THESIS

INTER-INDIVIDUAL VARIATION WITHIN SOCIAL GROUPS: HOW METABOLIC RATE SHAPES THE PACE OF LIFE

Submitted by

Stephen G. Mugel

Department of Biology

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2020

Master's Committee:

Advisor: Dhruba Naug

Gregory L. Florant
Paul Ode

ABSTRACT

INTER-INDIVIDUAL VARIATION WITHIN SOCIAL GROUPS: HOW METABOLIC RATE SHAPES THE PACE OF LIFE

Metabolic rate (MR) is often cited as the fundamental rate which determines the rate of all biological processes by shaping energetic availability for the various physiological, behavioral, and life-history traits that contribute to performance. Furthermore, the metabolic theory of ecology posits that performance at any level of biological organization is a function of the MR of its constituent units. It has therefore been suggested that MR drives the widely observed covariance among these different levels of phenotypic traits. However, much of the work on this topic has relied on pairwise correlational analysis on a handful of traits at a time, leaving an important gap in our understanding regarding the functional links that shape this phenotypic covariance, often referred to as pace-of-life. Furthermore, at a collective level, this has led to significant attention regarding how MR scales across group size, but considerably less attention has been paid to how heterogeneity in MR among constituent units shapes collective outputs.

Using honeybees as a model, we measured a large number of behavioral, physiological, and life-history traits in individual bees and used a path analysis to demonstrate that variation in metabolic rate plays a fundamental proximate role in driving the covariance among these traits. We combined this with a factor analysis in a structural equation model framework to characterize the overall phenotypic covariance or the pace-of-life axis in honeybees. We discuss the importance of these findings in the context of how interindividual variation in terms of slow–fast phenotypes

may drive the phenotype of a group and the functional role metabolic rate might play in shaping division of labor and social evolution.

Building on this work, we leveraged the well-characterized differences in MR associated with 'F' ('fast') and 'S' ('slow') malate dehydrogenase (MDH) alleles to breed homozygous genotypes of bees expressing high (FF) and low (SS) MR in addition to heterozygotes (SF), thought to express an intermediate phenotype. We then mixed progeny from these lines to create experimental groups with four different phenotype compositions: monomorphic FF, monomorphic SS, monomorphic SF, and a polymorphic group type at a 1 FF: 1 SS ratio. We then measured MR, energetic intake, thermoregulation in cold and heat stress, and survival of these groups in a high and low resource environment. Monomorphic fast groups outperformed monomorphic slow and polymorphic groups, which performed worse than expected on most traits. We quantified the effect of heterogeneity on polymorphic group performance using the 'diversity effect,' an analytical technique often used in ecosystem ecology to compare the productivity of diverse ecosystem assemblages to null expectations set by the constituent species when living alone. Diversity effects can be partitioned selection and complementarity effects and understand the mechanisms through which biodiversity acts on ecosystem productivity. We applied this technique in a novel way to show how each group-level performance trait is influenced by MR *morph* diversity through different processes. We also found that MR was strongly correlated to the other traits, especially in the low resource environment. We discuss these results in the context of how MR plays an important role in shaping division of labor and social evolution.

These studies provide empirical support for the theoretical idea that metabolic rate acts as a proximate driver of phenotypic covariance among a number of physiological, behavioral, and life-history traits at the individual level, and that behavior acts as a mediator for how metabolic

rate affects life history. In addition, using honeybees as an experimental model for these studies establishes a framework for asking questions regarding how these individual-level phenotypic covariance patterns lead to observed phenotypic covariance patterns at the colony level that have functional consequences for division of labor and social evolution. The results of these studies therefore contribute toward a better understanding of the rules of life that shape processes across different levels of biological organization. Our use of different structural equation modeling approaches for inferring heuristics and proximate causal relationships among multiple phenotypic traits also informs future research efforts on this topic. We also present a novel approach to experiments that explore functional group level performance traits through partitioning the effects of inter-individual heterogeneity.

ACKOWLEDGEMENTS

I would like to thank Dhruba Naug, Catherine Tait, Julian Cassano, Abbie Reade, Maybellene Gamboa, and our huge team of undergrads Brielle Hermstedt, Michael Matthews, Abby Hernandez, Natalie Namba, Amy Zimmermann, and John Starineiri for all the help to collect the data presented here. I would like to thank Greg Florant and Paul Ode for their helpful questions and insightful comments in improving this work. I would also like to thank my lovely fiancée Diana Jaramillo-Nuñez, for her unwavering support and understanding of long nights and days conducting lab work, and my parents and sisters Richard and Tammy, Margaret and Emily Mugel for their support and care throughout this experience. Finally, I would like to thank the other wonderful graduate students, faculty, and staff of the Biology Department who have made my time at Colorado State University a transformative experience.

TABLE OF CONTENTS

CHAPTER 1: METABOLIC RATE SHAPES PHENOTYPIC COVARIANCE AMONG PHYSIOLOGICAL, BEHAVIORAL, AND LIFE-HISTORY TRAITS IN HONEYBEES

Introduction

Metabolic rate (MR) is often considered to be the fundamental biological variable which determines the rate at which organisms acquire, process, and expend energy (Brown et al. 2004). However, the functional significance of both interspecific and intraspecific differences in MR is far from clear and has long been a topic of wide interest. In terms of within-species variation, two hypotheses predict how MR, by influencing energy acquisition and energy allocation, may in turn influence behavior and life-history (Biro and Stamps 2010). A higher MR, by acting as a metabolic engine, can allow for greater energy acquisition, making possible higher levels of behavioral performance and growth rates. However, a higher MR may also require higher energy allocation toward its maintenance and thereby exert a negative effect on these same performance variables (Careau et al. 2008). This dichotomy shows how MR can have a central but complex role in driving the widely observed covariance among physiological, behavioral, and life-history traits (Biro and Stamps 2008; Glazier 2015; Krams et al. 2017), often referred to as the pace-of-life syndrome (POLS) model (Ricklefs and Wikelski 2002; Réale et al. 2010). This model integrates the entire covariance pattern into a single slow – fast phenotypic axis, wherein 'slow' individuals live longer but more cautious lives, are shyer, more social, and have lower MR, whereas 'fast' individuals adopt a life-fast, die-young lifestyle, engaging in riskier, bolder behaviors, and a higher MR.

The putative central role of MR in driving such covariance and shaping the POLS axis has, however, rarely been empirically investigated in an integrative fashion using large suites of traits

(Careau and Garland 2012; Glazier 2015). In fact, the focus has remained largely on ultimate causation – framed around a trade-off between current and future reproduction, or between survival and current reproduction (Wolf et al. 2007). Functionally, such life history trade-offs have been suggested to be mediated through risk-taking behaviors, defined as behaviors that increase resource acquisition but with costs such as increased mortality, disease exposure, energy expenditure, or lost mating opportunities (Dammhahn et al. 2018; Mathot and Frankenhuis 2018), a prediction that is supported by some empirical studies (Sol et al. 2018; Jacques-Hamilton et al. 2017; Careau et al. 2009). However, most of these studies are based primarily on pairwise correlational analysis on a handful of traits, leaving an important gap in our understanding regarding how these traits are functionally integrated through an underlying set of causal relationships, in particular the putative proximate role of MR in shaping the slow–fast axis.

More recently, structural equation model approaches have been increasingly used to simultaneously assesses the strength and directionality of the relationships among the large number of traits that constitute the POLS model (Goulet et al. 2017; Krams et al. 2017; Santostefano et al. 2017; Jablonsky et al. 2018; Debecker and Stoks 2019). However, the limited data regarding the covariance between different phenotypic traits and MR have been fairly mixed – MR has been shown to be both positively and negatively associated with important life-history traits in different species (see Arnqvist et al. 2017; Royauté et al. 2018), indicating the need for further studies.

In this context, social insect colonies, which express considerable variation in MR, behavior, and life history among individual workers (Harrison and Fewell 2002; Jeanson and Weidenmüller 2014), present an opportunity for understanding how these different types of

phenotypic variation might be functionally linked and contribute to variation in terms of the position of an individual on the slow – fast phenotypic axis. The phenotypic diversity among individuals in social insect colonies is considered functionally critical to division of labor and its adaptive plasticity (Oster and Wilson 1978), features that are often credited for their tremendous ecological success. Understanding how MR covaries with slow – fast worker phenotypes in terms of behavior and life history in social insect colonies can therefore provide insights into the functional mechanisms that have contributed to social evolution. Social insects also present an intriguing case for POLS theory because little is known about how variation with respect to the various pace-of-life traits at the individual level translates to the observed variation in the same traits at the colony level (Segev et al. 2017; Blight et al. 2016; Bengston et al. 2017).

The honeybee (*Apis mellifera*) as an experimental model allows enormous opportunities to address how MR is related to the covariance among different levels of phenotypic traits. A number of traits such as sensorimotor responses, gustatory responsiveness, associative learning and the rate of behavioral development such as age polyethism, are known to be associated with aspects of foraging behavior such as preference for pollen or nectar (Page et al. 2006), a few of which also show a significant association with MR in other studies (Harrison and Fewell 2002; Feuerbacher et al. 2003). In this study, we therefore measured MR (routine and flight metabolic rate) and multiple other physiological (gustatory responsiveness), behavioral (energetic consumption, activity level, nursing and social behavior), and life-history (eclosure weight, age of first foraging and lifespan) traits that constitute the POLS model (see Réale et al. 2010) and are also relevant for honeybees, to characterize the covariance among them. Using a structural equation model (SEM) framework, we first compared a set of path models to test the proximate causal role of MR in

driving the observed phenotypic covariance among these traits and then used a factor analysis to characterize the nature of the overall covariance across all the traits or the POLS structure at the phenotypic level.

Methods

A common garden experimental design was created by periodically introducing eight cohorts of 100–200 newly emerged bees from four source colonies of *Apis mellifera* into a single, free-foraging, queen-right colony housed in a 3-frame observation hive. Combs with brood were extracted from the source colonies just prior to adult emergence and kept in an incubator at 32°C and 60% RH and newly hatched bees were individually weighed, marked with a unique number tag on their thorax before being introduced into the observation colony (Fig. S1.1A).

The within-nest behavioral profile was measured as the proportion of time an individual bee was observed in each of the following states: activity (walking), rest (sitting or autogrooming), sociality (exchanging food with another adult or allogrooming), and brood care (head inside a brood cell). These were calculated from scans conducted on tagged bees in the observation hive every 15 minutes for 3 hours each in the morning and afternoon when these bees were 4-5 days old, resulting in a total of 48 possible scans (4 per hour x 3 hours x 2 sessions x 2 days) for each bee. Since every tagged bee could not be located in every scan and the actual number of scan samples was different for each bee, a bootstrapping procedure was used to randomly select and average across 100 iterations of 10 random scan samples for each bee, and only for those bees that had more than 10 scan samples (totaling 780 bees). The hive entrance was observed every other day in two 3-hour periods during normal foraging times to record the age of first foraging (AFF)

for each bee which became a forager, but this variable was collapsed into a conservative weekly measure due to the limited sample size in each sampling interval.

Tagged bees were collected randomly as they departed the hive on foraging trips after they initiated foraging, and their flight metabolic rate (FMR) was measured using a FoxBox respirometry setup (Sable Systems). Each bee was placed in a clear 250 mL sealed glass chamber maintained at 22.4 ± 0.8 °C and ambient air scrubbed of H_2O and CO_2 was run through the chamber at a constant rate of 750 mL/min for 10 minutes (Fig. S1.1B). The CO_2 concentration in the excurrent airflow was recorded every second and corrected for drift by subtracting baseline CO_2 readings taken prior to each recording. Flight was stimulated by shining a light above the chamber and lightly agitating the chamber as necessary. The behavior of the bee was monitored constantly throughout the assay and FMR was calculated only during the 60 seconds of continuous flight with the most stable (lowest variance) CO_2 production (Fig S1.2). Bees that did not fly for 60 continuous seconds were discarded. Each bee was weighed immediately and mass-specific FMR was calculated as the weight corrected mean CO_2 production (mL hr^{-1} g^{-1}) which was transformed into a weight-corrected power output (mW g^{-1}) by multiplying it by 21.4 J mL^{-1} CO_2 and dividing by 3600 J hr^{-1} W^{-1} (Feuerbacher et al. 2003).

Following the FMR assay, each bee was harnessed in a plastic straw using a small wire, satiated with 30% sucrose, and maintained in an incubator at 25°C overnight. Each bee was then assayed for its gustatory responsiveness score (GRS) by touching its antennae with sucrose solution in an ascending series of concentrations (0.1%, 0.3%, 1.0%, 3.0%, 10%, 30%) and recording each instance of proboscis extension (Page et al. 1998). Each bee was then satiated with

30% sucrose and placed in the incubator for 2 hours to ensure a post-absorptive state in each bee. Following this, the routine (standard) metabolic rate (RMR) of each bee was measured by placing each bee in a dark 50 mL chamber through which ambient air scrubbed free of CO_2 and H_2O was passed at a constant rate of 250 mL/min. CO_2 production was recorded as described above and RMR was calculated for the continuous 2 minute period with the lowest average CO_2 production – a period which in all our samples included more than one cycle of CO_2 production, representing one respiration cycle (Fig. S1.3). Mass-specific RMR (mass independent data in Fig. S4) was then calculated as a weight-corrected power output as described above.

Following the RMR assay, each bee was individually housed in a small wooden cage (4 x 2 x 1.5 cm) fitted with a modified centrifuge tube and fed *ad libitum* a 30% sucrose solution and placed in an incubator at 25°C and 40% RH. Daily sucrose consumption was calculated for 5 days by measuring weight change of the feeder after correcting for any evaporative weight loss. This average daily consumption was transformed into a weight-corrected energetic equivalent representing energetic intake (1 mg sugar = 16.7 J). The survival of each bee was monitored daily until its death, and lifespan data were used only for those bees that survived for more than 2 days to discount any stress related death.

Data Analysis

All statistical analyses were performed in R (version 3.4.1, R core team). Structural equation models (SEM) were used to analyze the covariance and causal relationships among life-history, behavioral and physiological traits for the 148 bees on which at least eight of the ten traits were measured. The covariance structure depicting the correlation among all traits is calculated

using single, not repeated, measures for a given bee owing to the limitation of the number of traits measured in each individual within the relatively short honeybee lifespan. Therefore, since between- and within- individual variance components could not be parsed, the covariance observed here is at the level of the phenotype rather than being reflective of the underlying genetic structure. That said, any environmental components of the observed variation are expected to be minimal given our common garden design.

First, in order to test the hypothesis that metabolic rate drives differences in the observed phenotypic covariance structure, path analysis, a specific application technique of SEMs, was used to test multiple alternative putative causal pathways among the various traits. With ten traits that were measured, the number of possible path models is very large and so only a subset of them was tested. Path models were selected *a priori* as two general classes: models in which physiology influences life-history and behavior (Class A) and models in which behavior influences physiology and life history (Class B). In all models, life-history traits were considered as the final dependent variables or endpoints in the causal pathway. Within each class, we postulated three *a priori* model schemas with paths that integrated the traits in different ways: a fully-integrated Schema 1 that integrates all the behavior, physiology, and life-history variables, e.g. in class A, physiology influences behavior and *both* in turn influence life-history, a separate endpoints Schema 2 in which, e.g., physiology drives both behavior and life-history but behavior and life-history are themselves unrelated, and a mediator Schema 3 in which, e.g. physiology acts on life history through behavior acting as a mediator (see Figs. 1, S1.5–S1.11). We used hypothesis testing to sequentially drop non-significant relationships from *a priori* models to better fit the data. All models, including a null model in which all traits were independent of each other, were ranked

using AIC model selection. The top models of each schema are presented in the supplement (Fig. S1.5–S1.11) and the overall top model across all schema is presented in Fig. 1.1.

We then employed factor analysis, another specific application of SEM, to characterize the overall phenotypic covariance among observed traits, or a pace-of-life axis, absent the complexities of a causal path model. First, a hierarchical SEM was constructed that separately assessed the phenotypic covariance among all physiological traits (a physiological axis), all behavioral traits (a behavioral axis), all life-history traits (a life-history axis) as three separate latent factors, and then assessed the overarching covariance across these three axes as a further latent factor, which we interpret as representing the phenotypic pace-of-life axis (Fig. 1.3A). A second SEM was then constructed with a more parsimonious approach, in which the covariance among all physiological, behavioral and life-history traits was simultaneously assessed as a latent factor representing the phenotypic pace-of-life axis without first grouping the traits into separate categories (Fig. 1.3B). All the measured trait variables were appropriately transformed for normality and scaled to achieve more uniform residual variance measures with models being estimated using the full information maximum likelihood estimator.

All SEMs were built and tested using the R package 'lavaan' (version 0.6-2, Rosseel 2018). To test for the effect of source colony from which the experimental bees came, linear models of those relationships found to be significant in the top supported path model were re-run separately with colony as a covariate (Table S1.1).

Results

The top supported model in our path analysis (ΔAIC = 0, Fig. 1.1) shows how metabolic rate could shape behavior, and in turn affect life-history traits, with additional direct effects of RMR on life-history. RMR had a significant positive effect on GRS and FMR while negatively influencing the proportion of time spent in brood care and positively affecting activity levels. FMR positively affected consumption, which in turn negatively influenced lifespan. Activity had a negative influence on brood care, which in turn had a negative influence on lifespan. RMR negatively influenced AFF, indicating that bees with higher RMR began foraging earlier in life. Other models tested, including the null model, garnered weaker statistical support based on AIC model selection (Fig. S1.5–S1.11). Linear models on all significant relationships from the top path model with colony included as a covariate revealed similar strength and statistical significance to the path model approach, with colony having no significant effect on any response variable (Fig. 1.2, Table S1.1).

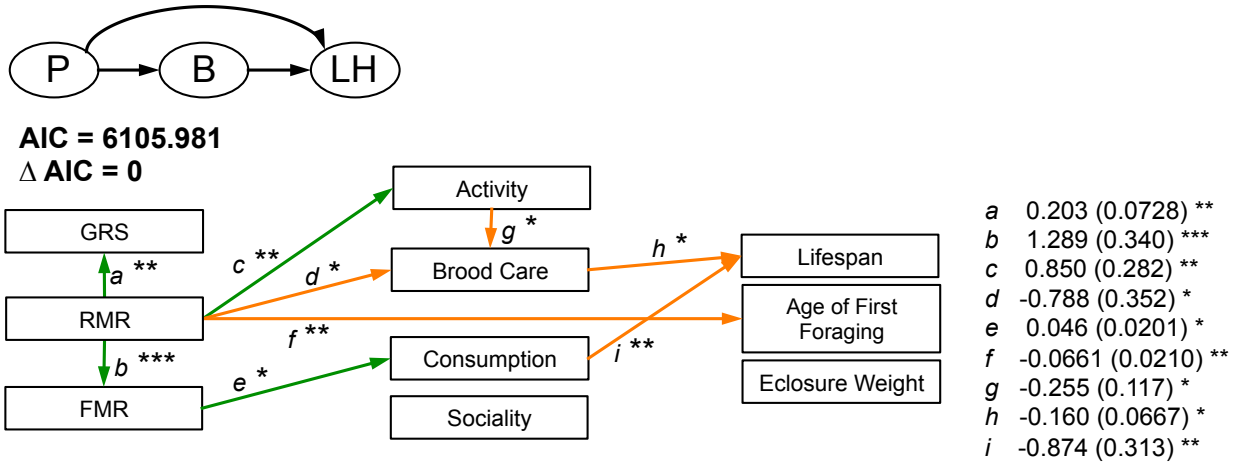

Figure 1.1. The top supported path analysis model tested for the role of metabolic rate on behavioral and life-history traits (ΔAIC = 0; additional models in Fig. S1.5–S1.11). Lines represent significant partial correlation coefficients (positive in green and negative in orange) with standard errors in parentheses, at $p \leq 0.001$ (***), $0.001 < p \leq 0.01$ (**), or $0.01 < p < 0.05$ (*) level.

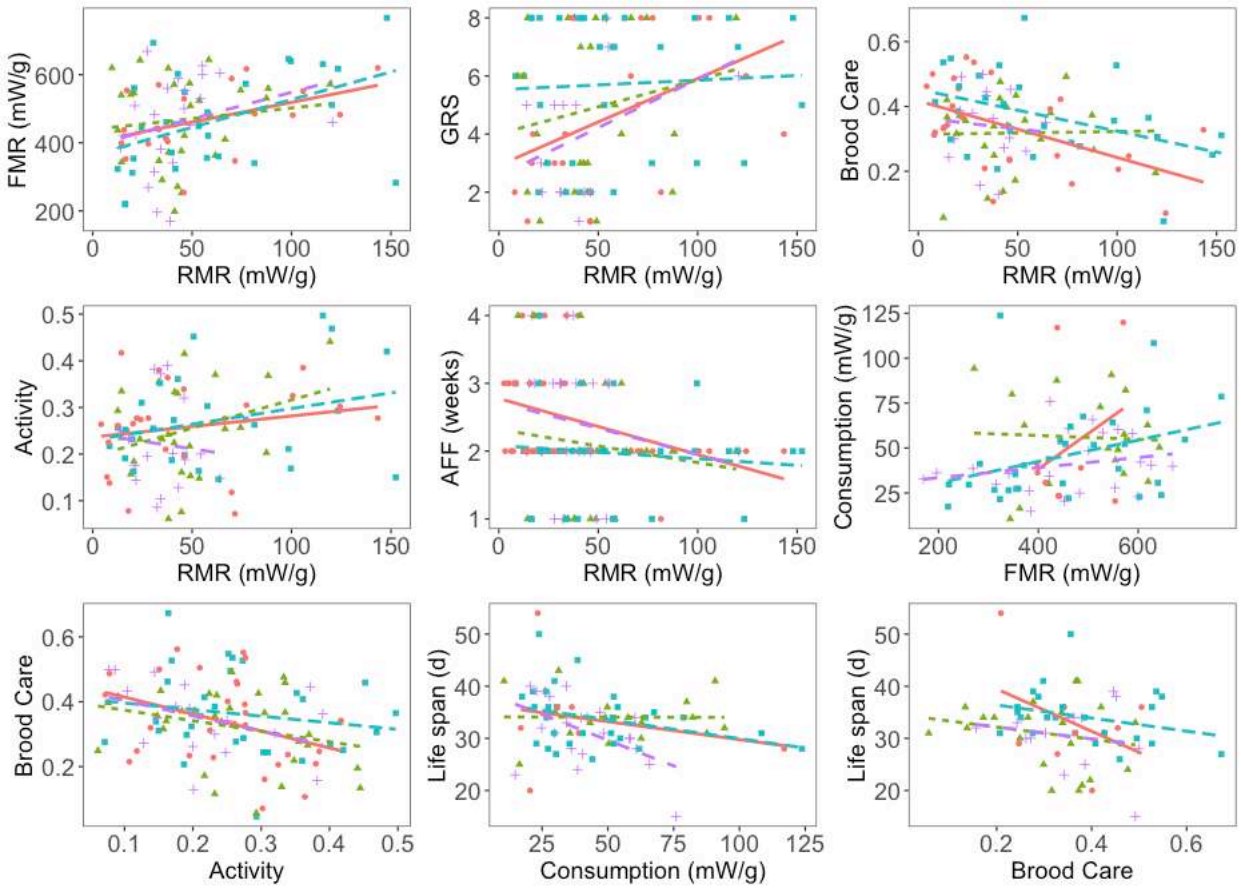

Figure 1.2. Relationships depicting the statistically significant paths from the top path model (Fig. 1.1), shown with untransformed raw data from source colonies 1-4 in red circles, green triangles, blue squares, and purple crosses, respectively, with regression lines shown in the same colors (print version: solid, short dash, long square dash, and long rounded dash, respectively). Each of these relationships (except FMR and consumption) upholds the patterns of the overall path model and source colony does not have a significant influence in any model (Table S1.1).

The structural equation model grouping behavioral, physiological, and life-history traits into separate categories showed support for a behavioral and a physiological axis, but not a life-history axis, nor an overarching phenotypic pace-of-life axis integrating all the three (Fig. 1.3A). FMR and GRS loaded significantly on the physiological axis, but RMR did not. None of the measured life-history parameters loaded significantly on the life-history axis. The more parsimonious SEM on the other hand, with a less restrictive structure that integrates all the traits simultaneously (Fig. 1.3B), showed significant positive loadings of RMR, FMR, GRS, activity

and consumption, indicating positive correlations among these traits, and significant negative loadings of AFF and brood care, indicating negative correlations with the aforementioned traits, directly on the phenotypic pace-of-life axis. The strength of the loadings indicates this pace-of-life axis to have a strong positive association with metabolic rate and activity and a strong negative association with brood care and AFF. The more parsimonious SEM (Fig. 1.3B) also exhibited a covariance structure among the traits that more closely resembled the nature and direction of the significant relationships seen in the top path model (Fig. 1.1) than the more restrictive SEM (Fig. 1.3A).

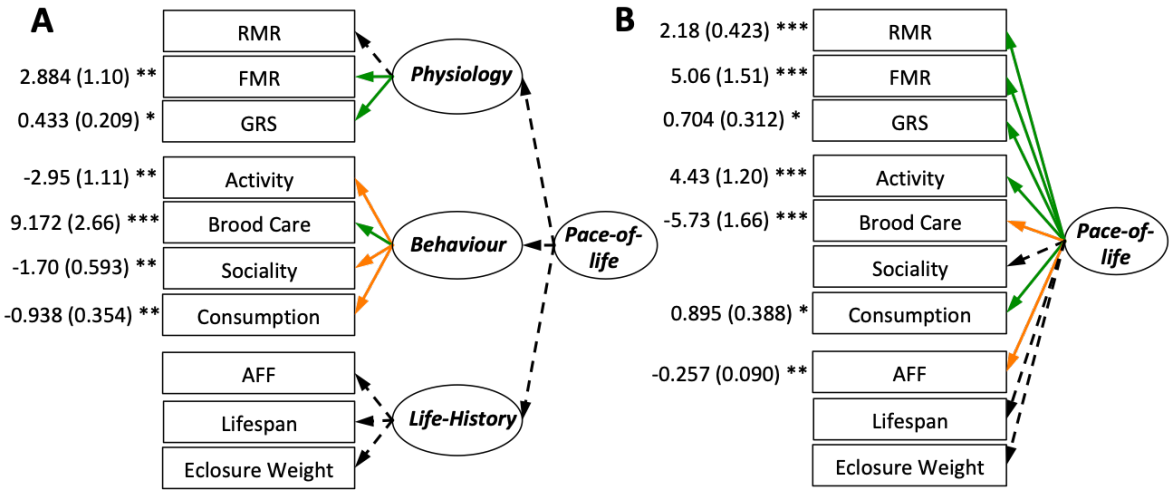

Figure 1.3. Structural equation models depicting the phenotypic covariance structure among measured variables (rectangles) manifested through unmeasured latent variables (ovals), with (A) the model in which traits are first grouped into assumed categories that were then loaded separately on a phenotypic pace-of-life axis, and (B) the model with a more parsimonious approach that simultaneously loads all traits on a phenotypic pace-of-life axis. Solid lines represent significant correlation coefficients with standard errors in parentheses, at $p \leq 0.001$ (***), $0.001 < p \leq 0.01$ (**), or $0.01 < p < 0.05$ (*) level. Orange arrows indicate negative correlation coefficients, green arrows indicate positive correlation coefficients while dotted lines represent non-significant correlation coefficients.

Discussion

The phenotypic covariance among the various traits shown here follows many, though not all, of the theoretical predictions of a phenotypic slow–fast pace-of-life axis linking behavioral,

physiological, and life-history traits in honeybees. The respective loadings of these traits on this phenotypic slow–fast axis indicate a 'fast' phenotype being manifested by higher MR, greater activity, engaging in less brood care, and making the risky life-history decision to leave the nest and begin foraging at an earlier age (Fig. 1.4). Conversely a 'slow' phenotype corresponds to lower MR, less activity, more brood care, and delaying foraging onset.

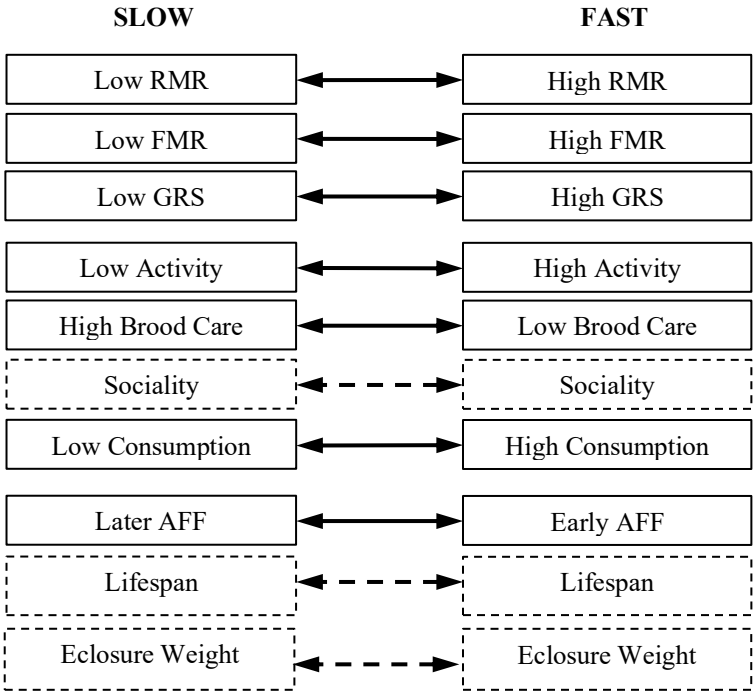

Figure 1.4. The phenotypic slow–fast axis in honeybees with solid boxes indicating traits which significantly covary along the axis and dotted boxes indicating traits which do not align with the axis.

By pairing a path analysis that assesses the causal relationship among a set of phenotypic traits with an assessment of the overarching nature of the phenotypic covariance that describes pace-of-life axis using a factor analysis, our study provides a comprehensive and robust description of the overall phenotypic pace-of-life axis than the more limited approach used in most previous studies. The factor analysis primarily provides a useful heuristic description of the nature of the covariance among all observed traits as the pace-of-life axis, while the path analysis provides a

more detailed picture of the proximate pathways that generate such a pattern. Some studies have interpreted pace-of-life as the covariance among three separate axes of physiology, behavior, and life-history (Debecker and Stoks 2019), which in our view ends up making unnecessarily restrictive assumptions, such as suggesting that all behavioral variables are more closely linked to one another than they are to any physiological or life-history variable, and so on. Our comparison of two model structures with and without such groupings and the higher support for the more parsimonious model suggests that it may be better to view the contribution of each trait to the slow–fast axis independently.

The best supported path model in our study suggests that variation in RMR is the causal variable underlying the variation in behavioral and life-history traits associated with the phenotypic pace-of-life axis. Furthermore, this path model indicates that behavioral traits primarily act as mediators influencing life-history, which is consistent with both theoretical predictions and empirical data regarding the pace-of-life theory (Réale et al. 2010; Santostefano et al. 2017; Dammhahn et al. 2018; Mathot and Frankenhuis 2018). Metabolic rate has often been cited as the fundamental biological rate that drives pace-of-life (Biro and Stamps 2010; Careau and Garland 2012; Arnqvist et al. 2017) and our results provide strong evidence in favor of this hypothesis. Our observation of a positive link between RMR and food intake also agrees with the prediction of this hypothesis regarding the association between the idling cost of metabolic rate and energetic demands (Nilsson 2002). These results provide support for the performance or acquisition model linking energetics to behavior in which the size of the metabolic machine, as expressed by RMR, determines energetic availability, in turn fueling activity levels (Careau et al. 2008). In this model, peak energetic output is also predicted to correlate with the baseline metabolic rate, seen here in

the positive association between FMR and RMR. Our results therefore suggest that RMR, by shaping energetic availability and idling costs, acts as the proximate driver of the behavioral and life-history axis that defines the variation in slow–fast pace-of-life phenotypes.

In honeybees, foraging behavior has been shown to be correlated with metabolic rate such that bees with higher FMR are more likely to engage in pollen foraging (Feuerbacher et al. 2003), a behavior that is also correlated with higher gustatory responsiveness (Pankiw and Page 2000). Our results showing the positive association between gustatory responsiveness and FMR as well as the positive effect of RMR on activity and its negative effect on brood care indicate how metabolic rate, by influencing both intranidal and foraging behavior, could be the fundamental determinant of how an individual contributes to colony performance. Our results therefore suggest that division of labor in social insect colonies could be shaped by variation in worker metabolic rate, underscoring how such variation might play a fundamental functional role in social evolution. Sociality, measured here as adult–adult contact through allogrooming and food exchange, was surprisingly not a part of the pace-of-life axis as expected. In retrospect, it however seems that our measurement of sociality in terms of social contacts, borrowed from general POLS theory, may have been somewhat narrow within our system, and that the contribution to the colony via foraging and brood rearing, both of which were found to be a part of the pace-of-life axis, are the more primary aspects of sociality.

Our study did not find support for a covariance between lifespan and other traits in the phenotypic pace-of-life axis, which could be due to how lifespan was measured by maintaining individuals in cages with *ad libitum* food. This likely minimized energetic demands and removed

the mortality risk associated with the natural environment. Lifespan may also be related to factors not measured here, such as oxidative damage, the effects of which can only be realized in the natural environment. Our lifespan measure therefore is more accurately a measure of senescence rather than survival which would require observations on free foraging bees, though the value of the lifespan measured here is quite similar to that measured in other studies (Rueppell et al. 2007). That said, our path analysis indicated that RMR acts as a proximate negative driver of brood care and positive driver of consumption behaviors (through FMR), both of which have a negative influence on lifespan. Although an inverse relationship between consumption and lifespan may seem somewhat counterintuitive at first, there is a large body of work regarding how caloric restriction positively contributes to lifespan in a wide variety of animals, including insects (Sohal and Weindruch 1996; Masoro 2005; Sinclair 2005). This shows how complex relationships among a large network of correlated traits can be revealed through path analytical approaches. The data also support previous findings that bees which engage more in brood care exhibit shorter lifespans (Amdam et al. 2009) and that the AFF of a honeybee worker exhibits a strong trade-off with survival and lifespan (Rueppell et al. 2007; Dukas 2008) and is therefore a risky life-history decision with a clear survival cost. Our data showing a strong covariance of AFF with the phenotypic pace-of-life axis is consistent with these findings and our path analysis indicating how RMR drives variation in AFF, potentially shaping the life-history trade-off observed between lifespan and AFF in other studies, shows how metabolic rate might be the underlying proximate mechanism shaping the pace-of-life axis.

It has been proposed that inter-individual differences in terms of a set of correlated traits, described as a foraging syndrome, in honeybees are based on sensorimotor differences, which in

turn reflect differences in signaling pathways, such as through biogenic amine cascades (Page et al. 2006). Biogenic amines such as octopamine, tyramine, dopamine, and serotonin play a major role in modulating gustatory responsiveness (Scheiner et al. 2002), a key variable associated with a number of behavioral differences and one that was a part of the pace-of-life axis measured here. The levels of the same biogenic amines have also been shown to change with age, suggesting that they are part of the signaling network that regulates age polytheism and division of labor in honeybees (Wagener-Hulme et al. 1999) and therefore could also be central to the pace-of-life axis. Similarly, juvenile hormone (JH), which is a key regulator of behavior in all insects and regulates age polyethism in many social insects (Robinson 1992), has in fact been referred to as a hormone which paces behavioral development in honeybees (Sullivan et al. 2000). How such hormonal and signaling pathways covary with MR are therefore key questions that can be explored further to provide an even more comprehensive picture of the causal pathways that shape the pace of life axis.

In social insects, the observed variation among colonies along the slow–fast axis (Bengston, Shin and Dornhaus 2017; Segev et al. 2017) suggests that selection can act on the phenotypic covariance structure at the group level. This leads to the interesting question of how a slow–fast phenotype at the group level emerges from the multitude of individual slow–fast phenotypes comprising the group since phenotypic variation with respect to the various pace-of-life traits, such as risk sensitivity and exploration tendency, is expressed at both the individual (Mayack and Naug 2011; Katz and Naug 2015) and the group level (Wray, Mattila and Seeley 2011; Blight et al. 2016). This question aligns well with the main thesis of the metabolic theory of ecology that the structural and functional properties at any level of biological organization are

explained by the variation in the metabolic rate of its components (Brown et al. 2004). Given our finding that the metabolic rate of an individual could play a fundamental proximate role in driving its behavioral profile and life-history trajectory, the question therefore becomes how the functional properties of a social group are emergent outcomes of the variation in metabolic rate among its members, a question that remains largely unexplored (Katz and Naug 2020).

The covariance structure across a set of traits that defines the slow–fast pace-of-life axis demonstrates biological constraints that may restrict the range of responses of an individual to environmental perturbation. However, social living concomitant with the inter-individual variation among the group members along the slow–fast axis, can make it possible to override some of these restrictions, allowing the group a greater range of responses. This interesting interplay of phenotypic covariance structures across different scales in social systems is a complex phenomenon that lacks a strong coherent framework. Future experimental work focusing on studying the functional properties of experimental groups with known distributions of slow–fast phenotypes can make important contributions to our understanding about group living and social evolution.

CHAPTER 2: METABOLIC RATE DIVERSITY INFLUENCES GROUP PERFORMANCE IN HONEYBEES

Introduction

The metabolic theory of ecology (Brown et al. 2004) argues that the metabolic rate (MR) at any level of biological organization, from cells to societies and ecosystems, is a composite function of the MR of its constituent parts. Extensive work has explored the widely observed allometry with respect to the scaling relationships between MR and size (Gillooly et al. 2001; West et al. 2002; White et al. 2019; Waters et al. 2010). In contrast, much less attention has been paid to the heterogeneity in MR among the constituent parts of a biological unit and to questions regarding how variation in MR among lower level units influences performance at a higher level of organization (Konarzewski and Diamond 1995; Konarzewski and Książek 2013; Woods 2014). This is an important consideration in the context of the large heterogeneity in MR that is common among biological units at any level of organization, whether they are different organs and tissues comprising an organism, or the different species that constitute an ecosystem.

At the level of a single organism, MR has been hypothesized to determine the rate at which it acquires and processes energy and is often considered the fundamental driver of its performance and productivity, or its pace of life (Careau et al. 2008; Glazier 2015). Performance or acquisition models of pace of life predict performance to be positively associated with the output of the metabolic engine (MR), while alternative allocation models predict negative associations based on trade-offs set by allocating a fixed energetic output towards different tasks, both of which functionally link MR to a suite of behavioral and life-history traits (Biro and Stamps 2010; Careau

and Garland 2012). MR has recently been shown to be associated with intraspecific differences in behavior and life-history in a variety of animals (Burton et al. 2011; Petterson et al. 2016; Krams et al. 2017; Mugel and Naug *in review*). However, how such interindividual differences in behavior and life history linked to variation in MR can shape performance parameters at the group level has rarely been addressed. Furthermore, since MR is directly related to energy demand and expenditure, the relationship between MR and performance is likely to show a strong interaction with resource availability, in turn suggesting that heterogeneity in MR within a group will interact with the resource environment to determine group-level performance (Katz and Naug 2020).

The effect of heterogeneity on group performance can be quantified as the 'diversity effect,' or the deviation in performance of a polymorphic group from the null expectation set by the performance of the same morphs in monomorphic groups (Loreau et al. 2001). This effect can be partitioned into two additive components: (1) a selection effect, or the disproportionate effect of a single morph, and (2) a complementarity effect, or the manner in which interactions between the different morphs influence the performance of the group. Both these effects can be either positive or negative, the former being indicative of asymmetric performance of one morph while the latter showing the action of niche partitioning or interference between the two. The diversity effect has been a major focus of community and ecosystem level questions regarding the benefits of biodiversity in terms of resilience, productivity and stability (Petchey and Gaston 2002; Cadotte et al. 2013), but has seldom been used as a framework for understanding how behavioral or physiological variation within a group contributes to collective performance (Takahashi et al. 2018). By partitioning the effects of group composition with respect to interindividual differences

in MR (and any co-varying traits) on group performance, productivity, and life-history, this approach lends itself extremely well to exploring the role of MR in shaping sociality.

Social insect colonies are often thought of as superorganisms in which the constituent individuals express considerable diversity in physiology, behavior, and life-history (Jeanson and Weidenmuller 2015), and yet contribute to fitness at the group level through their collective performance (Kennedy et al. 2017). Social insects are also seen to exhibit between-group variation at the colony level with respect to a number of behavior and life-history traits (Wray et al. 2011; Blight et al. 2016), which are outcomes of differences in this collective performance. Since selection primarily acts at the colony level in these groups (Seeley 1997; Fewell and Page 2000), how heterogeneity shapes division of labor and collective behavior is a question of significant interest in social evolution. Despite such phenotypic variability at multiple scales, research regarding the role of MR on sociality in insect colonies have almost exclusively focused on the question of metabolic scaling with size (Fewell and Harrison 2016; Waters et al. 2010), leaving unanswered the important question regarding the effects of metabolic diversity.

In the honeybee, *Apis mellifera*, workers exhibit considerable covariation in behavior, life-history, and physiology (Page et al. 2006; Tait and Naug 2020), including those related to differences in energetics and MR (Mayack and Naug 2011; Katz and Naug 2015; Harrison and Fewell 2002; Feuerbacher et al. 2003; Mugel and Naug *in review*). Using this as a background, in the current study we use genetic lines of honeybees with low and high MR, to explore how heterogeneity within a group in terms of MR interacts with the resource environment to shape group performance.

Methods

Experimental Design

Genetic lines of honeybees with different metabolic rates (MR) were raised based on the well-known slow (S) – fast (F) variation in malate dehydrogenase (MDH-1) allotypes (Coehlo and Mitton 1988; Harrison and Fewell 2002), in which the S and the F allele are associated with low and high MR, respectively. The allotypes of queens were determined by sampling and assaying the MDH-1 allele of six haploid drone offspring using gel electrophoretic techniques (providing a 95% chance of correctly identifying queen genotype; Feuerbacher et al. 2003). Six to ten different drones of known allotypes from multiple source colonies were then used to artificially inseminate queens of known allotypes to create crosses of the following 4 types: SS x S, FF x F (hereafter called SS and FF queens for the respective types of brood they produce), SS x F and FF x S (hereafter called SF queens that produce SF brood). The allotypes of these queens were verified from the young larvae produced from the first set of eggs laid by these queens. Colonies of each allotype were then set up using standard package bees of *Apis mellifera* and a queen of a specific type (FF, SS, SF). Frames with mature brood were extracted from multiple colonies of each type, kept in an incubator at 35 °C and 60 % Relative Humidity (RH) and newly emerged adults of each allotype from these were mixed together to create replicates of the four experimental group compositions described below.

The following experimental group compositions were created: monomorphic FF, monomorphic SS, monomorphic SF, and polymorphic SS and FF (1:1 ratio). Each group was made up of 50 newly emerged bees and housed in two conjoined plastic cages separated by a 0.3 cm gauge wire mesh with 25 bees on each side of the mesh, allowing contact and exchange of food

and pheromones across the entire group (Fig. S2.1). In the polymorphic group, the FF and the SS bees were separated on each side of the mesh, which allowed us to measure the performance of each morph separately in a polymorphic context. Each cage was made out of darkened plastic so that the bees were minimally disturbed by outside light and had a dimension 10 cm x 10 cm x 5 cm with wax comb material on the inside walls, holes for air circulation and attached with a modified syringe filled with sucrose solution from which the bees could feed *ad libitum*. One day prior to the performance assays, all groups were standardized to a size of 40 bees (20 bees per side) to allow an opportunity to supplement those groups with any early mortality with additional bees of the same age and allotype. Any group with 50% or higher mortality was discarded from the performance measurements. All groups were housed in an incubator maintained at 28 °C and 60% RH. Different replicates of each group composition were further assigned randomly to one of two resource environment treatments: a high resource environment of 30% w/w sucrose solution (HRE) and a low resource environment of 15% w/w sucrose solution (LRE). Thus, there were a total of 8 experimental treatments (4 group compositions x 2 resource environments).

Performance assays

MR was measured for each group by separately measuring each of the two halves of the group using flow through respirometry when the bees were 10 ± 2 days old. Each cage side with 20 bees was placed in a 380 mL sealed glass chamber and dry, CO_2 free air was run through the chamber at a rate of 800 mL/min and CO_2 was measured in the excurrent airstream with a FoxBox gas analyzer (Sable Systems). Each subgroup of bees was weighed immediately after respirometry and its mass-specific MR was calculated as the mean CO_2 production (in mL hr^{-1} g^{-1}) for the most stable 4 continuous minutes of CO_2 production (i.e. the 4 minutes with the lowest variance). This

MR value was transformed into a power output for all analyses (in mW g^{-1}) by multiplying it by 21.4 J per mL^{-1} CO_2 and dividing by 3600 J hr^{-1} W^{-1} (Feuerbacher et al. 2003; mass independent data shown in Fig. S2.2-2.3). Since the MR of each subgroup was measured separately to quantify the contribution of each morph in the polymorphic context, MR of the whole group was calculated as the sum of the two power outputs.

Food consumption of each subgroup was measured for 3 days following the MR measurement by recording the volume change in the feeding syringe in each cage and correcting for any evaporation and the actual number of bees alive on that day. The consumption volumes were converted to an energetic intake equivalent using concentrations and a conversion factor of 1 mg sugar = 16.7 J. Once again, the energetic intake of the whole group was calculated as a sum of the consumptions of the two subgroups.

The thermoregulatory performance of each group was measured over two periods of 45 minutes each, by placing each group of 40 bees in an incubator set to 18° C for a cold stress assay, and 38° C for a heat stress assay (± 10° C of the thermal neutral housing temperature at 28° C), in a random order. All assays were performed during daytime and remained the same across all treatments. A thermal probe (HOBO Systems) was inserted adjacent to the wax foundation in each side of the cage, which recorded the temperature every 10 seconds for 45 minutes (270 data points) for each subgroup. The temperature within an empty cage was measured simultaneously as a control and was subtracted from the measurement for each subgroup to produce a temperature residual. The thermoregulatory performance of the whole group was calculated by averaging the mean residual temperatures of each subgroup over the 45-minute period. Since better

thermoregulation in the heat stress assay is indicated by lower residual temperatures, for ease of interpretation and alignment with other traits in which greater values indicate better performance, these values were multiplied by -1.

The number of surviving bees in each group was checked daily and groups were terminated at the end of 35 days or when fewer than 10 bees remained alive (< 25% remaining), whichever occurred earlier. The repeatability of MR and thermoregulatory performance was assessed two days following their first measurement in only those groups in which fewer than 4 bees had died since the first measurement.

Data analysis

A total of 225 groups were assembled from 16 source colonies of the 3 allotypes with < 10% mortality on day 10 for the MR and thermoregulation assays, yielding a final sample size of 19-31 groups for each of the 8 treatments.

First, a linear model was used to test the interaction and main effects of group composition (monomorphic SS, FF, SF and polymorphic 1 SS: 1 FF) and resource environments (high and low) on MR, thermoregulatory performance under cold and heat stress, food consumption, and median survival. Post-hoc Tukey-adjusted *t*-tests were performed using the '*emmeans*' R package. See supplement for follow-up Kaplan-Meier survival analysis and Cox proportional hazard models on data pooled by treatment (Figs. S2.3-S2.4). Pearson's correlation tests were used to explore the correlations between the different traits within each environment after pooling data across the group compositions.

The effect of group composition on performance was further examined by calculating a diversity effect: the deviations in performance (ΔP) of a polymorphic group from a null expectation based on the performance of the same morphs in monomorphic groups. The expected performance of a 1 SS: 1 FF polymorphic group is the mean of that for SS and FF monomorphic groups. This expected value was compared to the observed polymorphic group performance values with a one-sample *t*-test. ΔP can be further partitioned into a complementarity effect (the effect of interactions between different morphs, $N\overline{\Delta RP}\bar{M}$) and a selection effect (the disproportionate effect of any one morph, $Ncov(\Delta RP, M)$):

$$\Delta P = N\overline{\Delta RP}\bar{M} + Ncov(\Delta RP, M)$$

where,

$$\sum_i \Delta \mathrm{R}P_i M_i = \sum_i \mathrm{R}P_{O,i}\, M_i - \sum_i \mathrm{R}\, P_{\mathrm{E},i}$$

where N is the number of morphs, $\bar{M}$ is the mean performance of all monomorphic groups, and $\overline{\Delta RP}$ is the mean deviation in observed relative performance of all morphs in polymorphic groups as compared to their performance in monomorphic groups. $\Delta \mathrm{R}P_i$ is calculated as the difference between the expected relative performance of morph I ($\mathrm{R}P_{E,i}$) and the observed relative performance of morph I ($\mathrm{R}P_{O,i}$) in the polymorphic group, ($\mathrm{R}P_{O,I} - \mathrm{R}P_{E,i}$). $\mathrm{R}P_{E,I}$, is determined by the ratio of morph I in the polymorphic group, and $\mathrm{R}P_{O,I}$, is the observed performance of I in the polymorphic groups ($P_{O,i}$) divided by the performance of morph I in its monomorphic group (M_i). All data analyses were completed in R.

Results

Metabolic Rate

A linear model indicated a significant interaction between group composition and resource environment influencing MR ($F_{3, 217} = 5.41$, $p = 0.001$; main effects of resource environment: $F_{1,217} = 49.51$, $p < 0.001$; group composition: $F_{3,217} = 10.62$, $p < 0.001$, Fig. 2.1). Post-hoc tests revealed that in the LRE, the monomorphic FF groups exhibited a significantly higher MR than the monomorphic SS groups ($t_{217} = 2.69$, $p = 0.03$), but was similar to monomorphic SF and polymorphic groups (see Table S1 for all comparisons). The MR of the monomorphic SS groups was significantly lower than polymorphic groups ($t_{217} = 2.77$, $p = 0.03$) though similar to monomorphic SF groups, the two of which were statistically similar to each other (Fig. 2.1).

Post-hoc tests revealed that compared to the LRE, MR was significantly higher in the HRE for all groups except the polymorphic group (FF: $t_{217} = 7.04$, $p < 0.001$; SS: $t_{217} = 5.04$, $p < 0.001$; SF: $t_{217} = 5.44$, $p < 0.001$; Polymorphic: $t_{217} = 1.49$, $p = 0.14$). In the HRE, the monomorphic FF groups exhibited a significantly higher MR than all other groups (SS: $t_{217} = 4.21$, $p < 0.001$; SF: $t_{217} = 3.43$, $p = 0.004$; Polymorphic: $t_{217} = 5.31$, $p < 0.001$), while the monomorphic SS, SF, and polymorphic groups were not significantly different from one another (Fig. 2.1).

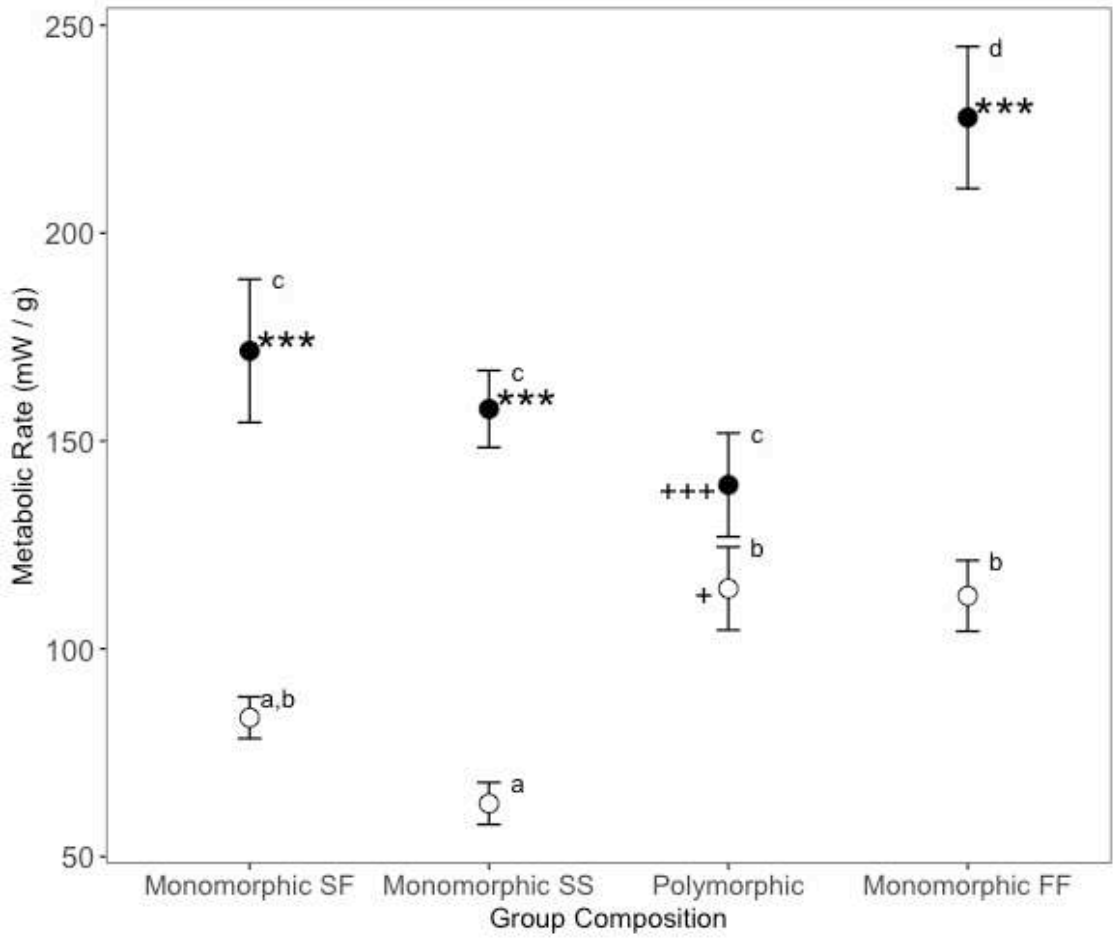

Figure 2.1. Group level MR is affected by an interaction between group composition and resource environment. Points represent mean ± SE in high (filled) and low (open) resource environments. Asterisks denote a significant difference across resource environments for a group composition (* $0.01 < p < 0.05$; ** $0.001 < p < 0.01$; *** $p < 0.001$). Different letters represent significant differences between different group compositions within a resource environments (*a,b* for low, *c,d* for high). Crosses (+) denote a significant deviation in performance of the polymorphic group from null expectation based on a one-sample *t*-test.

In the LRE, the observed polymorphic group mean MR of 114.5 mW g^{-1} was significantly higher than the expected MR of 87.75 mW g^{-1} ($t_{28} = 2.68$, $p = 0.01$, Fig. 2.1), and in the HRE, the observed polymorphic group mean MR of 139.4 mW g^{-1} was significantly lower than the expected MR of 192.75 mW g^{-1} ($t_{27} = 4.28$, $p < 0.001$, Fig. 2.1). These significant deviations from expectation, or diversity effects, were comprised of a non-significant selection effect ($t_{28} = 0.79$, $p = 0.43$), and a significant positive complementarity effect ($t_{28} = 2.63$, $p = 0.01$; Fig. 2.6) in the

LRE, and a non-significant selection effect (t_{28} = 1.89, p = 0.07) and a significant negative complementarity effect (t_{28} = 3.58, p = 0.001) in the HRE (Fig. 2.6).

Energetic Intake

Energetic intake was not significantly predicted by an interaction between resource environment and group composition ($F_{3, 218}$ = 0.92, p = 0.43), nor the main effect of group composition ($F_{3, 218}$ = 1.59, p = 0.19) but was predicted by resource environment ($F_{1, 218}$ = 17.9, p < 0.001; Fig. 2.2). Therefore, all group types were statistically similar within each resource environment (see Table S2.2 for post-hoc tests). Energetic intake was significantly higher for all groups in the HRE (FF: t_{218} = 4.24, p < 0.001; SS: t_{218} = 3.94, p < 0.001; SF: t_{218} = 4.49, p < 0.001; Polymorphic: t_{220} = 2.37, p = 0.02; Fig. 2.2).

In the LRE, the energetic intake of 16.6 mW g^{-1} in polymorphic groups was not significantly different from the expected value of 14.5 mW g^{-1} (t_{28} = 1.26, p = 0.22). However, in the HRE, the energetic intake of 21.15 mW g^{-1} in polymorphic groups was significantly lower than the expected value of 24.05 mW g^{-1} (t_{28} = 2.30, 1.78, p = 0.02; Fig. 2.2). The non-significant diversity effect in LRE was comprised of non-significant selection and complementarity effects (t_{28} = 0.41, p = 0.69, and t_{28} = 1.28, p = 0.21, respectively). The significant negative diversity effect in HRE was comprised of both non-significant selection and complementarity effects (t_{26} = 1.19, p = 0.24, t_{26} = 1.56, p = 0.13, respectively; Fig. 2.6).

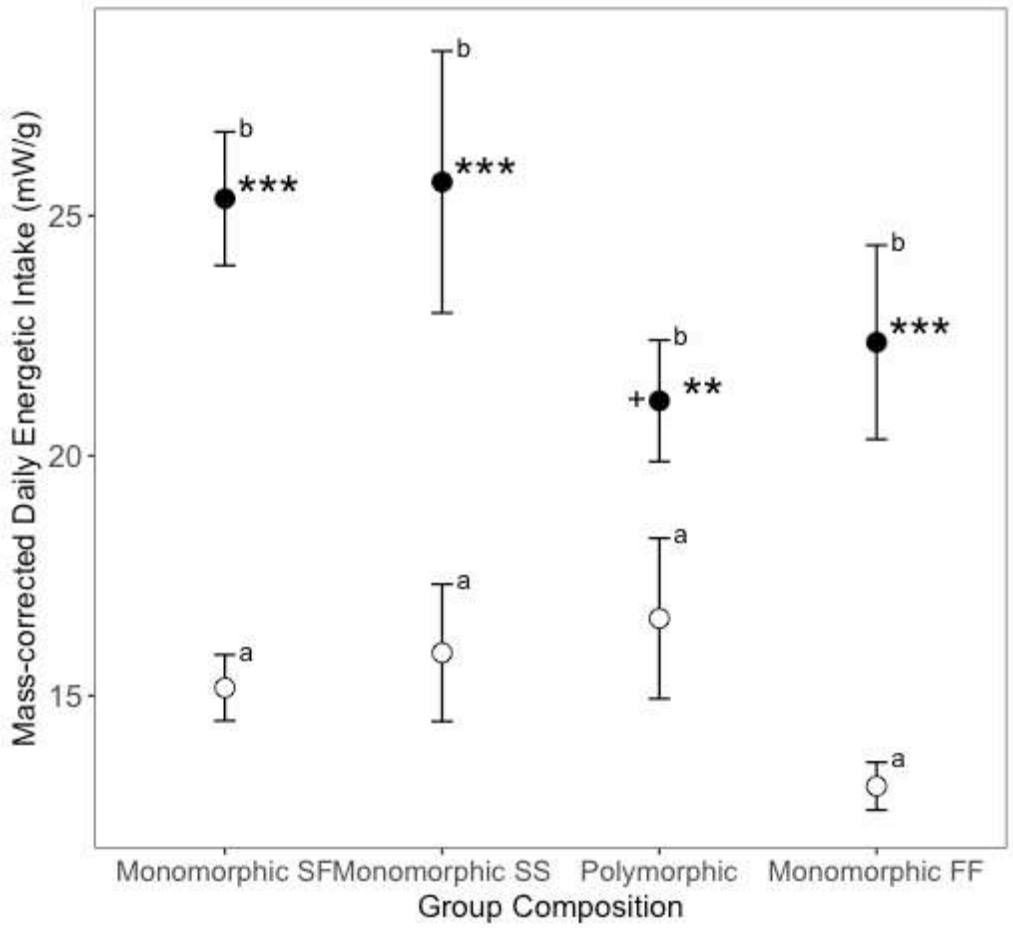

Figure 2.2. Group level energetic intake is affected by resource environment but not group composition. Points represent mean ± SE in high (filled circles) and low (open circles) resource environments. Asterisks (*) denote a significant difference across resource environments for a group composition, and crosses (+) denote a significant deviation in performance of the polymorphic group from null expectation. Different letters represent significant differences across different group compositions within a resource environment (*a* for low, *b* for high).

Thermoregulatory performance in cold and heat stress

Thermoregulatory performance of the group under cold stress was not influenced by a resource environment by group composition interaction ($F_{3, 213} = 2.58$, $p = 0.05$), nor a group composition main effect ($F_{3, 213} = 0.66$, $p = 0.58$) but was influenced by the resource environment ($F_{1, 213} = 21.41$, $p < 0.001$). Post-hoc *t*-tests revealed that in the LRE, the monomorphic FF groups had significantly better thermoregulatory performance with a higher mean residual temperature

than the monomorphic SS groups ($t_{213} = 2.83$, $p = 0.02$) and the polymorphic groups ($t_{213} = 2.73$, $p = 0.03$), but was not different from the monomorphic SF groups, the performance of which was similar to the other two groups (Table S2.3, Fig. 2.3). For all groups, performance was significantly better in the HRE (FF: $t_{213} = 4.63$, $p < 0.001$; SS: $t_{213} = 7.46$, $p < 0.001$; SF: $t_{213} = 6.38$, $p < 0.001$; Polymorphic: $t_{213} = 6.84$, $p < 0.001$; Fig. 2.3). In the HRE, all groups were statistically similar (Table S2.3).

In the LRE, the observed mean residual temperature of 0.53 °C for polymorphic groups was significantly lower than the expected value of 0.70 °C ($t_{28} = 2.15$, $p = 0.04$), indicating a significant negative diversity effect on thermoregulatory performance (Fig. 2.3). In the high resource environment, the observed mean residual temperature of 1.59 °C for polymorphic groups was not significantly different from the expected mean of 1.72 °C ($t_{27} = 1.15$, $p = 0.26$), indicating a non-significant diversity effect on thermoregulatory performance (Fig. 2.3). The significant negative diversity effect on thermoregulatory performance under cold stress in the LRE is comprised of a significant negative selection effect ($t_{28} = 2.45$, $p = 0.02$), and a non-significant complementarity effect ($t_{28} = 1.34$, $p = 0.19$; Fig. 2.6). The non-significant diversity effect in the HRE is comprised of a significant selection effect ($t_{28} = 2.95$, $p = 0.006$), and a non-significant complementarity effect ($t_{28} = 1.41$, $p = 0.17$; Fig. 2.6).

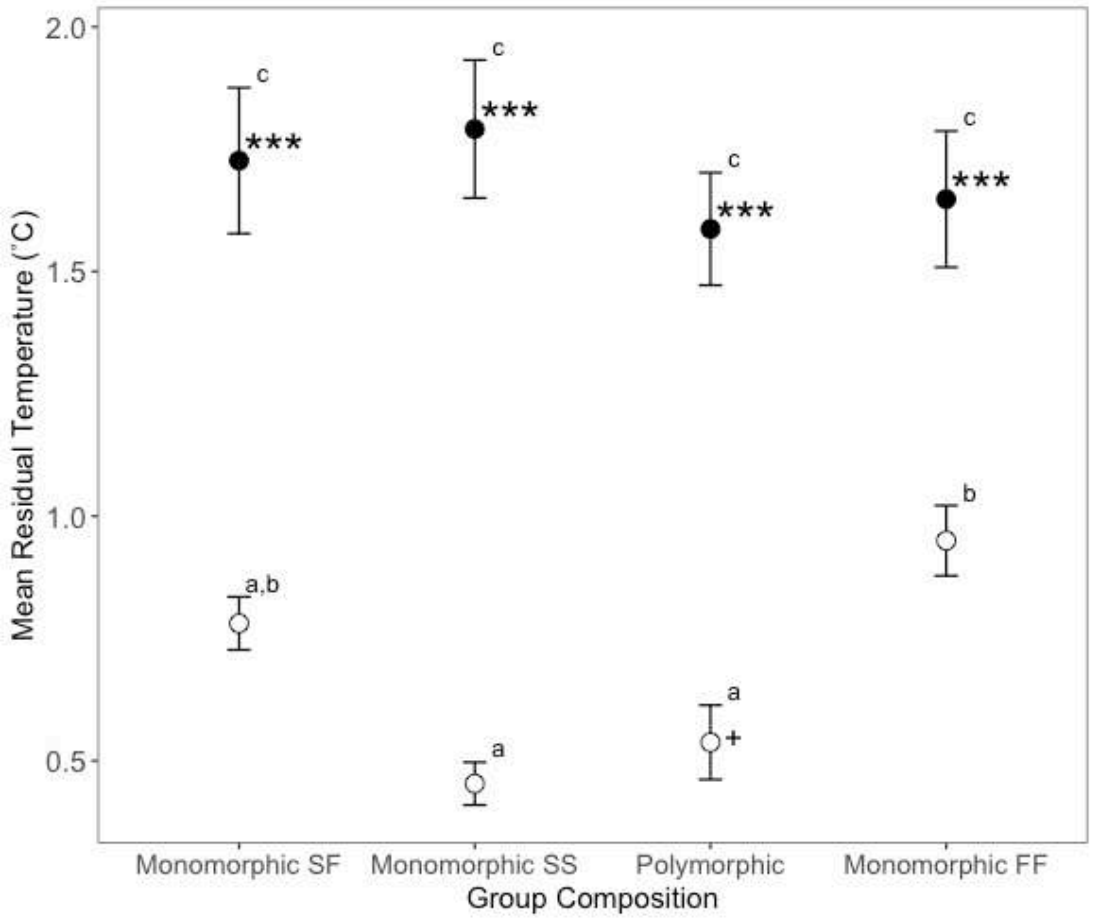

Figure 2.3. Thermoregulatory performance under cold stress is affected by resource environment but not group composition. Points represent mean ± SE in high (filled) and low (open) resource environments. Asterisks (*) denote a significant difference across resource environments for a group composition. Different letters represent significant differences across different group compositions within a resource environment (*a-b* for low, *c* for high). Crosses (+) denote a significant deviation in performance of the polymorphic group from null expectation.

Thermoregulatory performance under heat stress was not influenced by resource environment ($F_{1, 212}$ = 0.71, p = 0.40), group composition ($F_{3, 212}$ = 1.21, p = 0.16), nor an interaction between the two ($F_{1, 212}$ = 0.57, p = 0.63). All groups performed similarly both within and across the two resource environments (Table S4; Fig. 2.4). In the LRE, the observed (inverted) mean residual temperature of 0.17 °C for the polymorphic groups was not significantly different than the expected value of 0.23 °C (t_{27} = 1.48, p = 0.15), indicating a non-significant diversity

effect (Fig. 2.4). In the HRE, the observed (inverted) mean residual temperature of 0.12 °C in the polymorphic group was significantly lower than the expected value of 0.24 °C (t_{25} = 2.63, p = 0.01), indicating poorer than expected performance due to a significant diversity effect (Fig. 2.4).

In the LRE, the non-significant diversity effect was driven by a non-significant selection effect (t_{27} = 1.88, p = 0.07) and a non-significant complementarity effect (t_{27} = 1.42, p = 0.17, Fig. 2.6). The significant diversity effect on thermoregulatory performance under heat stress in the HRE was driven by a non-significant selection effect (t_{27} = 0.33, p = 0.74) and a significant negative complementarity effect (t_{26} = 2.74, p = 0.01, Fig. 2.6).

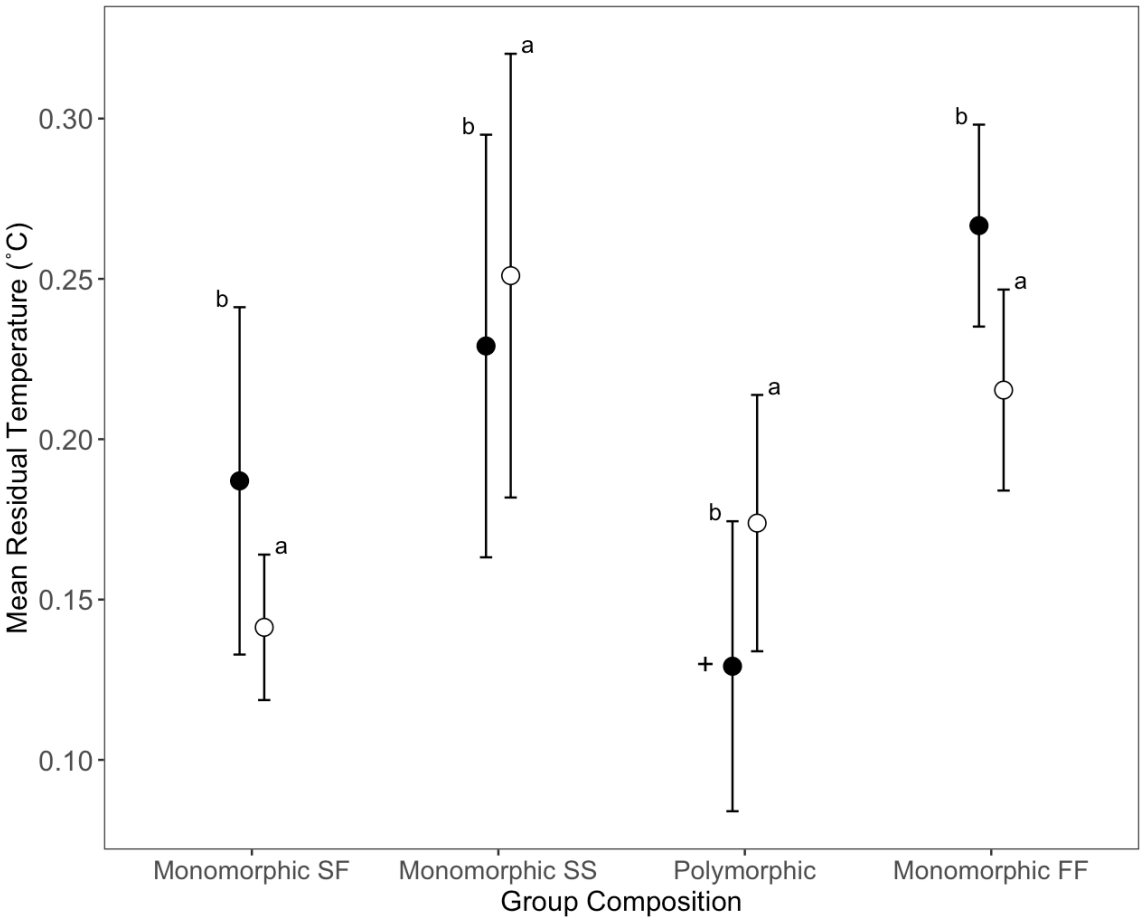

Figure 2.4. Thermoregulatory performance under heat stress is not affected by either resource

environment or group composition. Points represent mean residual temperatures ± SE (inverted for ease of understanding) in high (filled) and low (open) resource environments. Different letters represent significant differences across different group compositions within a resource environment (*a* for low, *b* for high). Crosses (+) denote a significant deviation in performance of the polymorphic group from null expectation.

Survival

The linear model on median survival time (longevity) revealed significant effects of group composition ($F_{3, 217} = 9.52$, $p < 0.001$), resource environment ($F_{1, 217} = 17.99$, $p < 0.001$), but not an interaction between the two ($F_{3, 217} = 1.41$, $p = 0.24$). In the LRE, the monomorphic FF groups had significantly higher survival than all other group compositions, which were all statistically similar to each other (FF-SS $t_{217} = 6.99$, $p < 0.001$; FF-SF $t_{217} = 6.14$, $p < 0.001$; FF-Polymorphic $t_{217} = 5.42$, $p < 0.001$; SS-SF $t_{217} = 1.63$, $p = 0.36$; Table S2.5; Fig. 2.5). All group compositions survived longer in the HRE than their counterparts in the LRE (FF: $t_{217} = 4.42$, $p < 0.001$; SS: $t_{217} = 6.34$, $p < 0.001$; SF: $t_{217} = 6.21$, $p < 0.001$; Polymorphic: $t_{217} = 5.67$, $p < 0.001$; Fig. 2.5). In the HRE, the monomorphic FF groups had significantly higher survival than all other group compositions, which were all statistically similar to each other (FF-SS $t_{217} = 4.80$, $p < 0.001$; FF-SF $t_{217} = 4.18$, $p < 0.001$; FF-Polymorphic $t_{217} = 3.83$, $p < 0.001$; Table S2.5; Fig. 2.5).

In the LRE, the observed median survival of 14.6 days for the polymorphic groups was significantly lower than the expected 15.7 days ($t_{28} = 2.94$, $p = 0.006$), while in the HRE the polymorphic groups exhibited similar median survival of 19.0 days to the expected 20.05 days ($t_{27} = 1.83$, $p = 0.07$; Fig. 2.5). The significant negative diversity effect on survival in the LRE was comprised of selection and complementarity effects that were both non-significant ($t_{28} = 1.84$, $p = 0.07$; $t_{28} = 0.34$, $p = 0.74$, respectively). The non-significant diversity effect in the HRE was

comprised of non-significant selection and complementarity effects ($t_{28} = 1.61$, $p = 0.12$; $t_{28} = 1.04$, $p = 0.31$, respectively; Fig. 2.6).

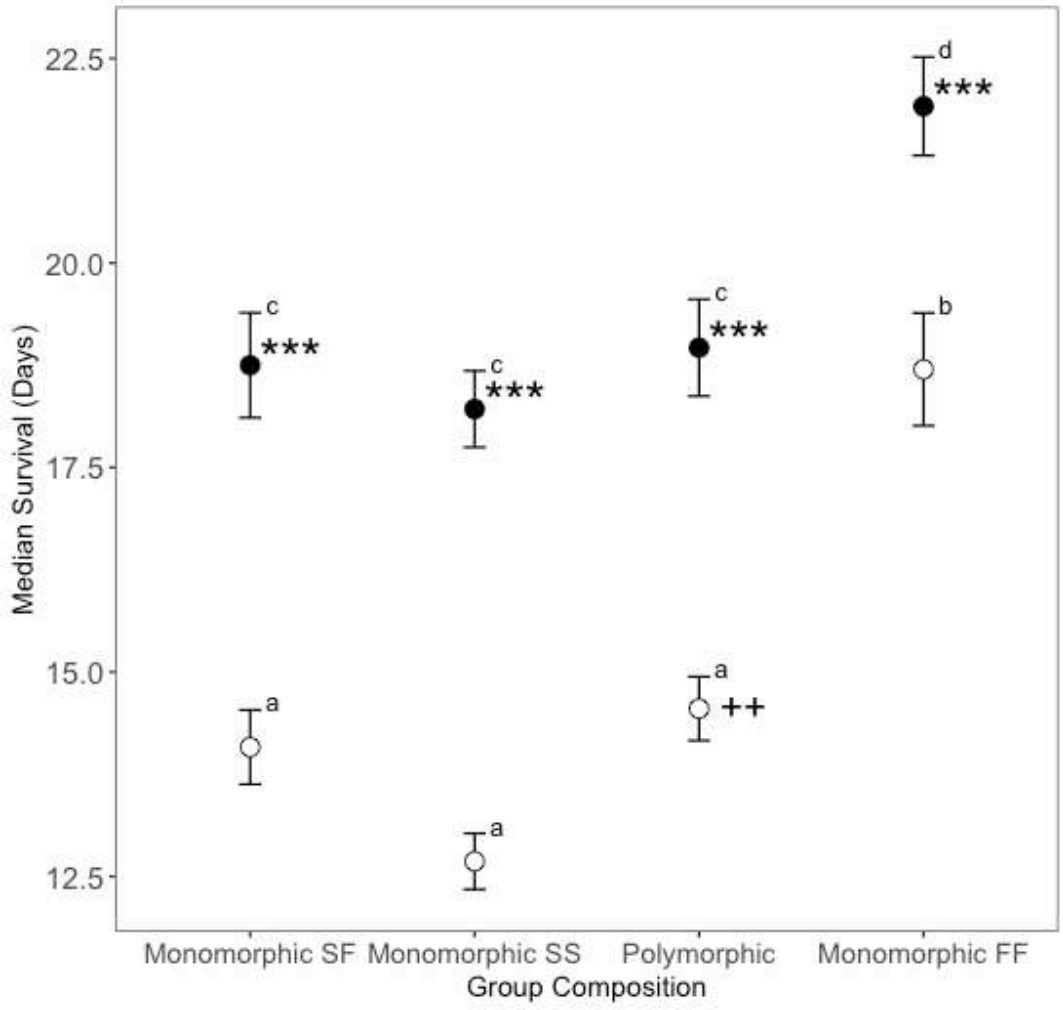

Figure 2.5. Median survival is affected by resource environment and group composition. Points represent mean ± SE in high (filled) and low (open) resource environments. Asterisks (*) denote a significant difference across resource environments for a group composition. Different letters represent significant differences across different group compositions within a resource environment (*a-b* for low, *c-d* for high). Crosses (+) denote a significant deviation in performance of the polymorphic group from null expectation.

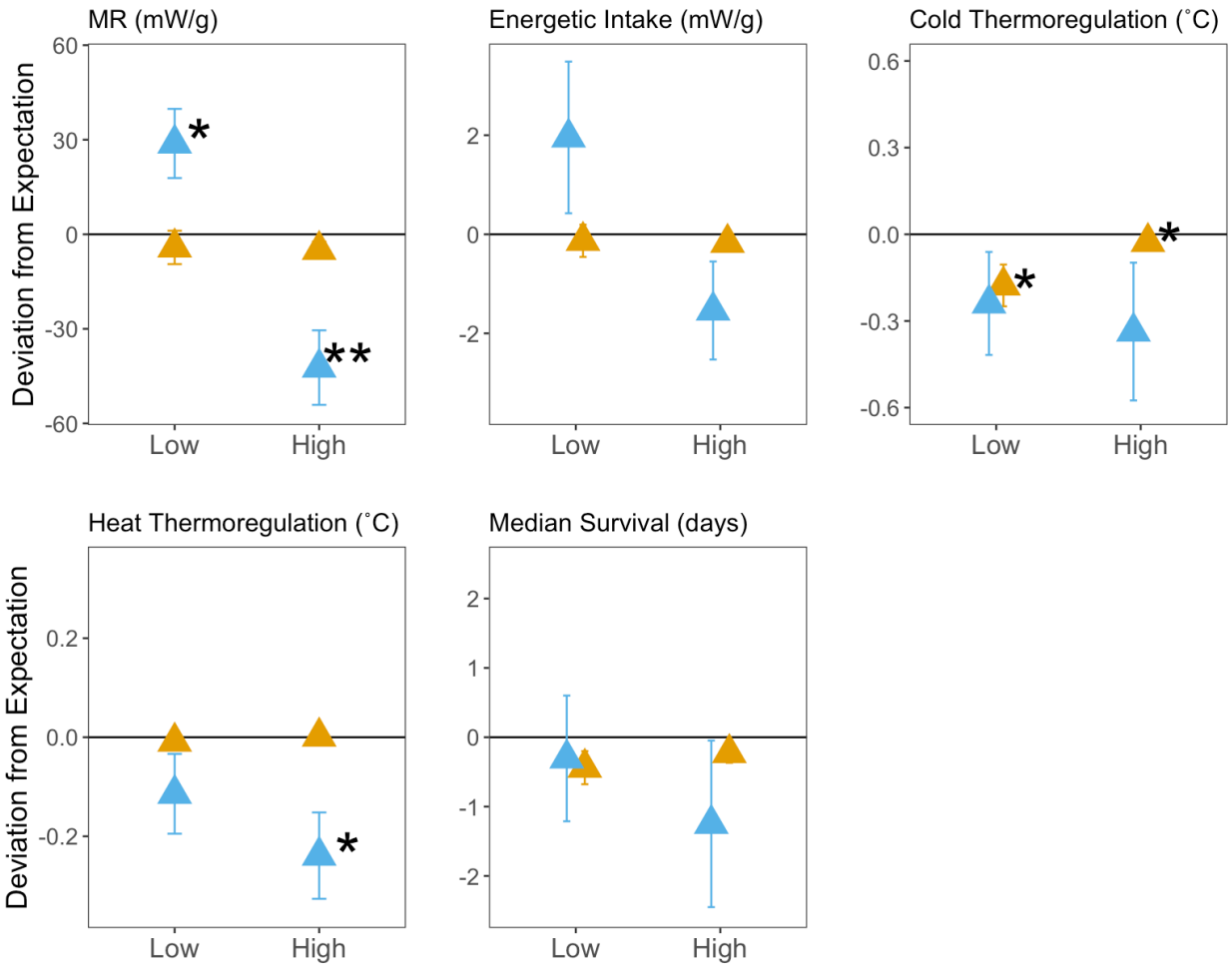

Figure 2.6. The diversity effects for all measured traits parsed into selection (orange) and complementarity (blue) components across low and high resource environments with SE bars. Components significantly different from 0 are depicted with asterisks (one sample t-tests, * $0.01 < p < 0.05$, ** $0.001 < p < 0.01$). Y-axes represent deviations from null expectation of the polymorphic groups based on 1 FF: 1 SS composition.

Table 2.1. Summary of diversity effects and their components for all measured traits across the two resource environments. Significant diversity effects, or the deviation of polymorphic performance from null expectation, and their selection and complementarity components are shown at the * $0.01 < p < 0.05$, ** $0.001 < p < 0.01$, and *** $p < 0.001$ levels.

Performance Measure	Resource Environment	Diversity Effect	Selection Component	Complementarity Component
MR	Low	* *Positive*	*n.s.*	* *Positive*
	High	*** *Negative*	*n.s.*	** *Negative*
Energetic Intake	Low	*n.s.*	*n.s.*	*n.s.*
	High	* *Negative*	*n.s.*	*n.s.*
Thermoregulation (Cold stress)	Low	* *Negative*	** *Negative*	*n.s.*
	High	*n.s.*	* *Negative*	*n.s.*
Thermoregulation (Heat stress)	Low	*n.s.*	*n.s.*	*n.s.*
	High	* *Negative*	*n.s.*	* *Negative*
Median Survival	Low	** *Negative*	*n.s.*	*n.s.*
	High	*n.s.*	*n.s.*	*n.s.*

Correlation between traits

Pearson's pairwise correlations on data pooled across group compositions, separated by resource environment, revealed that in the LRE, MR was significantly positively correlated with energetic intake ($r = 0.22$, $t_{107} = 2.33$, $p = 0.02$), thermoregulatory ability in cold stress ($r = 0.21$, $t_{106} = 2.13$, $p = 0.03$) and survival ($r = 0.38$, $t_{107} = 4.22$, $p < 0.001$). Survival was also significantly positively correlated with thermoregulatory ability in cold stress ($r = 0.55$, $t_{106} = 6.83$, $p < 0.001$) and heat stress ($r = 0.22$, $t_{106} = 2.15$, $p = 0.03$), and significantly negatively correlated with energetic intake ($r = -0.25$, $t_{107} = 2.57$, $p = 0.01$). No other traits were significantly correlated (Fig. 2.7A). In the HRE, MR was significantly positively correlated with energetic intake ($r = 0.29$, $t_{114} = 3.1$, $p = 0.003$) and survival ($r = 0.32$, $t_{114} = 3.74$, $p < 0.001$), and survival was significantly positively correlated with thermoregulatory ability in cold stress ($r = 0.26$, $t_{112} = 2.96$, $p = 0.004$), while no other traits were correlated (Fig. 2.7B).

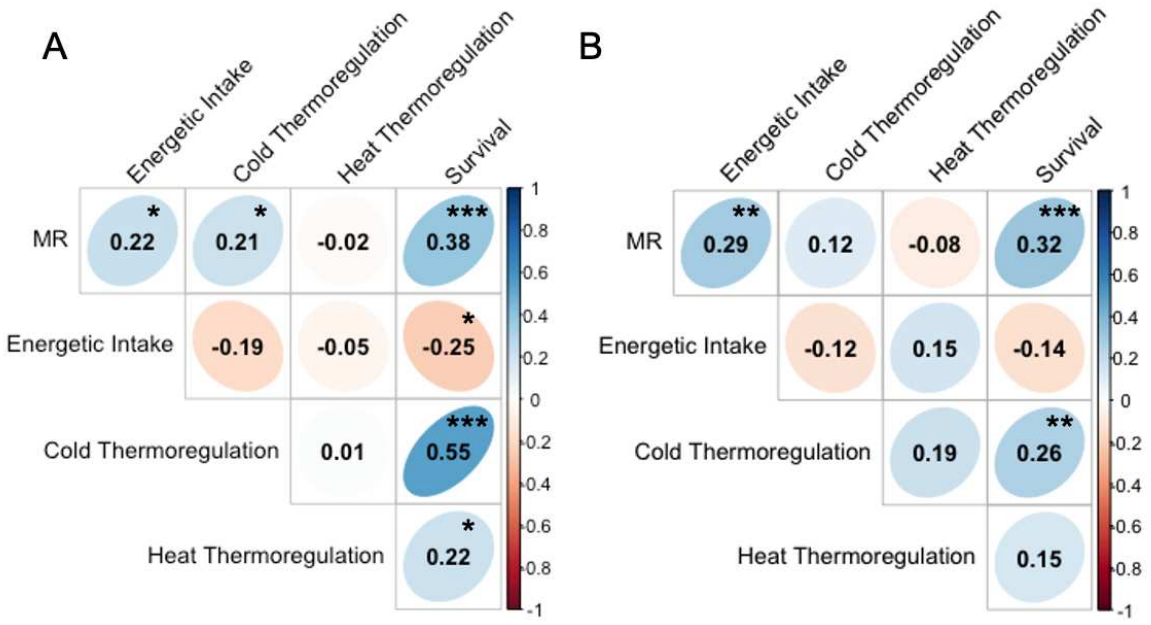

Figure 2.7. Pearson's correlation coefficients between traits pooled across group compositions in low (A) and high (B) resource environments. Significant coefficients shown with * $0.01 < p < 0.05$, ** $0.001 < p < 0.01$, and *** $p < 0.001$.

Discussion

The experimental design of our study substantiates and extends theoretical arguments that MR at any level of biological organization is a composite function of its component units (Brown et al. 2004). In this context, one striking result of this study is that the MR and the performance on every trait are statistically similar for the polymorphic 1 SS: 1 FF and the monomorphic SF groups. This is interesting because these two groups are similar with respect to the MDH allelic diversity, differing only in how the S and the F allele are allocated among individuals. From the framework of such a perspective, in this study we focus on how the performance of a biological unit is shaped by the heterogeneity in MR among its constituent units.

In the HRE, all groups exhibited a significantly higher MR and energetic intake, as well as better cold stress thermoregulation and longer survival than their counterparts in the LRE. While

the finding that performance is dependent on energetic availability is not itself surprising, the fact that there appears to be a robust group-level syndrome correlating all these traits that is more noticeable during resource restriction suggests these links to have energetic considerations at their root. Indeed, the positive association of MR with survival and thermoregulation traits extends the predictions of the acquisition or performance model, wherein the size of the metabolic engine, fueled by energetic intake, determines the energetic availability for thermoregulation and long-term survival (Careau et al. 2008; Biro and Stamps 2010) to the group level. Furthermore, in both environments, MR and energetic intake are positively correlated, extending the idea that higher MR is accompanied by higher maintenance costs for the upkeep of the metabolic engine (Nilsson 2002), to the group level.

The support for the performance model therefore suggests that morphs with a higher MR should outperform others, and indeed, the FF groups outperformed the SS groups on nearly every performance measure here, in both resource environments. This raises the provocative question as to the advantages conferred by the 'slow' (S) allele that maintain its presence in the population. Among natural honeybee populations, a latitudinal and thermal cline in F and S allele frequencies has been reported, suggesting a role of temperature mediated selection for MR (del Lama et al. 2004; Nielsen et al. 1994; Hatty and Oldroyd 1999). Given our results that 'fast' high MR workers are better thermoregulators per capita, though also demand more food, an intriguing possibility worth studying in greater detail is whether the 'slow' workers with lower MR and energetic demands are better equipped as a group to survive the long over-wintering period in higher latitudes without foraging through group-level thermoregulation and clustering a larger number of bodies that can be maintained with a lower total energy consumption.

Heterogeneous groups consisting of both 'slow' and 'fast' workers may therefore be better able to meet the complexity of demands faced by a honeybee colony over its annual colony cycle. Previous work has found MR negatively correlates with brood care and positively correlates to a longer intranidal period (Mugel and Naug *in review*), suggesting that slow workers with a lower MR serve other critical social roles in the colony division of labor. Fast workers perform more energetically demanding tasks, such as pollen foraging (Feurerbacher et al. 2003) and might be involved in scouting new resources (Tait and Naug 2020) while slower workers provide the workforce of recruits for extracting resources more efficiently from known locations. Honeybees often maximize foraging efficiency rather than gain rate (Schmid-Hempel et al. 1985), but whether individuals with different MR maximize different foraging currencies, or how colonies maximize energetic efficiency through division of labor among workers with different MR remains an interesting question. Such heterogeneity in the behavior of honeybee workers with different MR would suggest that colonies with mixed distributions of low and high MR phenotypes would outperform more monomorphic colonies.

Contrary to expectation, however, we did not find a positive diversity effect in the polymorphic group for nearly all traits in any of the two resource environments tested here. A number of studies have demonstrated the benefits of genetic diversity to population fitness (Takahashi et al. 2018), including in honeybees (Tarpy 2003; Jones et al. 2004; Oldroyd and Fewell 2007; Mattila and Seeley 2007), albeit for a handful of traits. However, it has also been shown that groups with too much heterogeneity may lack coordination in certain contexts, resulting in worse than expected collective output (Page et al. 1995; Neumann and Moritz 2000; Arathi & Spivak

2001; Jolles et al. 2020), and the polymorphic group tested here with a 1:1 ratio of the two MR allotypes represents the highest possible heterogeneity for a dimorphic scenario. The significant selection effects observed here suggest asymmetric performance of the two morphs, which can also lead to polymorphic populations to underperform under positive frequency dependent processes (Takahashi et al. 2018). However, the more intuitive appeal for the positive effects of diversity have led to such mechanisms remaining underexplored. While the polymorphic groups generally underperformed expectations, they did outperform the lowest performers, the monomorphic SS groups, suggesting a competitive advantage of heterogeneity against certain types of groups. It has been suggested that complex frequency-dependent processes may be involved in conferring an advantage to groups with a diversity of MR phenotypes, especially in a changing environment (Katz and Naug 2020), and empirical work testing this idea is needed.

Each trait responded differently to heterogeneity across the two resource environments, and significant diversity effects were shaped by both selection and complementarity components. Indeed, six of the ten trait measurements (5 traits x 2 environments) showed significant diversity effects, one of which was driven by selection effect, three by complementarity effects, and interestingly, two driven by the additive effect of both nonsignificant selection and complementarity effects. Significant selection effects suggest disproportionate performance of one of the morphs shapes group output, often referred to as ‘keystone individuals’ (Jolles et al. 2020). Here, negative selection effects indicate group thermogeneration ability conforms to its lowest performing members. Significant complementarity effects indicate complex feedback processes such as niche partitioning or facilitation for positive effects, or interference when negative. The negative complementarity in thermoregulation under heat stress suggest fanning behavior, used by

bees to lower hive temperature, to be a positive-frequency dependent process such that certain levels of heterogeneity could downregulate performance, which supports previous findings about the complexity of the fanning response (Garrison et al. 2018). Such group level processes interacted with the resource environment such that polymorphic group MR was significantly lower in the HRE and significantly higher in the LRE than expected, indicating the importance of understanding how group composition can show complex interactions with environmental parameters to shape performance. The significant negative diversity effect on survival was shaped by highly variable but negative complementarity and a marginally non-significant selection effect suggesting that in a group, survival may too conform to the lowest performing members, without consistent patterns of interaction between MR morphs.

This study uses a novel experimental and analytical framework for asking questions about how heterogeneity in a trait at the individual level shapes performance at the group level and parses any observed diversity effect into selection and complementarity components. We demonstrate that a number of important group-level performance measures are influenced by an interaction between the diversity in MR phenotypes within a group and the resource environment. Social insects provide excellent models to ask such questions regarding the influence of metabolic diversity at a higher level of biological organization as a large number of parameters can be measured in terms of individual and group level performance in experimentally created groups of different compositions. A natural extension of this study would be to create experimental colonies with known composition of MDH allotypes that allow for studying the foraging and life history dynamics at the colony level. Additionally, understanding how colonies with different levels of heterogeneity compete against one another at a landscape level in different types of resource

environments would be informative regarding the mechanisms that maintain the variation in MR both within and across populations in different environments.

REFERENCES

Arathi HS and Spivak M. 2001. Influence of colony genotypic composition on the performance of hygienic behaviour in the honeybee*, Apis mellifera L*. Animal Behaviour, 62(1), 57-66.

Amdam GV, Rueppell O, Fondrk MK, Page RE, and Nelson CM. 2009. The nurse's load: Early-life exposure to brood-rearing affects behavior and lifespan in honey bees (*Apis mellifera*). Experimental Gerontology*,* 44, 467-471.

Arnqvist G, Stojković B, Rönn JL, and Immonen E. 2017. The pace-of-life: A sex-specific link between metabolic rate and life history in bean beetles. Functional Ecology, 31, 2299-2309.

Bengston SE, Shin M, and Dornhaus A. 2017. Life-history strategy and behavioural type: risk-tolerance reflects growth rate and energy allocation in ant colonies. Oikos, 126, 556-564.

Biro PA and Stamps JA. 2008. Are animal personality traits linked to life-history productivity? Trends in Ecology & Evolution, 23, 361-368.

Biro PA and Stamps JA. 2010. Do consistent individual differences in metabolic rate promote consistent individual differences in behaviour? Trends in Ecology & Evolution, 25, 653-659.

Blight O, Albet DG, Cerdá X, and Boulay R. 2016. A proactive–reactive syndrome affects group success in an ant species. Behavioral Ecology, 27, 118-125.

Brown JH, Gillooly JF, Allen AP, Savage VM, and West GB. 2004. Toward a metabolic theory of ecology. Ecology, 85, 1771-1789.

Burton T, Killen SS, Armstrong JD, and Metcalfe NB. 2011. What causes intraspecific variation in resting metabolic rate and what are its ecological consequences? Proceedings of the Royal Society B: Biological Sciences, 278, 3465-3473.

Cadotte MW. 2013. Experimental evidence that evolutionarily diverse assemblages result in higher productivity. Proceedings of the National Academy of Sciences, 110, 8996 – 9000.

Careau V, Bininda‐Emonds ORP, Thomas DW, Réale D, and Humphries MM. 2009. Exploration strategies map along fast–slow metabolic and life‐history continua in muroid rodents. Functional Ecology, 23(1), 150-156.

Careau V, Garland Jr. T. 2012. Performance, personality, and energetics: correlation, causation, and mechanism. Physiological Biochemistry and Zoology, 85, 543-571.

Careau V, Thomas D, Humphries MM, Réale D. 2008. Energy metabolism and animal personality. Oikos, 117(5), 641-653.

Coelho JR and Mitton JB. 1988. Oxygen consumption during hovering is associated with genetic variation of enzymes in honey-bees. Functional Ecology, 2, 141-146.

Del Lama MA, Souza RO, Durán XAA, and Soares AEE. 2004. Clinal variation and selection on MDH allozymes in honeybees in Chile. Hereditas. 140, 149-153.

Dammhahn M, Dingemanse NJ, Niemelä PT, and Réale D. 2018. Pace-of-life syndromes: a framework for the adaptive integration of behaviour, physiology and life history. Behavioral Ecology and Sociobiology, 72, 62.

Debecker S, Stoks R. 2019. Pace of life syndrome under warming and pollution: integrating life history, behavior, and physiology across latitudes. Ecological Monographs, 89, 1332.

Dukas R. 2008. Mortality rates of honey bees in the wild. Insectes Sociaux, 55, 252-255.

Feuerbacher E, Fewell JH, Roberts SP, Smith EF, and Harrison JF. 2003. Effects of load type (pollen or nectar) and load mass on hovering metabolic rate and mechanical power output in the honey bee *Apis mellifera*. Journal of Experimental Biology, 206, 1855-1865.

Fewell JH and Page Jr RE. 2000. Colony-level selection effects on individual and colony foraging task performance in honeybees, *Apis mellifera L*. Behavioral Ecology and Sociobiology,

48(3), 173-181.

Fewell JH and Harrison JF. 2016. Scaling of work and energy use in social insect colonies. Behavioral Ecology and Sociobiology, 70(7), 1047-1061.

Garrison LK, Kleineidam CJ, and Weidenmüller A. 2018. Behavioral flexibility promotes collective consistency in a social insect. Scientific Reports, 8, 15836.

Gillooly JF, Brown JH, West GB, Savage VM, and Charnov EL. 2001. Effects of size and temperature on metabolic rate. Science, 293, 2248-2251.

Goulet CT, Thompson MB, Michelangeli M, Wong BBM, Chapple DG. 2017. Thermal physiology: A new dimension of the pace-of-life syndrome. Journal of Animal Ecology, 86, 1269-1280

Glazier DS. 2015. Is metabolic rate a universal 'pacemaker' for biological processes? Biological Reviews, 90, 377-407.

Harrison JF and Fewell JH. 2002. Environmental and genetic influences on flight metabolic rate in the honey bee, *Apis mellifera*. Comparative Biochemistry & Physiology A, 133, 323-333.

Hatty S and Oldroyd B. 1999. Evidence for temperature-dependent selection for malate dehydrogenase allele frequencies in honeybee populations. Journal of Heredity, 90, 565-568.

Jablonszky M, Szász E, Krenhardt K, Markó G, Hegyi G, Herényi M, Laczi M, Nagy G, Rosivall B, Szöllósi E, Török J, and Garamszegi LZ. 2018. Unravelling the relationships between life history, behaviour and condition under the pace-of-life syndromes hypothesis using long-term data from a wild bird. Behavioral Ecology and Sociobiology, 72, 52.

Jacques-Hamilton R, Hall ML, Buttemer WA, Matson KD, da Silva AG, Mulder RA, and Peters A. 2017. Personality and innate immune defenses in a wild bird: Evidence for the pace-of-life hypothesis. Hormones and Behavior, 88, 31-40.

Jeanson R, Weidenmüller A. 2014. Interindividual variability in social insects–proximate causes and ultimate consequences. Biological Reviews, 89, 671-687.

Jolles JW, King AJ, and Killen SS. 2020. The role of individual heterogeneity in collective animal behaviour. Trends in Ecology & Evolution, 35(3), 278-291.

Jones JC, Myerscough MR, Graham S, and Oldroyd BP. 2004. Honey bee nest thermoregulation: diversity promotes stability. Science, 305(5682), 402-404.

Katz K Naug D. 2015. Energetic state regulates the exploration–exploitation trade-off in honeybees. Behavioral Ecology, 26, 1045-1050.

Katz K and Naug D. 2020. A mechanistic model of how metabolic rate can interact with resource environment to influence foraging success and lifespan. Ecological Modelling, 416, 108899.

Kennedy P, Baron G, Qiu B, Freitak D, Helanterä H, Hunt ER, Manfredini F, O'Shea-Wheller T, Patalano S, Pull CD, and Sasaki T. 2017. Deconstructing superorganisms and societies to address big questions in biology. Trends in Ecology & Evolution, 32(11), 861-872.

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

Konarzewski M and Książek A. 2013. Determinants of intra-specific variation in basal metabolic rate. Journal of Comparative Physiology B, 183, 27-41.

Krams I, Niemelä PT, Trakimas G, Krams R, Burghardt GM, Krama T, Kuusik A, Mänd M, Rantala MJ, Mänd R, Kekäläinen J, Sirkka I, Luoto S, and Kortet R. 2017. Metabolic rate

associated with, but does not generate covariation between, behaviours in western stutter-trilling crickets, *Gryllus integer*. Proceedings of the Royal Society B, 284, 1851.

Loreau M and Hector A. 2001. Partitioning selection and complementarity in biodiversity experiments. Nature, 412(6842), 72-76.

Mathot KJ and Frankenhuis WE. 2018. Models of pace-of-life syndromes (POLS): a systematic review. Behavioral Ecology and Sociobiology, 72, 41.

Mattila HR and Seeley TD. 2007. Genetic diversity in honey bee colonies enhances productivity and fitness. Science, 317(5836), 362-364.

Mayack C and Naug D. 2011. A changing but not an absolute energy budget dictates risk-sensitive behaviour in the honeybee. Animal Behaviour, 82, 595-600.

Masoro EJ. 2005. Overview of caloric restriction and ageing. Mechanisms of Ageing and Development, 126, 913-922.

Mugel SG and Naug D. Metabolic rate shapes phenotypic covariance among physiological, behavioral, and life-history traits in honeybees. *In review.*

Neumann P and Moritz RFA. 2000. Testing genetic variance hypotheses for the evolution of polyandry in the honeybee (*Apis mellifera L.*). Insectes Sociaux, 47, 271-279.

Nielsen D, Page RE Jr., Crosland MWJ. 1994. Clinal variation and selection of MDH allozymes in honey bee populations. Experientia, 50.

Niemelä PT, Dingemanse NJ. 2018. On the usage of single measurements in behavioural ecology research on individual differences. Animal Behaviour, 145, 99-105.

Nilsson JA. 2002. Metabolic consequences of hard work. Proceedings of the Royal Society B, 269, 1735 – 1739.

Oldroyd BP and Fewell JH. 2007. Genetic diversity promotes homeostasis in insect colonies.

Trends in Ecology & Evolution, 22(8), 408-413.

Oster GF and Wilson EO. 1978. Caste and ecology in the social insects. Princeton University Press, Princeton.

Page RE, Robinson GE, Fondrk MK, and Nasr ME. 1995. Effects of worker genotypic diversity on honey bee colony development and behavior (Apis mellifera L.). Behavioral Ecology and Sociobiology, 36, 387-396.

Page RE, Erber J, and Fondrk MK. 1998. The effect of genotype on response thresholds to sucrose and foraging behavior of honey bees (*Apis mellifera L.*). Journal of Comparative Physiology A, 182, 489–500.

Page RE, Scheiner R, Erber J, Amdam GV, and Gerald PS. 2006. The development and evolution of division of labor and foraging specialization in a social insect (*Apis mellifera L.*). Current Topics in Developmental Biology, 74, 253-286.

Pankiw T and Page RE. 2000. Response thresholds to sucrose predicts foraging division of labor in honeybees. Behavioral Ecology and Sociobiology, 47, 265-267.

Petchey OL and Gaston KJ. 2002. Functional diversity (FD), species richness and community composition. Ecology Letters, 5(3), 402-411.

Pettersen AK, White CR, and Marshall DJ. 2016. Metabolic rate covaries with fitness and the pace of the life history in the field. Proceedings of the Royal Society B, 283(1831), 20160323.

Réale D, Garant D, Humphries MM, Bergeron P, Careau V, and Montiglio PO. 2010. Personality and the emergence of the pace-of-life syndrome concept at the population level. Philosophical Transactions of the Royal Society B, 365, 4051-4063.

Ricklefs RE and Wikelski M. 2002. The physiology/life-history nexus. Trends in Ecology & Evolution, 17, 462-468.

Royauté R, Berdal MA, Garrison CR, and Dochtermann NA. 2018. Paceless life? A meta-analysis of the pace-of-life syndrome hypothesis. Behavioral Ecology and Sociobiology, 72, 64.

Robinson GE. 1992. Regulation of division of labor in insect societies. Annual Review of Entomology, 37, 637-665

Rosseel Y. 2012. lavaan: An R Package for Structural Equation Modeling. Journal of Statistical Software, 48, 2.

Rueppell O, Bachelier C, Fondrk MK, and Page RE. 2007. Regulation of life history determines lifespan of worker honey bees (*Apis mellifera* L.). Experimental Gerontology, 42, 1020-1032.

Santostefano F, Wilson AJ, Niemelä PT, and Dingemanse NJ. 2017. Behavioural mediators of genetic life-history trade-offs: a test of the pace-of-life syndrome hypothesis in field crickets. Proceedings of the Royal Society B, 284, 20171567.

Scheiner R, Plückhahn S, Öney B, Blenau W, and Erber J. 2002. Behavioural pharmacology of octopamine, tyramine and dopamine in honey bees. Behavioural Brain Research, 136, 545-553

Schmid-Hempel P, Kacelnik A, and Houston AI. 1985. Honeybees maximize efficiency by not filling their crop. Behavioral Ecology and Sociobiology, 17(1), 61-66.

Seeley TD. 1997. Honey bee colonies are group-level adaptive units. The American Naturalist, 150, 22-S41.

Segev U, Burkert L, Feldmeyer B, and Foitzik S. 2017. Pace-of-life in a social insect: behavioral syndromes in ants shift along a climatic gradient. Behavioral Ecology, 28, 1149-1159.

Sinclair DA. 2005. Toward a unified theory of caloric restriction and longevity regulation. Mechanisms of Ageing and Development, 126, 987-1002.

Sohal RS and Weindruch R. 1996. Oxidative stress, caloric restriction, and aging. Science, 273, 59-63.

Sol D, Maspons J, Gonzalez-Voyer A, Morales-Castilla I, Garamszegi LZ, and Møller AP. 2018. Risk-taking behavior, urbanization and the pace of life in birds. Behavioral Ecology and Sociobiology, 72(3), 59.

Sullivan JP, Jassim O, Fahrbach SE, and Robinson GE. 2000. Juvenile hormone paces behavioral development in the adult worker honey bee. Hormones and Behavior, 37, 1-14.

Tait C and Naug D. 2020. Cognitive phenotypes and their functional differences in the honeybee, *Apis mellifera*. Animal Behaviour. *In press*.

Takahashi Y, Tanaka R, Yamamoto D, Noriyuki S, and Kawata M. 2018. Balanced genetic diversity improves population fitness. Proceedings of the Royal Society B, 285, 1871.

Tarpy DR. 2003. Genetic diversity within honeybee colonies prevents severe infections and promotes colony growth. Proceedings of the Royal Society B, 270, 99-103.

Wagener-Hulme C, Kuehn JC, Schulz DJ, and Robinson GE. 1999. Biogenic amines and division of labor in honey bee colonies. Journal of Comparative Physiology A–Sensory Neural and Behavioral Physiology, 184, 471-479.

Waters JS, Holbrook CT, Fewell JH, and Harrison JF. 2010. Allometric scaling of metabolism, growth, and activity in whole colonies of the seed-harvester ant *Pogonomyrmex californicus*. American Naturalist, 176, 501-510.

West GB, Woodruff WH, and Brown JH. 2002. Allometric scaling of metabolic rate from molecules and mitochondria to cells and mammals. Proceedings of the National Academy of Sciences, 99, 2473-2478.

White CR, Marshall DJ, Alton LA, Arnold PA, Beaman JE, Bywater CL, Condon C, Crispin TS, Janetzki A, Pirtle E, and Winwood-Smith HS. 2019. The origin and maintenance of metabolic allometry in animals. Nature Ecology & Evolution, 3(4), 598-603.

Wolf M, Van Doorn GS, Leimar O, and Weissing FJ. 2007. Life-history trade-offs favour the evolution of animal personalities. Nature, 447(7144), 581.

Woods HA. 2014. Mosaic physiology from developmental noise: within-organism physiological diversity as an alternative to phenotypic plasticity and phenotypic flexibility. The Journal of Experimental Biology, 217, 35-45.

Wray MK, Mattila HR, and Seeley TD. 2011. Collective personalities in honeybee colonies are linked to colony fitness. Animal Behaviour, 81, 559-568.

APPENDIX 1: SUPPLEMENTARY MATERIAL FOR CHAPTER 1

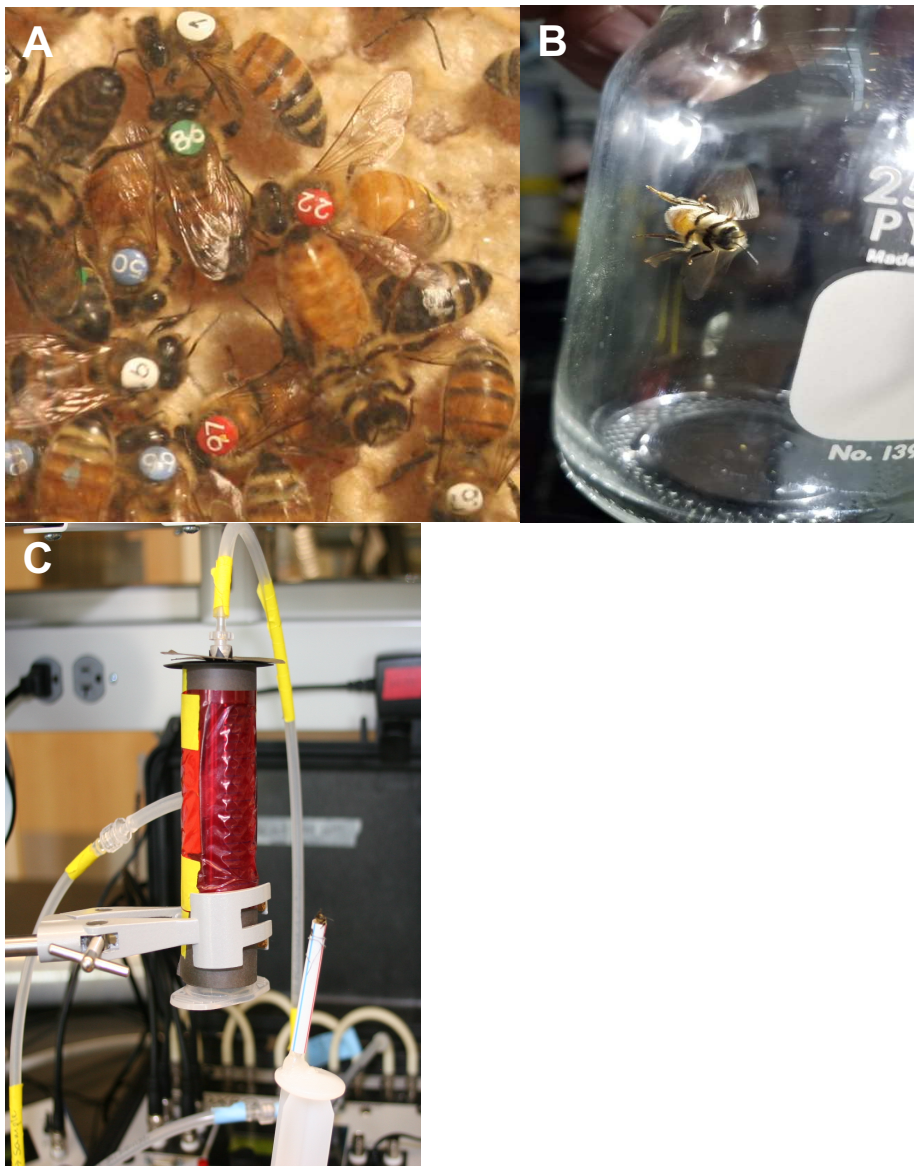

Figure S1.1. Photographs of (A) tagged bees in the common garden observation hive; (B) a bee in the FMR flight chamber; and (C) the RMR set-up with the FoxBox in the background, darkened chamber and strapped bee in the foreground.

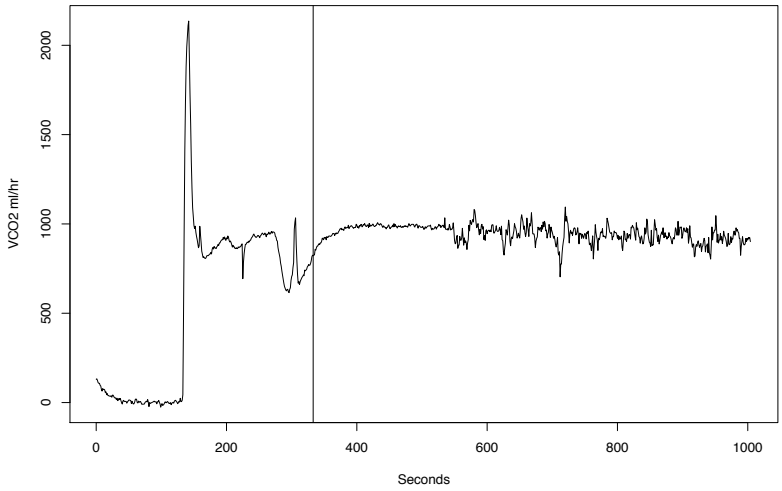

Figure S1.2. Example FMR VCO_2 trace over the entirety of the FMR assay. Vertical bar represents when flight was initiated, which continued for the rest of the assay.

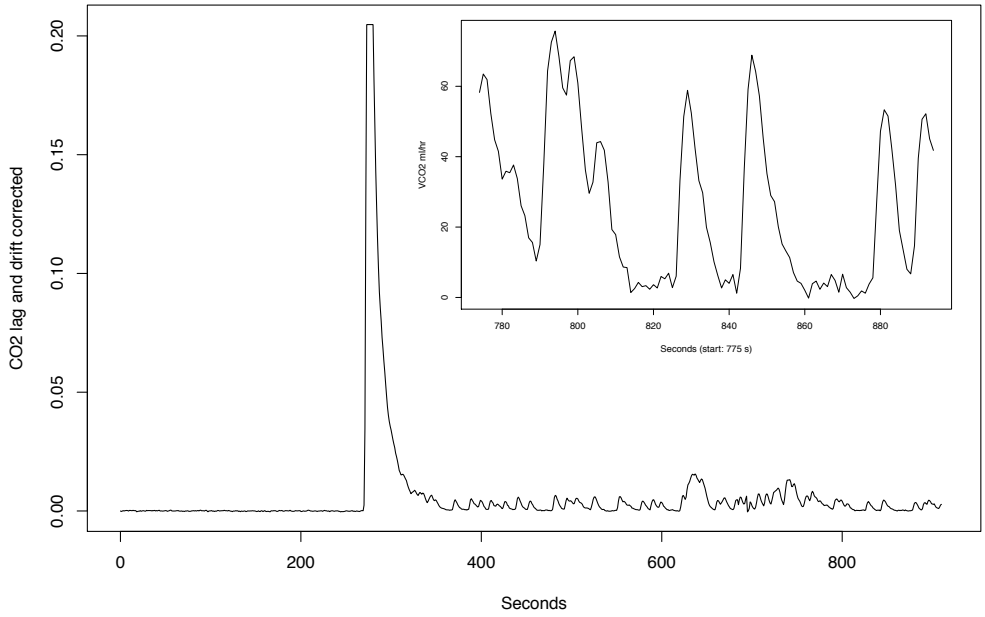

Figure S1.3. Example CO_2 production trace of RMR data. The main figure shows the VCO_2 produced over the entire 10 minute trial including the beginning reading from baseline chamber to correct for drift and zero the readings, and the CO_2 spike when the gas analyzer was switched from reading the baseline chamber to the experimental chamber (discarding the first two minutes following that). The insert panel shows the VCO_2 during the selected two minutes with the lowest variance.

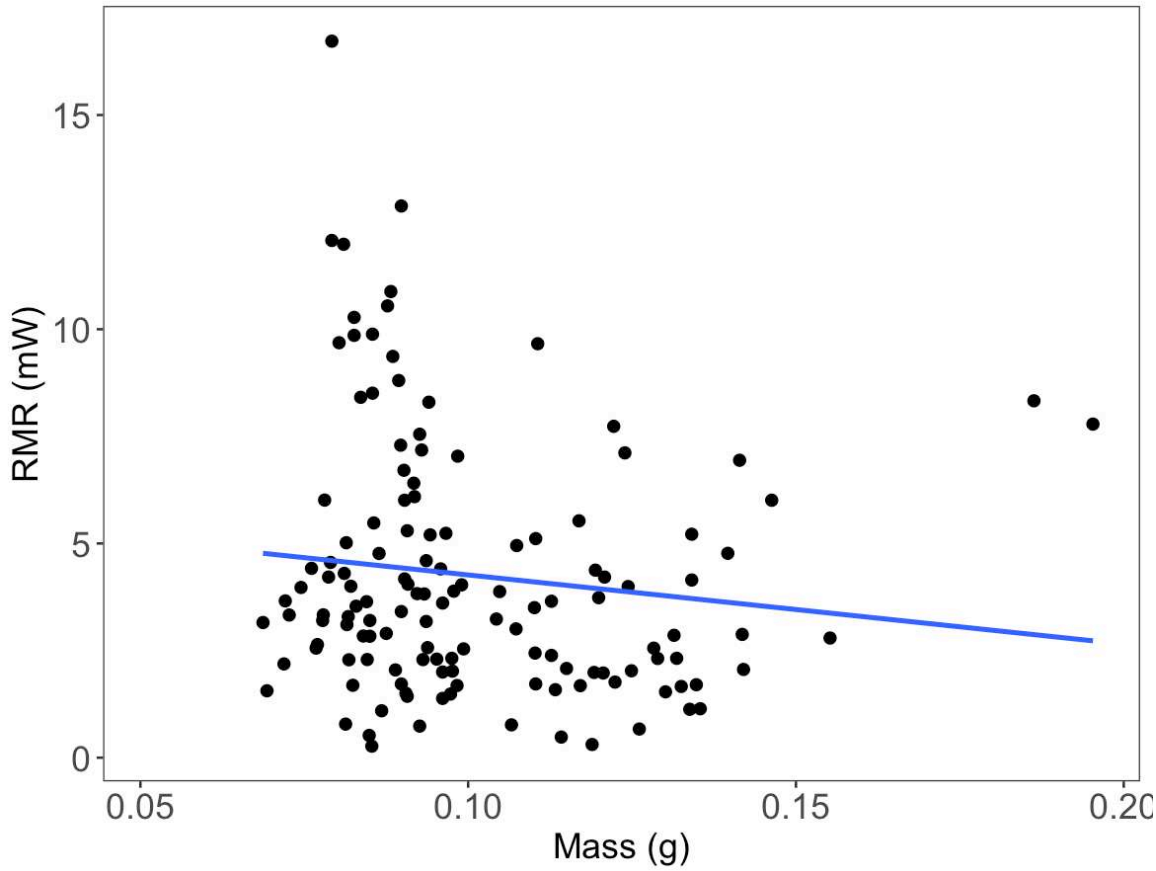

Figure S1.4. RMR as a function of body mass with points representing individual bees, and the regression line representing a linear model on log-transformed data ($F_{1,132}$ = 2.3, p = 0.136, y = -0.48x + 0.096, r^2 = 0.017).

For figures S1.5–S1.11 AIC and ΔAIC scores are presented adjacent to each diagram.

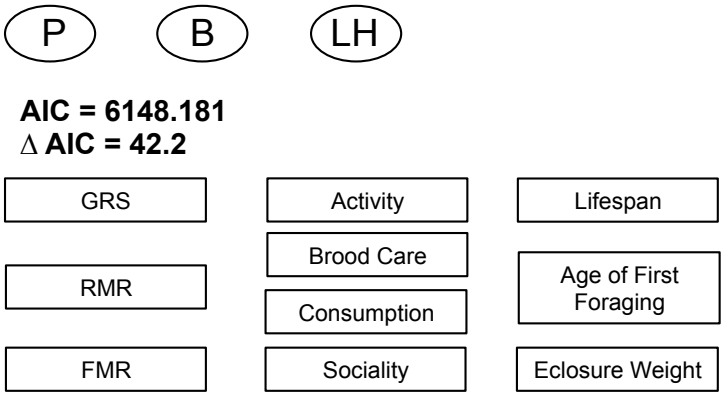

Figure S1.5. Null model wherein all measured variables are unrelated to one another.

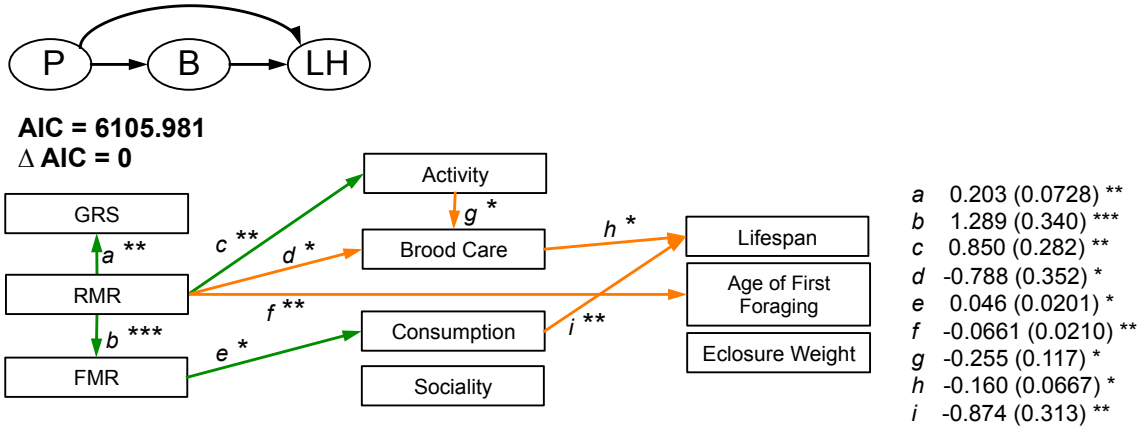

Figure S1.6. Class A, Schema 1: Physiology influences behavior and in turn both influence life-history. This is the overall top model and is presented in the main text (Fig. 2). Solid lines represent significant partial correlation coefficients (positive in green and negative in orange), at the $p < 0.05$ (*), ≤ 0.01 (**), and ≤ 0.001 (***) level.

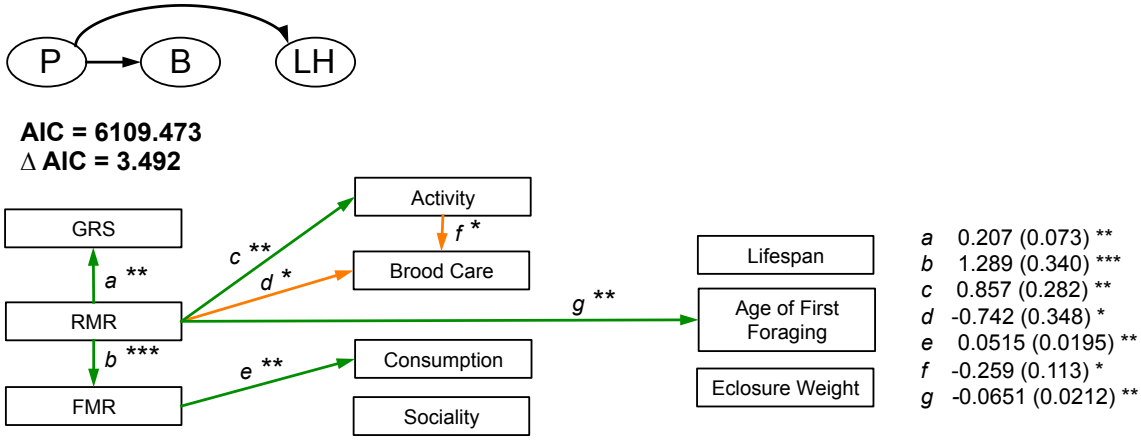

Figure S1.7. Class A, Schema 2: Physiology influences both behavior and life-history, and behavior and life-history are unrelated. Solid lines represent significant partial correlation coefficients (positive in green and negative in orange), at the $p < 0.05$ (*), ≤ 0.01 (**), and ≤ 0.001 (***) level.

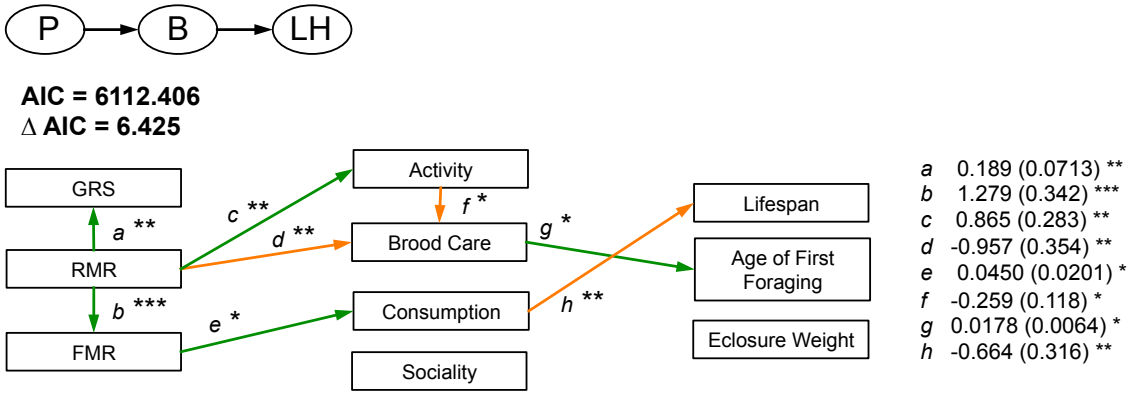

Figure S1.8. Class A, Schema 3: Behavior acts as a mediator between physiology and life-history. Solid lines represent significant partial correlation coefficients (positive in green and negative in orange), at the $p < 0.05$ (*), ≤ 0.01 (**), and ≤ 0.001 (***) level.

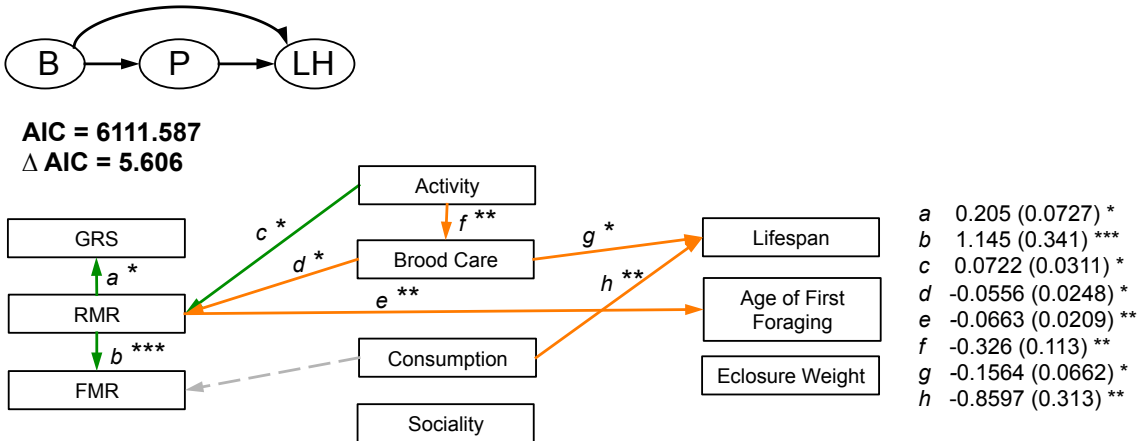

Figure S1.9. Class B, Schema 1: Behavior influences physiology and in turn both influence life-history. Solid lines represent significant partial correlation coefficients (positive in green and negative in orange), at the $p < 0.05$ (*), ≤ 0.01 (**), and ≤ 0.001 (***) level and dashed gray lines represent non-significant correlation coefficients ($p \geq 0.05$).

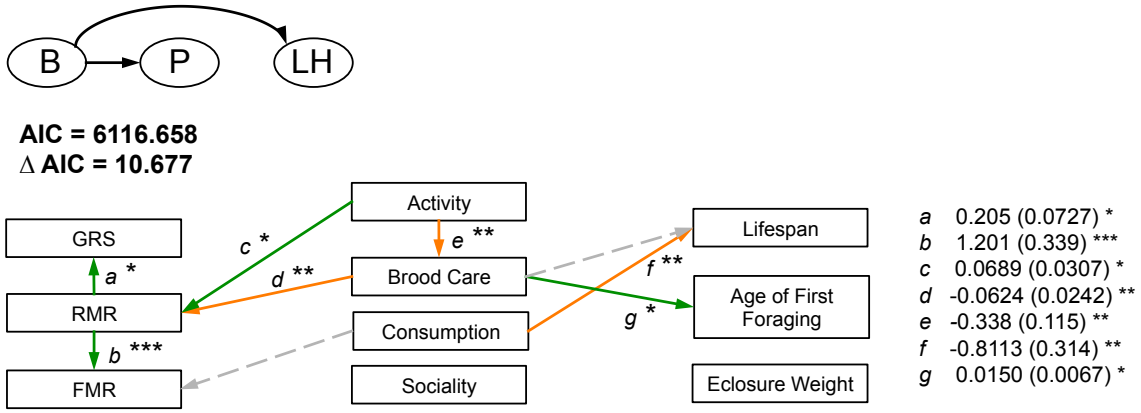

Figure S1.10. Class B, Schema 2: Behavior influences both physiology and life-history, and physiology and life-history are unrelated. Solid lines represent significant partial correlation coefficients (positive in green and negative in orange), at the $p < 0.05$ (*), ≤ 0.01 (**), and ≤ 0.001 (***) level and dashed gray lines represent non-significant correlation coefficients ($p \geq 0.05$).

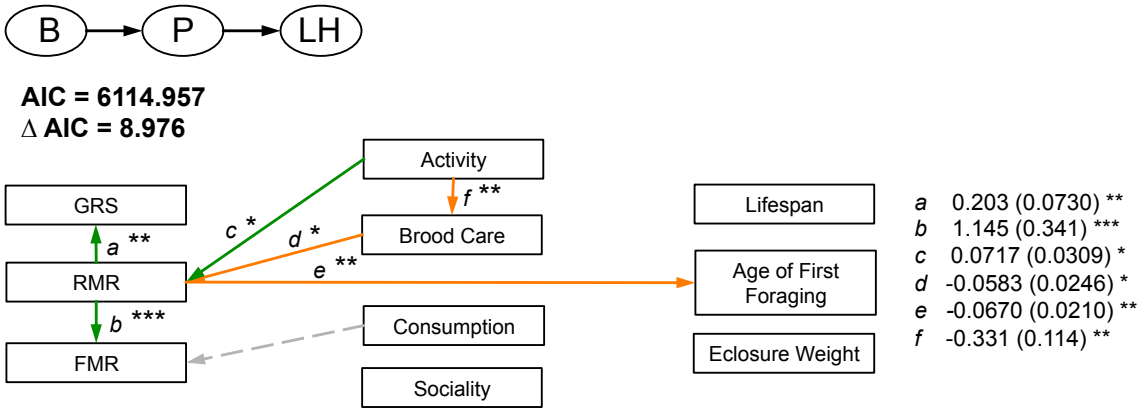

Figure S1.11. Class B, Schema 3: Physiology acts as a mediator between behavior and life-history. Solid lines represent significant partial correlation coefficients (positive in green and negative in orange), at the $p < 0.05$ (*), ≤ 0.01 (**), and ≤ 0.001 (***) level and dashed gray lines represent non-significant correlation coefficients ($p \geq 0.05$).

Table S1.1. Linear and generalized linear model results of significant phenotypic relationships from the best-supported path model (Fig. 1.1), with colony included as a covariate. Regression coefficients are shown for significant predictors. Transformed values were used in these analyses, and because each of these variables is tested independently of the rest, the exact values of the coefficients are less meaningful and therefore should be interpreted with caution. Source colony was used as a covariate in all these models and was found to have no significant effect on any of the relationships. All relationships tested from the top supported path model with the exception of FMR and consumption were found to have the same approximate significance and strength as in the integrative approach of the path SEMs, which upholds the strength and validity of the path SEMs as an integrative approach to understanding how multivariate relationships are related to one another in a causal manner.

Response	Predictor	*F*	*df*	*p*	*β*
FMR	**RMR**	**11.6**	**1**	**< 0.001 *****	**1.32**
	Colony	0.202	3	0.894	
GRS	**RMR**	**4.21**	**1**	**0.0433 ***	**0.156**
	Colony		3	0.216	
Activity	**RMR**	**4.60**	**1**	**0.0347 ***	**0.637**
	Colony	0.426	3	0.735	
Brood care	**RMR**	**11.2**	**1**	**0.00121 ****	**-1.22**
	Colony	1.90	3	0.135	
AFF	**RMR**	**5.47 †**	**1**	**0.0194 ***	**-0.0530**
	Colony	5.85 †	3	0.119	
Consumption	FMR	3.55	1	0.0633 *n.s.*	
	Colony	1.07	3	0.367	
Brood Care	**Activity**	**9.03**	**1**	**0.00334 ****	**-0.354**
	Colony	0.715	3	0.545	
Lifespan	**Brood Care**	**4.27**	**1**	**0.0429 ***	**-0.145**
	Colony	2.10	3	0.110	
Lifespan	**Consumption**	**5.77**	**1**	**0.0189 ***	**-0.703**
	Colony	1.41	3	0.248	

† Chi-square statistic from generalized linear model rather than *F* statistic

APPENDIX 2: SUPPLEMENTARY MATERIAL FOR CHAPTER 2

Methods

Experimental design

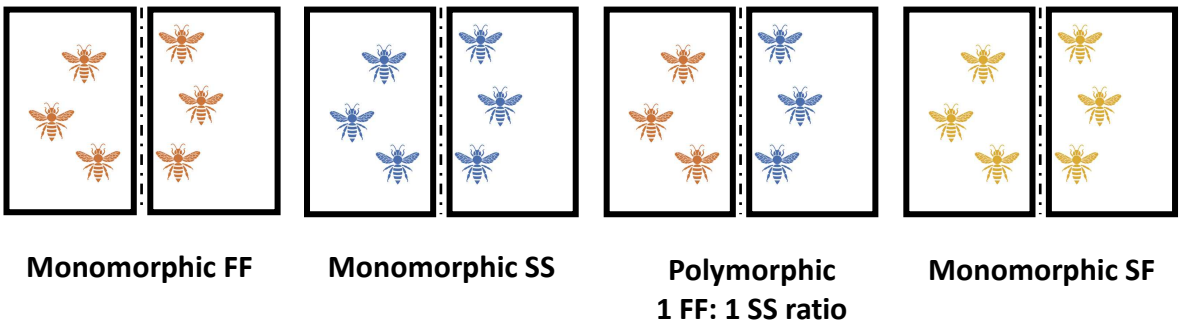

Figure S2.1. Schematic of housing and experimental group composition treatments. This novel experimental cage design allowed for separate measurements to be taken as each subgroup (each set of 20 bees on either side of the wire mesh) could be separated for a brief time sufficient to characterize each morph type's contribution to total group performance and still maintain a functionally unified social group of 40 bees. FF bees (red) and SS bees (blue) could be separated on each side of the wire mesh (dotted line) in a polymorphic group yet exchange pheromones and food and contact one another across this mesh divide. FF, SS, and SF (purple) groups were also separated by the wire mesh to maintain uniformity in experimental design across composition treatments.

Analysis

Median survival analysis using a linear model framework was followed by fitting a Cox proportional hazard regression in which groups in each treatment combination were pooled into a single cohort for a total of eight cohorts (4 composition types x 2 resource quality levels), using composition type and resource environment as fixed effects. This analysis pools the bees of replicate groups and treats the individual bee as the unit of analysis for time to death, and thus comparisons are intended only to highlight proportional hazards between the eight treatment categories and should be interpreted with caution as variation within treatments is not accounted

for as it is in the median survival analysis. A Kaplan-Meier plot was also generated to visualize the survival differences between these pooled cohorts of each treatment combination (Fig. S2.3).

Results

Metabolic Rate

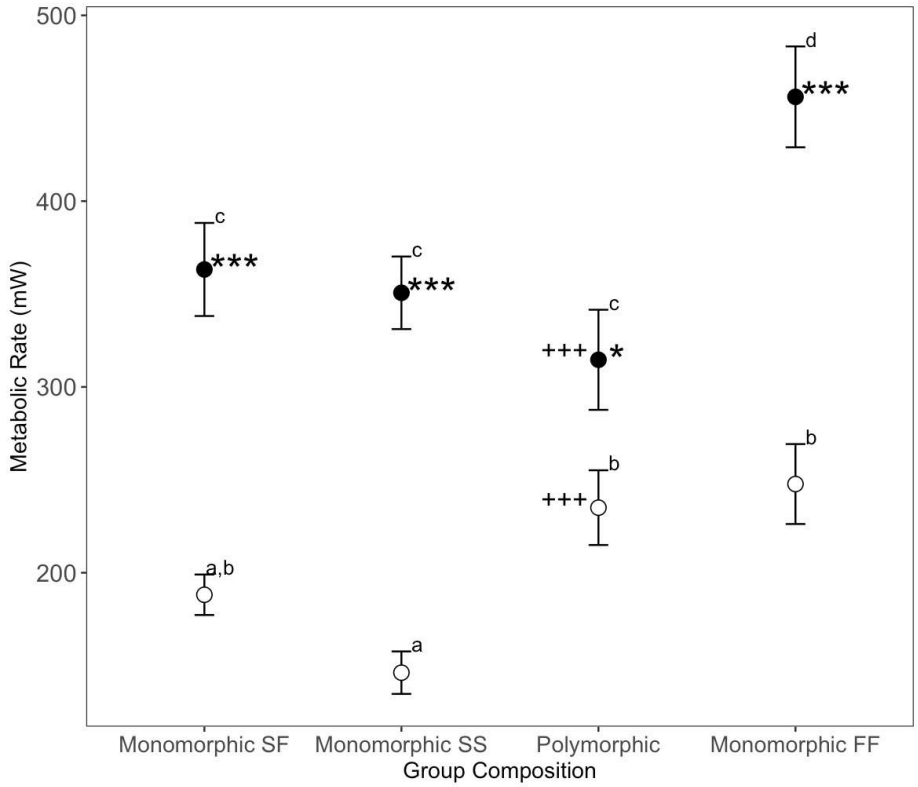

Figure S2.2. Group mass-independent MR of experimental groups (mW) is affected by an interaction between group composition and resource environment (interaction: $F_{3,\,217} = 3.76$, $p = 0.01$; main effects of resource environment: $F_{1,217} = 48.57$, $p < 0.001$; group composition: $F_{3,217} = 7.94$, $p < 0.001$). Points represent mean ± SE in high (filled) and low (open) resource environments. Asterisks denote a significant difference across resource environments for a group composition (* $0.01 < p < 0.05$; ** $0.001 < p < 0.01$; *** $p < 0.001$). Different letters represent significant differences between different group compositions within a resource environments (*a,b* for low, *c,d* for high). Crosses (+) denote a significant deviation in performance of the polymorphic group from null expectation based on a one-sample *t*-test.

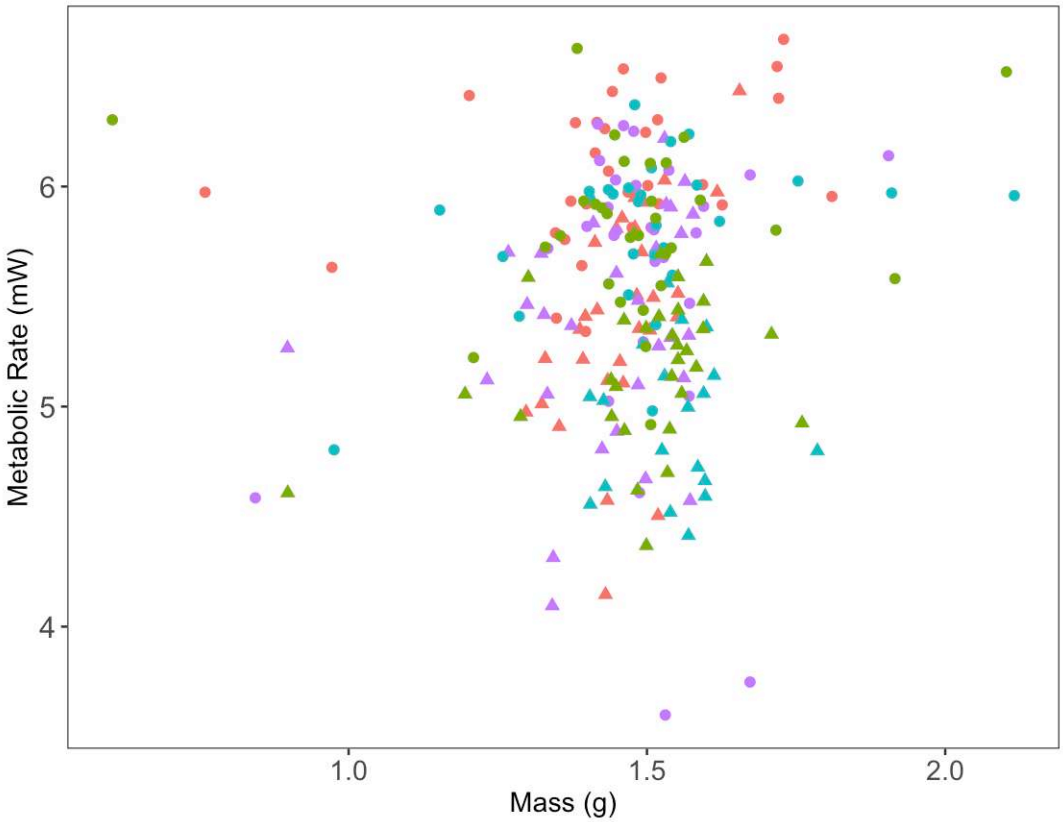

Figure S2.3. Log-log plot of group MR (mW) as a function of group mass. Points represent group values of monomorphic FF (red), monomorphic SS (blue), monomorphic SF (green), and polymorphic (purple) in both high (circles) and low (triangles) resource environments. Mass was not a significant predictor of MR ($F_{1, 223} = 2.85$, $p = 0.09$), thus mass-specific MR was used for further analyses.

Table S2.1. Post-hoc Tukey-adjusted *t*-test comparisons from the linear model on mass-specific Metabolic Rate (mW/g) with predictors of an interaction between composition type and resource level as well as main effects of each factor ('emmeans' R package). Compact letter display (CLD) represent statistically similar performance measures within low (*a-b)*, and high resource environments (*c-d*), which are depicted on Fig. 2.1. Bolded comparisons are significant at the 0.05 level.

Factor Level		Estimate (95% CI)	CLD	Factor Level Comparison	*t*	*df*	*p*
Low Resource Environment	**Monomorphic FF**	112.7 (89.9, 135.5)	*b*	**Monomorphic SS**	**2.69**	**217**	**0.038**
				Monomorphic SF	1.81	217	0.27
				Polymorphic	0.11	217	0.99
				Monomorphic FF (High Resource)	**7.04**	**217**	**<0.001**
	Monomorphic SS	62.8	*a*	Monomorphic SF	1.12	217	0.68

		(34.2, 91.4)					
				Polymorphic	**2.77**	**217**	**0.031**
				Monomorphic SS (High Resource)	**5.04**	**217**	**<0.001**
	Monomorphic SF	83.4 (61.0,105.9)	*a ,b*				
				Polymorphic	1.90	217	0.232
				Monomorphic SF (High Resource)	**5.44**	**217**	**<0.001**
	Polymorphic	114.5 (91.3, 137.7)	*b*				
				Polymorphic (High Resource)	1.49	217	0.138
High Resource Environment	**Monomorphic FF**	227.8 (205.0, 250.6)	*d*				
				Monomorphic SS	**4.21**	**217**	**<0.001**
				Monomorphic SF	**3.43**	**217**	**0.004**
				Polymorphic	**5.31**	**217**	**<0.001**
	Monomorphic SS	157.7 (134.1, 181.3)	*c*				
				Monomorphic SF	0.84	217	0.835
				Polymorphic	1.08	217	0.703
	Monomorphic SF	171.7 (148.9, 194.5)	*c*				
				Polymorphic	1.94	217	0.216
	Polymorphic	139.4 (115.8, 163.0)	*c*				

Energetic Intake

Table S2.2. Post-hoc Tukey-adjusted *t*-test comparisons from the linear model on energetic intake (daily consumption in mW/g) linear model with predictors of an interaction between composition type and resource level as well as main effects of each factor ('emmeans' R package). Compact letter display (CLD) represent statistically similar performance measures within low (*a)*, and high resource environments (*b*), which are depicted on Fig. 2.2. Bolded comparisons are significant at the 0.05 level.

Factor Level		**Estimate (95% CI)**	**CLD**	**Factor Level Comparison**	***t***	***df***	***p***
Low Resource Environment	**Monomorphic FF**	13.1 (10.1, 16.2)	*a*				
				Monomorphic SS	**1.12**	**218**	**0.68**
				Monomorphic SF	1.17	218	0.64
				Polymorphic	1.59	218	0.39
				Monomorphic FF (High Resource)	**4.24**	**218**	**<0.001**
	Monomorphic SS	15.9 (12.1, 19.7)	*a*				
				Monomorphic SF	0.10	218	0.99

				Polymorphic	0.29	218	0. 99
				Monomorphic SS (High Resource)	**3.94**	**218**	**<0.001**
	Monomorphic SF	15.7 (12.7, 18.6)	*a*				
				Polymorphic	0.44	218	0.97
				Monomorphic SF (High Resource)	**4.49**	**218**	**<0.001**
	Polymorphic	16.6 (13.5, 19.7)	*a*	**Polymorphic (High Resource)**	**2.37**	**218**	**0.019**
High Resource Environment	**Monomorphic FF**	22.4 (19.3, 25.4)	*b*				
				Monomorphic SS	1.52	218	0.43
				Monomorphic SF	1.37	218	0.52
				Polymorphic	0.21	218	0.99
	Monomorphic SS	25.7 (22.6, 28.8)	*b*				
				Monomorphic SF	0.16	218	0.99
				Polymorphic	1.70	218	0.33
	Monomorphic SF	25.4 (22.3, 28.4)	*b*				
				Polymorphic	1.56	218	0.41
	Polymorphic	21.9 (18.8, 25.0)	*b*				

Thermoregulation

Table S2.3. Post-hoc Tukey-adjusted *t*-test comparisons from the linear model on thermoregulation in cold stress (mean residual temperature in ˚C) linear model with predictors of an interaction between composition type and resource level as well as main effects of each factor ('emmeans' R package). Compact letter display (CLD) represent statistically similar performance measures within low (*a-b)*, and high resource environments (*c*), which are depicted on Fig. 2.3. Bolded comparisons are significant at the 0.05 level.

Factor Level		**Estimate (95% CI)**	**CLD**	**Factor Level Comparison**	***t***	***df***	***p***
Low Resource Environment	**Monomorphic FF**	0.950 (0.741, 1.158)	*b*				
				Monomorphic SS	**2.83**	**213**	**0.026**
				Monomorphic SF	1.14	213	0.66
				Polymorphic	**2.73**	**213**	**0.034**
				Monomorphic FF (High Resource)	**4.63**	**213**	**<0.001**
	Monomorphic SS	0.453 (0.18, 1.16)	*a,*				
				Monomorphic SF	1.87	213	0.24
				Polymorphic	0.477	213	0.96
				Monomorphic SS (High Resource	**7.46**	**213**	**<0.001**
	Monomorphic SF	0.781 (0.58, 0.99)	*a,b*				
				Polymorphic	1.63	213	0.367
				Monomorphic SF (High Resource)	**6.38**	**213**	**<0.001**

	Polymorphic	0.538 (0.33, 0.75)	*a*	**Polymorphic (High Resource)**	**6.83**	**213**	**<0.001**
High Resource Environment	**Monomorphic FF**	1.65 (1.44, 1.86)	*c*	Monomorphic SS	0.93	213	0.79
				Monomorphic SF	0.522	213	0.95
				Polymorphic	0.38	213	0.98
	Monomorphic SS	1.79 (1.57, 2.01)	*c*	Monomorphic SF	0.42	213	0.97
				Polymorphic	1.31	213	0.56
	Monomorphic SF	1.73 (1.52, 1.94)	*c*	Polymorphic	0.92	213	0.80
	Polymorphic	1.59 (1.37, 1.80)	*c*				

Table S2.4. Post-hoc Tukey-adjusted *t*-test comparisons from the linear model on thermoregulation in heat stress (mean residual temperature in ˚C, multiplied by -1 such that better thermoregulatory performance corresponds to higher values) linear model with predictors of an interaction between composition type and resource level as well as main effects of each factor ('emmeans' R package). Compact letter display (CLD) represent statistically similar performance measures within low (*a)*, and high resource environments (*b*), which are depicted on Fig. 2.4. Bolded comparisons are significant at the 0.05 level.

Factor Level		**Estimate (95% CI)**	**CLD**	**Factor Level Comparison**	***t***	***df***	***p***
Low Resource Environment	**Monomorphic FF**	0.22 (0.13, 0.30)	*a*	Monomorphic SS	0.51	212	0.96
				Monomorphic SF	1.22	212	0.61
				Polymorphic	0.67	212	0.91
				Monomorphic FF (High Resource)	0.84	212	0.40
	Monomorphic SS	0.25 (0.14, 0.36)	*a*	Monomorphic SF	1.57	212	0.40
				Polymorphic	1.10	212	0.70
				Monomorphic SS (High Resource)	0.31	212	0.76
	Monomorphic SF	0.14 (0.06, 0.23)	*a*	Polymorphic	0.53	212	0.95
				Monomorphic SF (High Resource)	0.75	212	0.46
	Polymorphic	0.174 (0.09, 0.26)	*a*	Polymorphic (High Resource)	0.69	212	0.49
High Resource Environment	**Monomorphic FF**	0.67 (.018, 0.35)	*b*	Monomorphic SS	0.60	212	0.93
				Monomorphic SF	1.29	212	0.57
				Polymorphic	2.17	212	0.14

Monomorphic SS	0.23 (0.14, 0.32)	*b*				
			Monomorphic SF	0.67	212	0.91
			Polymorphic	1.55	212	0.41
Monomorphic SF	0.18 (0.10, 0.27)	*b*				
			Polymorphic	0.91	212	0.80
Polymorphic	0.13 (0.04, 0.22)	*b*				

Survival

Table S2.5. Post-hoc Tukey-adjusted *t*-test comparisons from the linear model on median survival (days) linear model with predictors of an interaction between composition type and resource level as well as main effects of each factor ('emmeans' R package). Compact letter display (CLD) represent statistically similar performance measures within low (*a-b)*, and high resource environments (*c-d*), which are depicted on Fig. 2.5. Bolded comparisons are significant at the 0.05 level.

Factor Level		**Estimate (95% CI)**	**CLD**	**Factor Level Comparison**	***t***	***df***	***p***
Low Resource Environment	**Monomorphic FF**	18.7 (17.6, 19.8)	*b*	**Monomorphic SS**	**6.99**	**217**	**<0.001**
				Monomorphic SF	**6.14**	**217**	**<0.001**
				Polymorphic	**5.42**	**217**	**<0.001**
				Monomorphic FF (High Resource)	**4.24**	**217**	**<0.001**
	Monomorphic SS	12.7 (11.4, 14.0)	*a*				
				Monomorphic SF	1.63	217	0.36
				Polymorphic	2.15	217	0.14
				Monomorphic SS (High Resource)	**6.34**	**217**	**<0.001**
	Monomorphic SF	14.1 (13.0, 15.1)	*a*				
				Polymorphic	0.62	217	0.93
				Monomorphic SF (High Resource)	**6.21**	**217**	**<0.001**
	Polymorphic	14.6 (13.5, 15.6)	*a*	**Polymorphic (High Resource)**	**5.67**	**217**	**<0.001**
High Resource Environment	**Monomorphic FF**	21.9 (20.9, 23.0)	*d*	**Monomorphic SS**	**4.80**	**217**	**<0.001**
				Monomorphic SF	**4.18**	**217**	**<0.001**
				Polymorphic	**3.83**	**217**	**0.001**
	Monomorphic SS	18.2 (17.1, 19.3)	*c*				
				Monomorphic SF	0.69	217	0.90
				Polymorphic	0.96	217	0.77
	Monomorphic SF	18.8 (17.7, 19.8)	*c*				
				Polymorphic	0.28	217	0.99
	Polymorphic	19.0 (17.9, 20.1)	*c*				

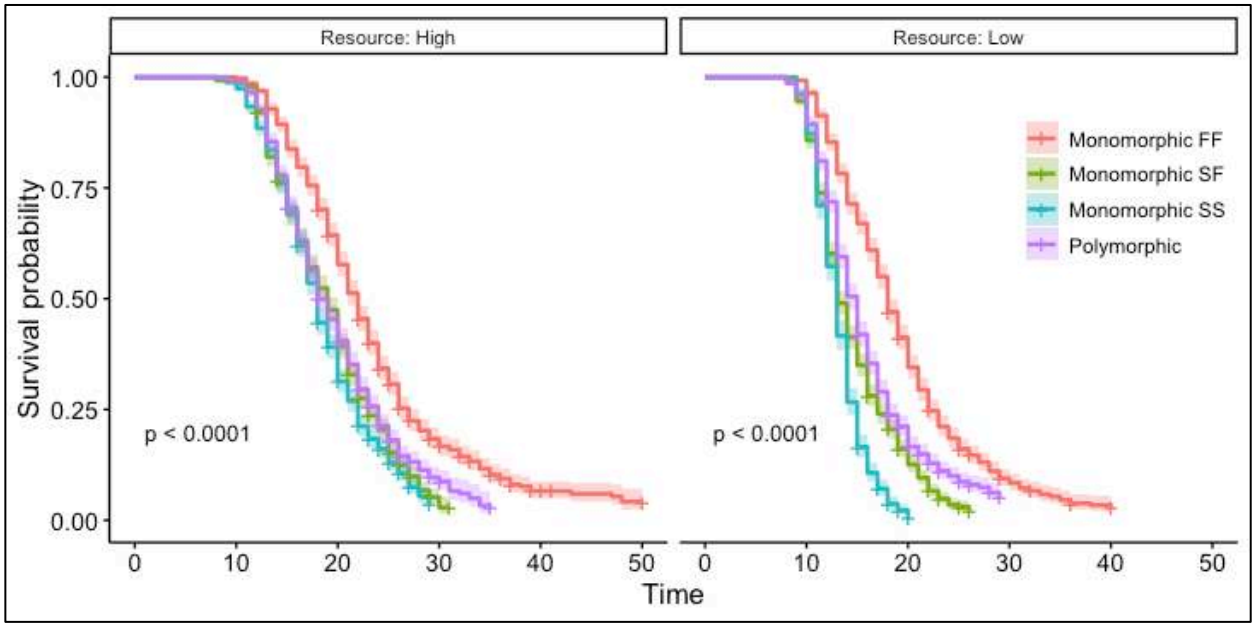

Figure S2.4. Kaplan-Meier plot comparing survival probabilities through time for pooled treatment groups within high (left panel) and low (right panel) resource environments. Cox regression hazard ratios for the same model are depicted in Fig. S2.5.

The Cox hazard ratio model revealed that compared to the high resource environment, the low resource environment had a 2.26 times higher survival risk (Wald $z = 35.04$, $p < 0.001$). Compared to the monomorphic FF group, which had the highest survival, the monomorphic SS group had a 2.60 times greater survival risk (Wald $z = 28.03$, $p < 0.001$), the monomorphic SF group had a 1.99 times higher survival risk (Wald $z = 22.09$, $p < 0.001$), and the polymorphic group had a 1.65 times higher survival risk (Wald $z = 15.69$, $p < 0.001$; Fig. S2.5). A non-significant interaction term was dropped from the model prior to analysis presented here.

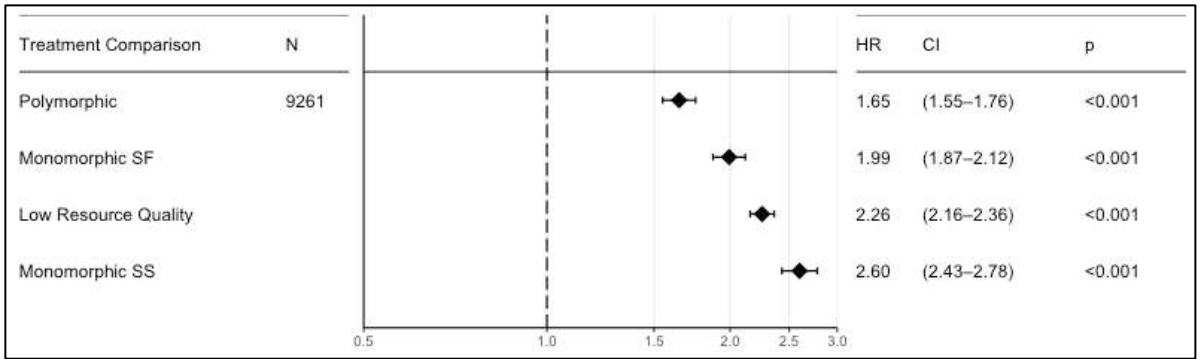

Figure S2.5. Survival hazard ratios ranked lowest to highest for pooled Cox regression analysis. Reference groups were the high resource environment and the monomorphic FF composition for resource level and composition type, respectively.

Table S2.6. All diversity effects and components.

		Diversity Effect					Selection Component				Complementarity Component			
Performance Measure	**Resource Environment**	**Expected Value**	**Observed Value (SE)**	***t***	***df***	***p***	***Observed Value (SE)***	***t***	***df***	***p***	***Observed Value (SE)***	***t***	***df***	***p***
MR (mW/g)	**Low**	**87.75**	**114.5 (11.8)**	**2.68**	**28**	**0.012 ***	-4.17 (5.30)	-0.79	28	0.44	**28.81 (11)**	**2.63**	**28**	**0.014 ***
	High	**192.75**	**139.4 (12)**	**4.28**	**27**	**<0.001 *****	-4.93 (2.62)	-1.89	25	0.07	**-42.29 (11.8)**	**3.58**	**25**	**0.0015 ****
Energetic Intake (mW/g)	Low	14.5	16.6 (1.6)	1.26	28	0.22	-0.020 (0.049)	0.41	28	0.69	0.293 (0.23)	1.28	28	0.21
	High	**24.05**	**21.15 (1.6)**	**2.3**	**28**	**0.03 ***	-0.0515 (0.043)	1.19	26	0.24	-0.462 (0.30)	1.56	26	0.13
Thermo-regulation (Cold stress, ˚C)	**Low**	**0.702**	**0.538 (0.108)**	**2.16**	**28**	**0.04 ***	**-0.177 (0.072)**	**2.45**	**28**	**0.021 ***	-0.239 (0.178)	1.34	28	0.19
	High	1.72	1.59 (0.109)	1.15	28	0.26	**-0.0265 (0.009)**	**-2.95**	**28**	**0.006 ****	-0.337 (0.238)	1.41	28	0.17
Thermo-regulation (Heat stress, ˚C)	Low	-0.233	-0.174 (0.045)	1.48	27	0.15	0.0089 (0.0047)	1.88	27	0.071	0.114 (0.081)	1.42	27	0.17
	High	**-0.248**	**-0.129 (0.046)**	**2.63**	**25**	**0.014 ***	-0.0013 (0.004)	0.33	26	0.75	**0.239 (0.87)**	**2.74**	**26**	**0.011 ***
Median Survival (days)	**Low**	**15.7**	**14.6 (0.55)**	**2.94**	**28**	**0.006 ****	-0.439 (0.238)	1.84	28	0.076	-0.305 (0.91)	0.34	28	0.74
	High	20.05	19 (0.56)	1.83	27	0.078	-0.228 (0.141)	1.61	28	0.12	-1.25 (1.2)	1.04	28	0.31