among-individual variation is the raw material natural selection acts on and is a key prerequisite for plasticity to evolve. Both genetic and permanent environmental effects can influence the development and maintenance of physiological structures such as enzymes and cell membranes required for responding to changes in temperature, thus maintaining individual variation in thermal plasticity. Our results suggest that the thermal reaction norm for metabolic rate has the potential to undergo natural selection if variation among-individuals has a heritable component. This will require verification by estimating the heritability of the thermal reaction norm of metabolic rate, which can be logistically challenging but crucial to understand the adaptive evolution of plasticity.

## Cross-temperature correlations of metabolic rate

Metabolic rate measured at different temperatures was positively correlated at the within- and among-individual level. Irrespective of temperature, some individuals have consistently higher energy turnover, while others do not. Consistent individual differences in metabolic rate can promote temporally consistency in behavioural traits (Biro & Stamps, 2010; Careau et al 2008). Our results show that while individuals differ in their plasticity (I x G), their rank order in metabolic rate are maintained across different thermal environments. Consistency in rank order also implies that there are not trade-offs in thermal reaction norms at the within- or among individual level. While it is intuitive to consider allocation trade-offs in thermal reaction norms, for example, an individual may allocate resources to form more thermally stable enzymes at one temperature compared to another, our finding contradicts this hypothesis (Uberfleas and Angilletta paper). Trade-offs may not manifest under acute changes in temperatures because immediate responses are strongly governed by the thermodynamics of biochemical reactions rather than phenotypic adaptation. Instead, trade-offs in whole-organism metabolism to different thermal environments may occur with the capacity for individuals to acclimate. Thermal acclimation requires allocating resources to remodel different aspects of the physiological system. This finite pool of resources could be determined by genetic differences or even permanent environment differences at development and may determine trade-offs in plastic acclimation responses (Beaman et al, TREE paper). Detecting whole-organism performance trade-offs following acclimation at the individual level would be fruitful and insightful avenue to pursue in order to understand the evolution of thermal reaction norms shape.

The strength of cross-temperature correlations determines how strongly responses at other temperatures are indirectly selected on. To our surprise, the strength of correlation of metabolic rate was not homogeneous across the temperature gradient. This can give rise to non-linear forms of thermal reaction norms if there is a heritable component underlying these phenotypic correlations. For example, if selection were to operate on metabolic responses at 30ºC, correlated selection on responses at neighbouring temperatures such as 28ºC and 32ºC would be selected upon more strongly than compared to metabolic rate at 22ºC. While the mechanisms that determine cross-temperature correlations are unclear, measurement error at low temperatures may have added more noise to the data, resulting in weaker correlations. Strong correlations imply that metabolic rate at different temperatures may be under ‘modular’ control. Mechanisms such as heat shock proteins that allow enzymes to be more structurally stable at hot temperatures may be recruited when an animal experiences rapid increases in temperature which could result in a strong positive correlation between hot temperatures (Somero, 1995 and Fields 2001). Thermal tolerance of difference metabolic enzymes and proteins may be a plausible explanation for cross-temperature correlations, however characterising the enzymes or expression of heat shock proteins at different temperatures may be logistically challenging but a direct way of testing this hypothesis.

## Considerations of using both approaches

* Amount of data points vs number of environments
* Assumptions of linearity
* Function-valued models are useful in providing insight in whether individuals consistently vary in plasticity, the character-state approach does not directly address this. Without assuming that phenotypic variance is contingent to an intercept-slope covariance, the character-state approach showed that individual reaction norms are not strictly linear (Fig. 3b). Acclimation could result in non-linear reaction norms. In support of this, the character-state approach showed very little change in repeatability between 24ºC – 26ºC. which suggests that individuals are converging to similar physiological responses to compensate under lab conditions
* Cross correlation estimates poorly estimated using matrix algebra with FV models

# Conclusion

[Summarise key findings]

We emphasise the importance of considering the hierarchical structure of datasets when trying to estimate mass-scaling exponents at multiple scales of biological variation. Moreover, while we do not advocate the use of any single approach in modelling plasticity, our primary aim is to illustrate how differences in assumptions of the function-valued and character-state approaches can influence the inferences we draw from them.

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

# References

Notes:

* Could this be due to fasting and metabolism different types of energy reserves e.g. Carbs, proteins, and fats
* Catabolism processes with some of these energy reserve types and this ramps up MR
* Going through periods of intermittent fasting, physiological system may want to reduce MR over time and this can drive WI effects
* An individual can modify 3 times of MR over such a small scale so huge compared to taxonomic scale
* Huge range in empirical studies
* While a wide variety of non-linear reaction norms exist in nature, we will focus our discussion on linear reaction norms for simplicity sake.
* The biological interpretation of evolutionary constraints of the approaches is contingent on our limited understanding of the mechanistic basis of these covariances and how selection operates on labile traits, regardless
* Long history of trying to explain variation in metabolic using a unifying math model – but this is just not enough, way too simplistic, nonetheless a good place to start
* MR determines energy budget pools, animals faced with trade offs, behaviour in a certain way, make decisions about life history in a certain way to optimise this balance of energy.
* MR is undoubtedly one of the most labile traits
* erroneously tease out the hierarchical effects and misinterpret
* Current climate change projections predict harsher and greater fluctuations in temperature regimes in the coming years putting natural populations, particularly of ectotherms at risk.
* An individual’s body composition can fluctuate throughout its lifetime, particularly during periods of food limitation, which will undoubtedly influence an individual’s metabolic rate.
* Individuals can adaptively switch from carbohydrate-based fuels to lipid-based as it determines. Carbohydrate, lipids and protein are the main types of metabolic fuel and requires different amount of energy to breakdown.
* Individuals of the same weight can also differ in their body composition catabolising different energy reserves can result in different metabolic rates ().
* An individual’s body mass can fluctuate drastically within its lifetime and may reflect temporal changes in body composition.
* these examples demonstrate the need to explore the how among- and within-individual variation in mass affects metabolic rate across different temperature environments.
* An individual’s body mass can fluctuate drastically within its lifetime, these can be in response to temporal changes in growth, diet, seasonality and reproductive activity. These fluctuations in body mass will undoubtedly affect an individuals metabolic rate and individuals may adaptively adjust their energetic expenditure to conserve energy. This intra-individual variation in mass and its effects on metabolic has been largely been neglected in the metabolic theory literature. Interestingly, intra-individual exponents are often greatly than one, implying when individuals increase in mass, their metabolic rate increases disproportionally higher than a lower mass. Neglecting to account for hierarchical variation in scaling relationships will confound within-individual effects and among-individual effects. Distinguishing between inter- and intra- individual effects allows to new hypotheses to be formulated about the mechanisms that drive broad scale patterns from the bottom up ({vandePol:2009em}).
* Whole-organism metabolic rate, undoubtedly a very labile trait, determines how an animal optimises their energy expenditure to competing processes such as reproductive, growth or maintenance (Careau:2008fi}).
* Across broad taxonomic groups, metabolic rate scales with mass following a ¾ power relationship which suggests that mass effects on metabolic rate is dictated my common mechanisms (reviewed {Glazier:2005ei, Glazier:2015fr}).Ttemperature influences metabolic rate through its effects on the rate of biochemical reactions for and varies according to the Boltzmann factor (*e-E/kT*) for any given body mass where MR = metabolic rate, M = body mass, E = the activation energy of metabolism, T = absolute temperature, k = Boltzmann’s constant and B0 is a normalisation constant independent of M and T. This equation explicitly assumes that metabolic rate scales with The equation also assumes that temperature influences metabolic rate through its effects on the rate of biochemical reactions for and varies according to the Boltzmann factor (*e-E/kT*) for any given body mass. This generalised equation also makes the implicit assumption that relationship between temperature and metabolic rate is identical across different hierarchical and taxonomic levels because the kinetics of a reaction is underpinned by same thermodynamic mechanism (See Clarke 2004 for in depth discussion). Accumulating evidence from intraspecific studies has shown that mass-scaling relationships do not adhere to the ¾ power law and often interact with temperature ({Barneche:2016ke}). This may be because reaction
* body mass and metabolic rate can vary drastically among individuals of the same population and within an individual, moreover multi-level variation in mass-scaling relationships was largely neglected during the development of unifying theory of metabolic ecology
* Body size is fundamental in governing an individual’s metabolic rate {Brown:2004hp, Gillooly:2001cg}. Metabolic theories assume that metabolic rate relates to body mass following a power relationship, irrespective of organism size. However, these theories neglect to consider that individuals of the same body mass can have very different energy expenditure and an individual’s body mass can fluctuate drastically within its lifetime. Individuals of the same weight can vary in their internal anatomy or body composition which requires different energy expenditure to maintain physiological processes. For example, in a study of six inbred lines of lab mice, lines that have relatively higher basal metabolic rate after correcting for body mass differences, were characterised by larger small intestine, heart, liver and kidney {Konarzewski:1995cu}. An individual’s body composition can determine the type of energetic fuel utilised by the metabolic system as well as the lipid content of cell membranes which in turn can affect metabolic rate. These mechanisms can maintain variation in mass among individuals and within individual and alter the population scaling relationship with metabolic rate. In order to properly understand how metabolic rate scales with mass at higher levels of biological variation, variability in mass needs to be accounted for at the among- and within- individual level.
* Temperature is also an important factor in determining metabolic rate. The interactive effects of body mass and temperature on metabolic rate is less well understood.
* There is also a growing number of studies that show mass-scaling changes with temperature