The Journal of Experimental Biology 215, 1651-1661

doi:10.1242/jeb.066167

RESEARCH ARTICLE

Within-lifetime trade-offs but evolutionary freedom for hormonal and immunological traits: evidence from mice bred for high voluntary exercise

Cynthia J. Downs[1,*], Heidi Schutz[2,†], Thomas H. Meek[2,‡], Elizabeth M. Dlugosz[2], Wendy Acosta[2], Karen S. de Wolski[3], Jessica L. Malisch[3], Jack P. Hayes[1] and Theodore Garland, Jr[2]

[1]Program in Ecology, Evolution, and Conservation Biology and Department of Biology, University of Nevada, Reno, NV 89557, USA, [2]Department of Biology, University of California, Riverside, CA 92521, USA and [3]W. M. Keck Science Department of the Claremont Colleges, Claremont, CA 91711, USA

*Author for correspondence at present address: Mitrani Department of Desert Ecology, Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, 84990 Midreshet Ben-Gurion, Israel (downsc@gmail.com)

†Present address: Department of Biological Sciences, California Polytechnic State University, San Luis Obispo, CA 93407, USA

‡Present address: Diabetes and Obesity Center of Excellence, Department of Medicine, University of Washington, Seattle, WA 98109, USA

Accepted 25 January 2012

SUMMARY

Chronic increases in circulating corticosterone (CORT) generally suppress immune function, but it is not known whether evolved increases necessarily have similar adverse effects. Moreover, the evolution of immune function might be constrained by the sharing of signaling molecules, such as CORT, across numerous physiological systems. Laboratory house mice (*Mus domesticus* Linnaeus) from four replicate lines selectively bred for high voluntary wheel running (HR lines) generally had baseline circulating CORT approximately twofold higher than in four non-selected control (C) lines. To test whether elevated baseline CORT suppresses the inflammatory response in HR mice, we injected females with lipopolysaccharide (LPS). All mice injected with LPS exhibited classic signs of an inflammatory response, including sickness behavior, loss of body mass, reduced locomotor activity (i.e. voluntary wheel running), enlarged spleens and livers, elevated hematocrit and elevated inflammatory cytokines. However, as compared with C mice, the inflammatory response was not suppressed in HR mice. Our results, and those of a previous study, suggest that selective breeding for high voluntary exercise has not altered immune function. They also suggest that the effects of evolved differences in baseline CORT levels may differ greatly from effects of environmental factors (often viewed as 'stressors') that alter baseline CORT during an individual's lifetime. In particular, evolved increases in circulating levels of 'stress hormones' are not necessarily associated with detrimental suppression of the inflammatory response, presumably as a result of correlated evolution of other physiological systems (counter-measures). Our results have important implications for the interpretation of elevated stress hormones and of immune indicators in natural populations.

Supplementary material available online at http://jeb.biologists.org/cgi/content/full/215/10/1651/DC1

Key words: corticosterone, immune function, inflammatory response, selection experiment, trade-off, wheel running.

INTRODUCTION

The ability to respond to infectious agents can significantly affect Darwinian fitness. Although comparative, ecological and evolutionary physiologists have long been interested in the importance of physiological variation in nature, only recently has a serious appreciation for the ecological significance of immune function developed (Lochmiller and Deerenberg, 2000; Ricklefs and Wikelski, 2002; Zuk and Stoehr, 2002; Lee, 2006). Particular interest has focused on the question of whether immune function trades off with other aspects of physiology, and whether immune function mediates trade-offs among other physiological functions (Lochmiller and Deerenberg, 2000; Ricklefs and Wikelski, 2002; Tieleman et al., 2005; Martin et al., 2006; Martin et al., 2008).

In life-history theory, much effort has been dedicated to examining potential trade-offs among major components of life history and also among lower-level traits, including aspects of physiological function, that contribute to survival and reproduction (Stearns, 1992; Clobert et al., 1998; Roff, 2002). Energy is often the currency used to explore these trade-offs (Pough, 1980; Sibly and Calow, 1986; Bennett, 1987; Gittleman and Thompson, 1988; Stearns, 1989), but other mechanisms such as variation in signaling molecules that affect numerous traits might also play a key role in mediating trade-offs (Stearns, 1989; Zera and Harshman, 2001; Ricklefs and Wikelski, 2002; Robinson et al., 2010). Such trade-offs seem particularly plausible in relation to endocrine variation because various hormones (e.g. corticosterone, CORT) influence immune function and many other aspects of physiology and behavior, and might ultimately be involved in the regulation of suites of life-history traits (Harvey et al., 1984; Dallman et al., 1995; Sapolsky et al., 2000; Ricklefs and Wikelski, 2002).

CORT, primarily associated with the stress response, is also involved in the regulation of energy metabolism at rest and during physical activity [see the following and references therein (Dallman et al., 1995; Coleman et al., 1998; Girard and Garland, 2002)] and in the regulation of the immune response (Harvey et al., 1984; Sapolsky et al., 2000). High levels of circulating CORT promote the release of stored energy and alter immune function (Sapolsky et al., 2000; Jacobson, 2005). When elevated for a short time, CORT is thought to contribute to immediate survival of stressful situations by causing increased energy availability. Acute increases in CORT

also lead to redistribution of immune resources and enhancement of some measures of immune function (Martin, 2009). However, chronically high CORT levels produce a potential trade-off because CORT likely suppresses other essential activities, including immune function, growth and reproduction (Raberg et al., 1998; Moberg, 2000; Sapolsky et al., 2000; Wingfield and Kitaysky, 2002).

Additionally, CORT facilitates the return of physiological systems to homeostasis after an inflammatory response. Inflammation is a component of the innate immune system and a general response to microbes. Inflammation is physiologically costly as a result of both increased rates of energy use and potential damage to healthy cells (Klasing and Leshchinsky, 1999). As an immunosuppressor, CORT mediates some negative effects by regulating the strength of the inflammatory response (Munck and Náray-Fejes-Tóth, 1992; Baumann and Gauldie, 1994). Within an individual animal, both acute and chronic increases in CORT suppress immune function (Wingfield et al., 1998; Sapolsky et al., 2000; McEwen and Wingfield, 2003; Korte et al., 2005). However, evolutionary changes have led to differences in baseline levels of CORT among populations and among species (e.g. Abbott et al., 2003; Goymann and Wingfield, 2004; Romero, 2004; Bonier et al., 2009; Cockrem et al., 2009), and it is unclear whether or how such differences affect immune responses, particularly inflammatory responses. To test whether evolutionarily elevated baseline CORT suppresses immune function, we studied laboratory house mice (*Mus domesticus* Linnaeus) from lines that had been artificially selected for 55 generations for high voluntary wheel running (HR lines). Mice from the four replicate HR lines run 2.5–3.0 times more revolutions per day than those from four non-selected control (C) lines when given access to wheels (Garland et al., 2011a). In addition, these HR mice evolved baseline levels of circulating CORT that are approximately twice those of C mice (Malisch et al., 2007; Malisch et al., 2008). [Moreover, during nightly wheel running the HR mice experience exercise-associated elevations in CORT greater than those observed in C mice (Girard and Garland, 2002).] The elevated baseline CORT levels in HR mice potentially increase energy availability during exercise, and hence could be an adaptation (Malisch et al., 2007; Malisch et al., 2008), but the high CORT could also result in a chronic suppression of the inflammatory response (Malisch et al., 2009b), which could be viewed as a cost or trade-off.

In the present study, we challenged HR and C mice with lipopolysaccharide (LPS) from *Escherichia coli*. LPS is an endotoxin found on the cell wall of gram-negative bacteria. LPS imitates a bacterial infection and triggers an inflammatory response and all its associated costs in terms of energy and nutrients. However, LPS itself does not actually cause any sickness; thus, any changes in physiology are purely the result of the immune response (Elin and Wolff, 1976). Hence, by using a LPS inoculation we would be able to identify differences in inflammatory response between the HR and C lines.

The inflammatory response is a suite of physiological and behavioral responses that include decreased body mass, increased sickness behavior (e.g. decreased activity, piloerect fur, panting; Table 1), increased hematocrit, increased inflammatory cytokine levels, and increased spleen and liver masses (Bauss et al., 1987; Hart, 1988; Gabay and Kushner, 1999). Consequently, we measured the inflammatory response in multiple ways that would most broadly represent the suite of characteristics of the inflammatory response.

We hypothesized that HR mice would exhibit a reduced inflammatory response because they have twice as much circulating CORT under baseline conditions. If the HR mice did exhibit such

Table 1. Description and scoring of behaviors recorded during observations of mice

Behavior	Point value	Description of behavior
Body position		(A mouse could only exhibit one body position at a time.)
Rearing	–2	Mouse stands on hind legs with front legs in the air.
Hunched	1	Mouse sits with its front paws either on the ground or tucked near its chest. The mouse's head is often hanging down, but its chest is off the ground.
Sleeping	1	Either the mouse is lying down with its eyes closed or it is lying down with its head tucked under its body (eyes cannot be seen).
Lying down	2	Mouse lies down with its belly and chest on the ground. Lying down was assigned a higher point value than sleeping, because preliminary trials suggested that sick mice were often lying down and not sleeping (their eyes were open). It also took longer for sick mice to fall asleep if they ever slept.
Sprawled	3	Mouse is lying down with its legs and feet extended away from its body. Preliminary trails suggested that extremely sick mice performed this behavior.
Fur position		(A mouse could only exhibit one fur position at a time.)
Partial piloerect	2	Some hair on head, neck and back is erect, but not all.
Complete piloerect	3	Most hair on head, neck and back is erect.
Activities		
Climbing	–3	Mouse's feet are off the floor of the cage, and they are hanging onto the lid of the cage.
Digging	–2	Moving bedding with nose and feet.
Grooming	–1	Scratching, licking paws, chewing on fur.
Exploring	–1	Walking around cage.
Chewing	–1	Chewing (cage top, bedding, feces).
Sniffing	0	Sniffing air or bedding. Can see mouse's whiskers move as it inhales and exhales forcefully.
Panting	2	Mouse is breathing quickly. Movement of mouse's rib cage is rapid and obvious.
Ear trait		(A mouse could only exhibit one ear trait at a time.)
Red ear	2	Veins on the mouse's ear are obviously redder then normal.
Ear hemorrhage	3	Can see burst blood vessels in the mouse's ear (no mouse exhibited this behavior). This behavior was used as an indicator of extreme sickness. As no mouse exhibited this behavior, we know that we did not overdose any mouse with LPS.

Behaviours are listed by category. Point values for each behavior were determined during a preliminary study. Negative values indicate 'active' behaviors while positive values indicate 'sickness' behaviors. Neutral behaviors (e.g. standing still) were not recorded or scored and are not on the list. We recorded all behaviors that each mouse performed, so a mouse could be assigned behaviors from more than one category during a single observation (e.g. a mouse could be assigned rearing and sniffing, or lying down, partial piloerect and panting).

a reduction in inflammatory response, then changes in indicators of an inflammatory response would be expected to be lower than in C mice. Thus, when injected with LPS, HR mice should exhibit a smaller reduction in body mass, less sickness behavior, a smaller hematocrit reduction, lower circulating levels of cytokines, and smaller increases in spleen and liver mass than C mice. We quantified sickness behavior in two complementary ways: an observational behavioral assay (Table 1) and subsequent voluntary wheel running. We also quantified CORT levels in a later generation of mice to ensure that they were still higher in HR than in C mice, as expected from previous studies (Girard and Garland, 2002; Malisch et al., 2007; Malisch et al., 2008).

MATERIALS AND METHODS

Study animals

The mice used in the immune portion of this study were randomly sampled from generation 55 of an ongoing selection experiment for increased voluntary wheel running. A full description of the selection experiment is provided elsewhere (Swallow et al., 1998), and we give only a brief summary here. Overviews of the characteristics of HR mice can be found in previous publications (Rhodes et al., 2005; Garland et al., 2011a; Garland et al., 2011b).

Outbred Hsd:ICR house mice were used as the founder population (Swallow et al., 1998). Mice were assigned to one of four control lines (no selection) or one of four lines bred for high voluntary wheel running (HR mice). The selection criterion was the number of revolutions run on days 5 and 6 of a 6 day period during which mice had access to running wheels (Wahman-type activity wheels, 1.12 m circumference) at ~7–9 weeks of age. For the selected HR lines, the male and female from each family that ran the most were chosen as breeders. For control lines, a random male and female from each family were chosen as breeders. Within lines, male and female pairings were random with the restriction that siblings were not mated to each other. In each generation, 10 pairs of mice (families) were used to propagate the next generation of each line.

For this study, 18 adult females from each of the eight lines were randomly chosen at weaning (21 days of age). However, in one control line (lab designation line 1) only 12 females were available because of low birth rates in that generation. We used only females in this experiment to avoid possible effects of sex and also because females run more than males and have higher circulating CORT levels (Malisch et al., 2007). On average, mice were 10 weeks old at the start of the procedures; however, some line 1 mice were only 6 weeks old because of breeding difficulties in the prior generation. At all times, mice were maintained on a 12 h light:12 h dark cycle with lights on at 07:00 h, and they were provided with food (Harlan Teklad Rodent Diets 8604, Madison, WI, USA) and water *ad libitum* (except during behavioral observations, detailed below). Beginning at weaning, 2–4 mice were housed per cage. All procedures were approved by the Institutional Animal Care and Use Committee of the University of California Riverside, which is accredited by the Association for Assessment and Accreditation of Laboratory Animal Care (AUP A-20080018), and complied with US laws.

Experimental design

Mice from each line were equally divided among three treatments: LPS, sham and baseline. LPS mice received an i.p. injection of 60 μg of LPS (from *Escherichia coli* 0111:B4, product no. L4391, Sigma-Aldrich, St Louis, MO, USA) diluted in 0.9% sodium chloride solution (product no. S8776, Sigma-Aldrich) in a total volume of 0.1 ml. A dose of 60 μg was determined by a pilot study. At this dose, all injected individuals showed sickness behavior (Table 1) (Hart, 1988), but no mortality. Dose was not adjusted for body mass because the pilot study indicated mice representing the typical range of body masses for each line would all exhibit sickness behavior when given a dose of 60 μg. Sham mice were injected with 0.1 ml of 0.9% sodium chloride. Baseline mice received no injection. Six individuals from each line were assigned to each treatment, except for line 1. As only 12 individuals from line 1 were available, each treatment had 4 mice from line 1. Mice were randomly divided among 6 batches of up to 25 mice each, regardless of line or treatment, because behavior (details below) could be measured for a maximum of 25 mice at a time. Batches were stratified by age to minimize any potential effects of age.

General procedure

For each batch, mice were housed individually beginning the night before the experiment started. On the morning of the first experimental day, mice were weighed. LPS mice and sham mice received the appropriate injection at ~08:00 h. Whether LPS or sham mice were injected first was randomized. Individual cages were then placed in a predetermined, random order for behavioral observations, which started at 08:30 h. Each individual was observed 15 times and behavioral observations were conducted at 15 s intervals, so that each individual was observed either every 6 min or every 6 min and 15 s (depending on the size of the batch). Hence, for batch sizes of 25 mice all observations were completed within 1 h 34 min and batches of 24 were completed within 1 h 30 min. For the last batch of 20 mice, the observer waited between observation rounds so the timing between observations of each individual mouse was 6 min (similar to timing for batches of 24 mice). A single observer (C.J.D.), who was blind to the experimental treatment, performed all observations for all batches. All behaviors performed by a mouse during the time of observation were recorded. Point values for each behavior were determined during a pilot study that used an open-ended behavioral catalog and are described in Table 1. Mice injected with LPS are typically lethargic (Hart, 1988), so negative values were assigned to indicate 'active' behaviors whereas positive values indicate 'sickness' behavior. LPS also causes piloerection and panting (Hart, 1988), so these behaviors were assigned positive values. Behavior values were summed for the 15 observation points to yield one overall behavior score. A higher behavior score indicates a sicker mouse. Mice did not have access to food or water during the behavioral experiment, but bedding was present.

At ~10:00 h, mice were anesthetized with isoflurane and blood samples (two samples of 70 μl) were taken retro-orbitally using heparinized microhematocrit tubes (Hoff, 2000). Blood samples were taken approximately 2 h after injection to correspond with expected peak cytokine production (Zanetti et al., 1992). Mice were then weighed and put into standard housing cages with access to Wahman-type activity wheels (see Swallow et al., 1998) for 48 h (13:00 h on injection day to 13:00 h 2 days later). Computers recorded the number of revolutions in 1 min intervals for each wheel. Data were downloaded every 24 h and wheels were checked to ensure freedom of movement. Wheel data collected are described below (see 'Statistics').

After 48 h, mice were removed from the wheels and weighed. A second set of blood samples (two samples of 70 μl) was taken as above. Mice were then killed with CO_2, and spleen and liver were removed by dissection. Livers were weighed and dried at approximately 58°C for 48 h or until constant mass. Spleens were weighed and then flash frozen in liquid nitrogen and stored at –80°C for later analysis. Previous work found that two HR lines have small triceps surae complexes (mini-muscles) (Garland et al., 2002). Line

3 is fixed for mini-muscles and line 6 is polymorphic for mini-muscles. Because the mini-muscle phenotype is associated with other morphological and physiological traits (e.g. Garland et al., 2002; Houle-Leroy et al., 2003; Syme et al., 2005; Rezende, 2006), triceps surae were also dissected from line 6 mice. Triceps surae were then weighed, dried to a constant mass at approximately 58°C, and then weighed again to determine which individuals had the mini-muscle phenotype by inspection of graphs *versus* body mass.

Blood analysis

Immediately after collection, blood samples were centrifuged and plasma was frozen for further analysis. Mean hematocrit was calculated for each of the two sampling periods (before and after access to the running wheels). Wang et al. showed that hematocrit is elevated 2 and 3 h after mice are injected with LPS, but then returns to baseline levels by 4 h after injection (Wang et al., 1991). We calculated the difference between the mean hematocrit before and after mice were placed on wheels as a measure of recovery from the LPS injection.

Blood plasma from the pre-running sampling period was analyzed for four inflammatory cytokines: tumor necrosis factor-alpha (TNF-α), interleukin-6 (IL-6), interleukin-1β (IL-1β) and granulocyte-macrophage colony-stimulating factor (GM-CSF). TNF-α, IL-6 and IL-1β are upstream regulators of the inflammatory response (Beutler and Cerami, 1989; Baumann and Gauldie, 1994; Beutler, 2004). TNF-α and IL-1β are pro-inflammatory cytokines, whereas the anti-inflammatory properties of IL-6 actually help regulate the inflammatory response (Baumann and Gauldie, 1994; Xing et al., 1998; Gabay and Kushner, 1999). We measured all three cytokines because there are numerous, redundant pathways for mounting an inflammatory response (e.g. Fantuzzi et al., 1996; Gabay and Kushner, 1999), and it is possible that C and HR mice have evolved to utilize alternative cytokine pathways. GM-CSF up-regulates leukocyte proliferation in red bone marrow, a component of the inflammatory response (Gasson, 1991). Plasma was processed using a Mouse Inflammatory Four-Plex Antibody Bead Kit (Invitrogen Corporation, Camarillo, CA, USA). All protocols recommended by the kit manufacturer were followed, except we used 25 µl of assay dilutes and samples instead of the recommended 50 µl. Single samples were tested and compared with a standard curve for quantification. Cytokines were quantified in pg ml^{-1} using Luminex multiplex technology (Luminex 200, Luminex Corporation, Austin, TX, USA).

CORT

Males from generation 60 were used. As per usual procedures, mice were weaned at 21 days of age, toe-clipped for identification, and housed randomly in same-sex groups of four. Mice were 6–9 weeks old at the time of blood sampling. The light cycle was maintained at 12 h light:12 h dark (lights on at 07:00 h) and food and water were given *ad libitum*.

One male mouse from each of 10 families per line was sampled for this study (total N=80). As part of a separate study of the effects of restraint stress (see Malisch et al., 2007), these animals were sampled during the routine selection procedure, with the addition of a seventh day of wheel access. Following the sixth day of wheel access, half of the mice were removed from their cage, subjected to a 40 min restraint stressor (see Malisch et al., 2007), then returned to their cage with wheel access. The other half were left undisturbed.

Blood samples were taken from all 80 mice (20 HR restrained, 20 HR not restrained, 20 C restrained and 20 C not restrained) ~24 h post-restraint (between 14:30 h and 16:15 h). Blood was acquired from the retro-orbital sinus (Hoff, 2000) under isofluorane sedation (Malisch et al., 2009b). Precautions were taken to minimize stress to the mice before and during the bleeding process, and blood was acquired within 3 min of removal from the cage.

Blood samples were centrifuged (13,300 r.p.m.) at 4°C, then plasma was decanted and maintained at –80°C until assays. Total plasma CORT concentration was determined in duplicate with a commercially available kit following previous protocols (Malisch et al., 2008).

Statistics

Mixed-model nested ANOVA were used to assess variation in the innate immune responses. Hereafter, 'linetype' will indicate differences between HR and C mice and 'innate treatment' will indicate differences among baseline, sham and LPS treatments. The test of our main hypothesis is the interaction between linetype and innate treatment. For example, if HR mice had suppressed inflammatory responses, then we would expect TNF-α to be higher in C mice injected with LPS than in HR mice injected with LPS. However, we might not expect a difference in TNF-α in baseline or sham-injected mice.

The general model for each mixed model included linetype, innate treatment, the interaction of linetype × innate treatment, and a dummy variable indicating mini-muscle as fixed effects. In the general model, experimental batch, line nested within linetype and the interaction innate treatment × line nested within linetype were specified as random effects. We investigated all individuals with standardized residuals greater than |3| from all analyses and tried to determine why the residual was so large. Based on this analysis, we excluded two mice from all analyses: one mouse from the baseline treatment was excluded because of abnormally low wheel running. One mouse from the sham treatment was excluded because it had an extremely large change in mass indicating that an incorrect datum was recorded. Details of the model used for each response variable are presented below. Mixed models were run in SAS version 9.2 64-bit (SAS Institute Inc., Cary, NC, USA) using Proc Mixed.

Sickness behavior

Higher sickness behavior scores indicated sicker mice. Behavior score was modeled using the general model without alteration. We also investigated whether our behavior score could be predicted from the level of circulating cytokines TNF-α, IL-6, IL-1β and GM-CSF. We ran a separate model for each cytokine, by adding cytokine level as a predictor covariate to the general mixed model. Finally, we ran a principal components analysis (PCA) on the cytokine values and added the first principal component as a predictor covariate to the general mixed model if a linear combination describing the overall profile of circulating cytokines predicted sickness behavior.

Body mass

Body mass was measured twice, 2 h after injection (initial body mass) and after 48 h of wheel running (final body mass); both were used as response variables. The residuals for initial and final body mass we not normally distributed, so we log-transformed these data. Change in body mass, defined as initial body mass minus final body mass, was also calculated. Because change in body mass could be affected by both initial body mass and wheel running, we tested statistical models that included both covariates. They did not reveal anything different about linetype, innate treatment or the interaction between linetype and innate treatment, so we only present results from the general model.

Hematocrit

We measured hematocrit twice, 2 h after injection and after 48 h of wheel running. We also analyzed the change in hematocrit, defined as hematocrit after wheel running minus hematocrit 2 h after injection. These response variables were analyzed with the unaltered general model.

Organ masses

The statistical models for dry liver and wet spleen included the covariate log-transformed final body mass. The residuals for raw dry liver mass and wet spleen mass were not normally distributed; thus, we log transformed these data.

Cytokines

Raw and log-transformed TNF-α, IL-6, IL-1β and GM-CSF (pg ml^{-1}) exhibited large differences in variance among the innate treatments (i.e. baseline, sham and LPS treatments); thus, these data were rank transformed to equalize variance (Conover and Iman, 1981; Potvin and Roff, 1993). We removed one sham treatment mouse from all of the cytokine analyses because it had an extremely high IL-1β measurement indicating an anomaly in the cytokine measurements. We also performed a PCA on the cytokine data and used the first principal component as a response variable to determine whether the overall cytokine profile was different between the HR and C mice. The principal component scores had unequal variance between the HR and C mice, and so we rank transformed these values for analysis. We used two complementary approaches to test for cytokine differences between the HR and C mice. First, we used the mixed model approach described above. Four mice in the IL-1β data set and six mice in the IL-6 data set had standardized residuals greater than |3|. These mice were all from the LPS treatment and represent variation in the cytokine so we left these individuals in the analysis. Second, we modeled each innate treatment (LPS, sham or baseline) separately. In selection treatments, the line is the unit of replication, not the individual (Henderson, 1989; Garland and Rose, 2009); thus, selection experiments often have relatively low power to detect differences between selection treatments. In addition, ANOVA may have relatively low power to detect interaction effects (Wahlsten, 1990). Therefore, with this second analysis, we attempted to increase our power by testing for differences between C and HR mice within each innate treatment. For these analyses we also used mixed models, but the fixed effects were linetype and a dummy variable indicating mini-muscle; the random effects were experimental batch and line nested within linetype. The results for the second statistical method were similar for all of the cytokines, so we present only results for rank TNF-α to represent the second method.

Wheel running

We analyzed all four measures of wheel running for two time periods: day 1 (first 24 h) alone, day 2 (second 24 h) alone. Distance run was calculated as the number of revolutions each mouse ran on the wheel. Distance run data were square-root transformed for both days to normalize the residuals. Running duration was the number of 1 min intervals during which an individual tallied at least one revolution, and data from both days were square-root transformed to normalized residuals. Mean speed (revolutions min^{-1}) was calculated as the total number of revolutions divided by the number of active intervals. Mean speed data for both days were square-root transformed to normalize residuals. Maximum speed was calculated as the maximum number of revolutions in a 1 min interval, and was not transformed for either day. Before mice were put on wheels, we determined how much the wheel rotated after acceleration to a constant velocity. This measure of freeness was square-root transformed to normalize its distribution, and then used as a covariate in all the analyses of wheel running.

CORT levels

The effect of linetype, restraint and their interaction were tested with a mixed model in SAS Proc Mixed, each with 1 and 6 d.f. Covariates of age, time of day and bleed delay time were included.

RESULTS

As expected, sickness behavior (Fig. 1) was greater in LPS mice than in sham and baseline mice (Table 1; $F_{2,12}$=47.3, P<0.001), but did not significantly differ between HR and C mice (supplementary material Table S1; $F_{1,6}$=3.23, P=0.122). Unexpectedly, sickness behavior was greater in mice with the mini-muscle phenotype than in those with normal muscles (supplementary material Fig. S1A, Table S1; $F_{1,106}$=14.6, P<0.001). TNF-α ($F_{1,104}$=1.18, P=0.280) and the other cytokine measurements (including the principal component scores indicating overall cytokine profile; results not shown) were not significant predictors of sickness behavior at the level of individual variation in the ANCOVA models.

Initial body mass did not significantly differ among LPS, sham and baseline mice ($F_{2,12}$=0.36, P=0.707, Fig. 2A). However, body mass after wheel running was lower in LPS-treated mice than in sham and baseline mice ($F_{2,12}$=13.7, P<0.001, Fig. 2B). Thus, LPS-treated mice lost more mass than sham and baseline mice during the 48 h between mass measurements ($F_{2,12}$=48.7, P<0.001, Fig. 2C). Initial body mass, final body mass and change in body mass did not differ significantly between HR and C mice or between mice with and without mini-muscles (supplementary material Table S1).

Two hours after injection, LPS-treated mice had higher hematocrit than sham or baseline mice ($F_{2,12}$=35.1, P<0.001, Fig. 3A). After wheel running, there was no difference in hematocrit among these innate treatment groups ($F_{2,12}$=0.74, P=0.497, Fig. 3B). Thus, between the first and second measurements, hematocrit decreased

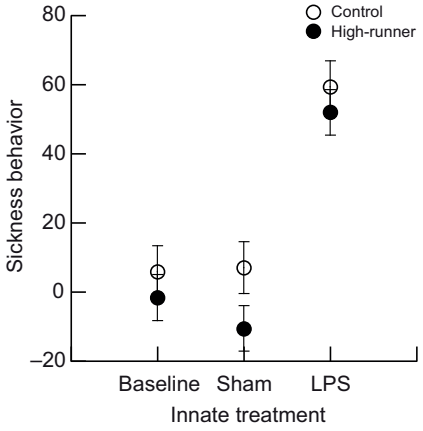

Fig. 1. Least squares means (±s.e.m.) for sickness behavior measured in high-runner and control lines of mice without injections (baseline mice), with sham injections and with lipopolysaccharide (LPS) injections. Higher scores indicate greater sickness behavior (see Table 1 for scoring scheme). Sickness behavior in LPS mice was significantly different (P<0.05) from that in baseline and sham mice, but there was no significant difference between high-runner and control lines (see supplementary material Table S1).

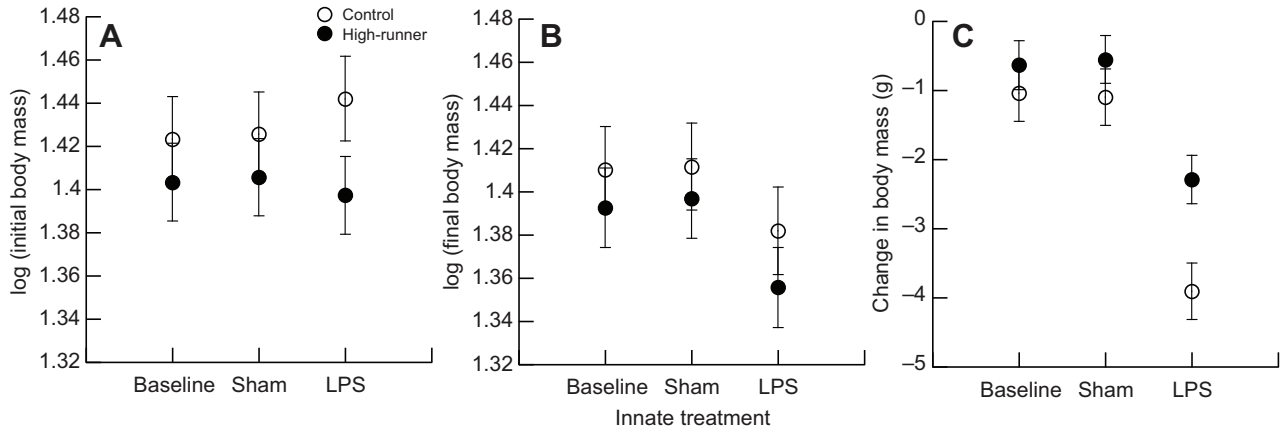

Fig. 2. Least squares means (±s.e.m.) for initial body mass (A), final body mass (B) and change in body mass (C) measured in high-runner and control mice without injections (baseline mice), with sham injections and with LPS injections. Final body mass and change in body mass in LPS mice were significantly different (P<0.05) from those in baseline and sham mice (supplementary material Table S1). Body mass was originally measured in grams.

more for the LPS-treated mice than for the sham and baseline mice ($F_{2,12}$=26.8, P<0.001, Fig. 3C). None of the hematocrit variables were significantly different between HR and C mice or between mice with and without mini-muscles (supplementary material Table S1).

The log body mass was a strong predictor of both log dry liver mass (β=1.27±0.09; $F_{1,105}$=208.2, P=<0.001) and log wet spleen mass (β=1.27±0.13; $F_{1,105}$=96.0, P<0.001). However, log dry liver mass and log wet spleen mass were greater in LPS mice ($F_{2,12}$=22.24, P<0.001, Fig. 4A; $F_{2,12}$=25.8, P<0.001, Fig. 4B, respectively). On average, mini-muscle mice had heavier livers (log dry mass, $F_{1,105}$=13.12, P<0.001, supplementary material Fig. S1B) and spleens (log wet mass, $F_{1,105}$=8.14, P=0.005, supplementary material Fig. S1C) than those with normal muscles. Neither organ mass differed significantly between HR and C mice (supplementary material Table S1).

LPS-treated mice had significantly higher cytokine levels than other mice (rank TNF-α: $F_{2,12}$=170.3, P<0.001; rank GM-CSF: $F_{2,12}$=75.0, P<0.001; rank IL-1β: $F_{2,12}$=158.0, P<0.001; rank IL-6: $F_{2,12}$=301.4, P<0.001, Fig. 5). Cytokine levels did not differ between HR and C mice or between mice with and without mini-muscles (supplementary material Table S1). When rank TNF-α of LPS, sham and baseline mice were examined separately, there was no significant difference between HR and C mice or between mice with and without mini-muscles (supplementary material Table S2). Results for all of the other cytokines were similar to those for rank TNF-α (not shown). The overall cytokine profile (i.e. the first principal component from a PCA of all of the cytokine data) had a quantitatively similar pattern to each individual cytokine (results not shown).

As shown in Fig. 6, mice from the HR lines ran significantly farther than those from control lines, irrespective of innate treatment, mainly because the former ran at higher speeds (see supplementary material Table S1 for statistical results). The interaction between innate treatment and linetype was significant for running duration on day 1 ($F_{2,12}$=4.87, P=0.028), Aside from this, the interaction term was never statistically significant for any response variable (supplementary material Table S1).

HR mice (least squares mean ± s.e.m., 36.5±2.93 ng ml^{-1}) had statistically higher plasma CORT concentrations (2-tailed, P=0.0158) than C mice (22.6±2.86 ng ml^{-1}), with no statistical effect of restraint (2-tailed, P=0.1451), no interaction between linetype and restraint (P=0.2742), and no significant effect of any covariate (all 2-tailed, P>0.17).

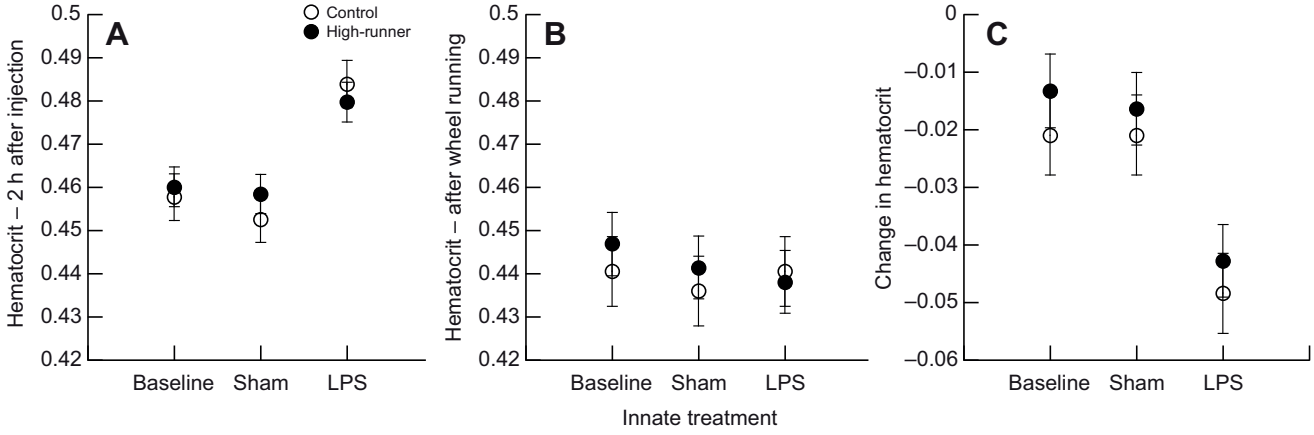

Fig. 3. Least squares means (±s.e.m.) for hematocrit measured 2 h after injection (A), hematocrit measured after wheel running (B) and change in hematocrit (C) measured in high-runner and control mice without injections (baseline mice), with sham injections and with LPS injections. Hematocrit 2 h after injection and the change in hematocrit in LPS mice were significantly different (P<0.05) from those in baseline and sham mice (supplementary material Table S1).

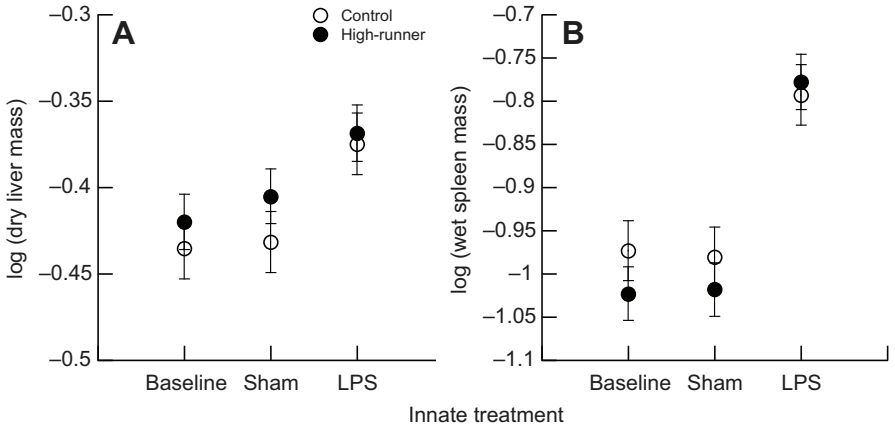

Fig. 4. Least squares means (±s.e.m.) for dry liver mass (A) and wet spleen mass (B) measured in high-runner and control mice without injections (baseline mice), with sham injections and with LPS injections. Based on ANCOVA adjusting for variation in body mass, both organ masses in LPS mice were significantly different (P<0.05) from those in baseline and sham mice (supplementary material Table S1).

DISCUSSION

As expected, mice injected with LPS exhibited classic signs of an inflammatory response, including sickness behavior, weight loss, increased hematocrit (2 h after injection), heavier livers and spleens, increased cytokine levels and decreased activity (voluntary wheel-running activity) (supplementary material Table S1; Figs 1–6). Sham-treated and non-manipulated 'baseline' mice showed no statistical differences from each other (results not shown). Hence, changes in physiology and behavior in the LPS-injected mice were attributable to an inflammatory response caused by the LPS, not simply the process of injection.

CORT has anti-inflammatory properties and can inhibit the synthesis and release of IL-1β, IL-6, TNF-α and GM-CSF (Munck and Náray-Fejes-Tóth, 1992; Sapolsky et al., 2000). Because HR mice have a baseline circulating CORT concentration approximately twice that of C mice (Malisch et al., 2007), we hypothesized that HR mice would have a reduced inflammatory response. The key statistic to test our hypothesis is the interaction between linetype and innate treatment. Sham and baseline mice were not immune challenged and differences between HR and C mice for these treatments were not expected. In contrast, given their increased CORT levels, we predicted that in response to injection with LPS, HR mice would have a weaker inflammatory response than C mice. Consequently, we predicted differences between HR and C mice only in the LPS-injected group. If these predictions were supported by the data, then a statistically significant interaction between linetype and innate treatment was expected.

Of the 21 response variables, one (running duration on day 1) had a significant (P<0.05) interaction and three (body mass, log-transformed wet spleen mass and wheel running on day 1) had marginally significant (P<0.1) interactions (supplementary material Table S1). These results could be interpreted as suggesting that our

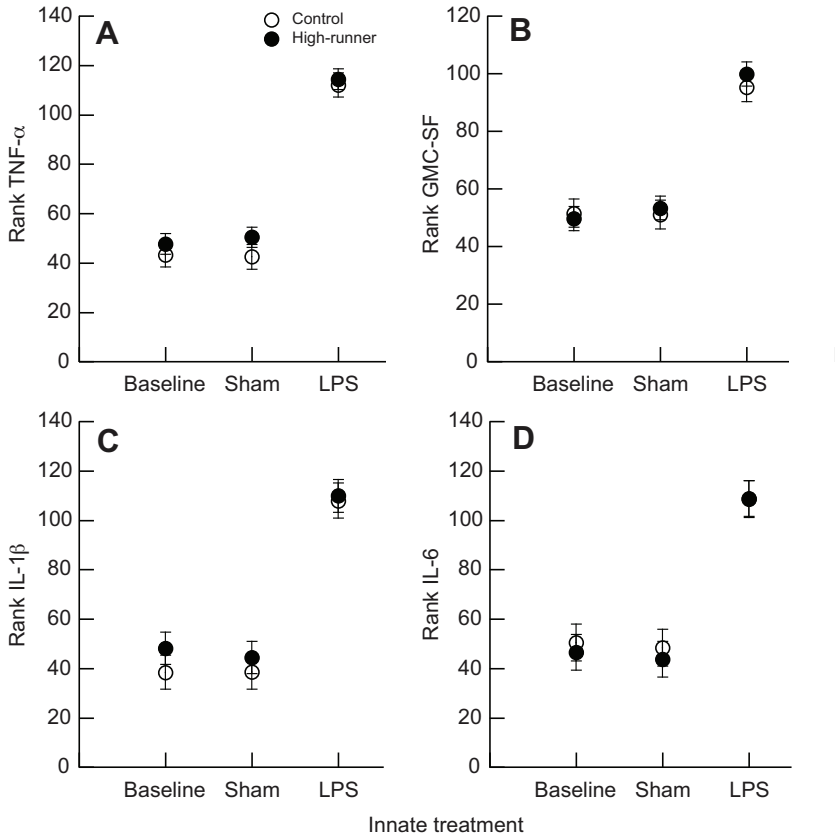

Fig. 5. Least squares means (±s.e.m.) for the cytokines tumor necrosis factor-alpha (TNF-α, A), granulocyte-macrophage colony-stimulating factor (GM-CSF, B), interleukin-1β (IL-1β, C) and interleukin-6 (IL-6, D) measured in high-runner and control mice without injections (baseline mice), with sham injections and with LPS injections. All cytokine data were originally measured in pg ml^{-1}, but were rank transformed for statistical analyses. All cytokines in LPS mice were significantly different (P<0.05) from those in baseline and sham mice, but there was no significant difference between HR and control lines, and no statistical interactions were detected (supplementary material Table S1).

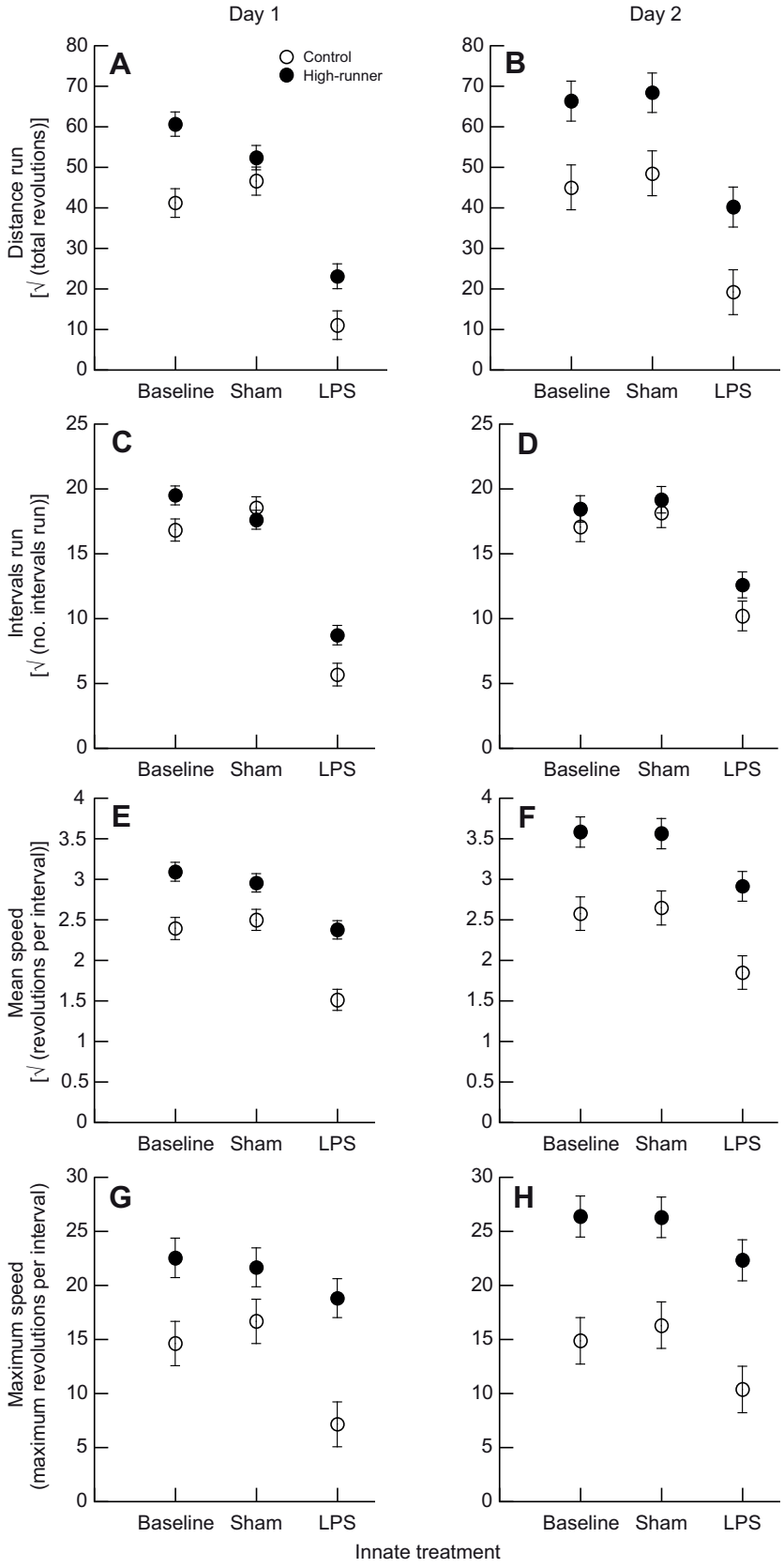

Fig. 6. Least squares means (±s.e.m.) for wheel running variables measured in high-runner and control mice without injections (baseline mice), with sham injections and with LPS injections. All measurements of wheel running were affected by LPS treatment (supplementary material Table S1). As expected from numerous previous studies (see Discussion), mice from the HR lines ran more total revolutions and at higher speeds than those from the non-selected control lines (supplementary material Table S1). The only treatment × linetype interaction that reached statistical significance was running duration on day 1 (P=0.028).

experiment lacked sufficient statistical power to detect a difference in inflammatory responses of HR and C mice. However, we think this is unlikely given the clearly non-significant interaction terms for the other 17 response variables, particularly the cytokine results (Fig. 5). Because they are upstream regulators of inflammatory response, the cytokine data are our best indicator of the strength of that response (Beutler and Cerami, 1989; Baumann and Gauldie, 1994; Beutler, 2004), and the interaction term is not significant for

these data (supplementary material Table S1). Furthermore, no difference existed between the C and HR mice when each innate treatment (baseline, sham and LPS) was investigated individually (supplementary material Table S2). Finally, type 1 error due to the large number of response variables could explain the single significant interaction we found for running duration on day 1. Thus, we conclude that inflammatory response does not significantly differ between HR and C mice.

The immune system has many redundant pathways for achieving the same end result (Suffredini et al., 1999; Parkin and Cohen, 2001), so it is possible that HR and C mice could have utilized different cytokine cascades for mounting an inflammatory response. However, we did not observe any differences in cytokine values between HR and C mice (supplementary material Tables S1, S2).

Our results are consistent with a previous study showing that, overall, HR and C mice did not differ in another measure of immune function – their ability to clear a parasitic nematode from the gut (Malisch et al., 2009b). Thus, HR mice do not have an impaired ability to clear micro- or macro-parasites. However, the immune system is complicated and multifaceted, with both innate and adaptive components, and its response depends on the specific type of infectious agent or parasite (Medzhitov, 2007). Further, one aspect of the immune system might be impaired while other aspects are enhanced to maintain overall immune function; therefore, it is important to study numerous aspects of the immune system (Norris and Evans, 2000; Matson et al., 2006). In this experiment, we tested only the difference in inflammatory response, a component of innate immune function. In contrast, Malisch et al. (Malisch et al., 2009b) implicitly tested both innate and adaptive responses, as both are required to clear nematodes from the gut (Maizels and Yazdanbakhsh, 2003; Anthony et al., 2007). Differences in adaptive immunity would result in differences between HR and C mice in their ability to clear parasites, but no difference in inflammatory response – i.e. we would expect a significant difference between HR and C mice for the study by Malisch et al. (Malisch et al., 2009b) but not for the present study. However, without testing adaptive immune function directly, we cannot rule out this explanation, so adaptive immune function should be tested directly in future studies of these mice.

When considering the results of Malisch et al. (Malisch et al., 2009b) with our present results, neither study suggests that evolved differences in baseline CORT lead to suppression of the immune response in HR mice as had been hypothesized. Instead, they suggest two potential scenarios regarding how selective breeding for voluntary wheel running has affected immunity in HR mice.

First, CORT levels may not have been elevated in the particular sample of HR mice we studied from generation 55, but we consider this explanation unlikely. The difference in baseline CORT levels between the HR and C mice was first detected in generation 21 (Girard and Garland, 2002), and it was reconfirmed in generations 35 and 39 (Malisch et al., 2007; Malisch et al., 2008), as well as in generation 60 (this experiment).

Second, the effects of evolved differences in baseline CORT levels may be very different from the effects caused by environmental factors that alter baseline CORT concentrations during the course of an individual's life. The difference in circulating CORT levels between HR and C represents a correlated response to artificial selection on running behavior (Girard and Garland, 2002; Malisch et al., 2007; Malisch et al., 2008). Because the endocrine system does not evolve in isolation from other organ systems, the relationships among CORT, voluntary exercise and immune function may be different in mice bred for high voluntary exercise *versus* those responding to a shorter-term stressor. For example, CORT is often released when a system's homeostasis is altered and/or when energy demands increase, such as during exercise (Chrousos and Gold, 1992; Stranahan et al., 2008). CORT released during exercise triggers a response that restores the energy balance (Girard and Garland, 2002; Droste et al., 2003; Stranahan et al., 2008). However, artificial selection for increased exercise may pose a threat to homeostasis because of the co-evolved increase in CORT level. If the new baseline level of CORT was high enough to initiate downstream responses to stress, then HR mice would exhibit chronic stress. In response, the set point for the level of circulating CORT required to trigger a stress response might have evolved in kind to a higher level. Under this scenario, the twofold increase in circulating CORT seen in HR mice would not cause a stress response and would not suppress immune function. In summary, our results suggest that there may be a fundamental difference in the downstream consequences of high CORT when the increase is evolved *versus* caused by short-term stress. Our results are also consistent with the hypothesis developed for testosterone that evolved differences in hormone levels will not always lead to negative pleiotropic effects (Hau, 2007). Thus, patterns regarding how hormones alter immune function might be more complicated than originally hypothesized (Fuxjager et al., 2011).

Mini-muscle mice

Mice with mini-muscles exhibited greater sickness behavior (supplementary material Fig. S1A) and had larger spleens and livers (Fig. 2B,C) than those with normal triceps surae muscles, suggesting an elevated inflammatory response in the former. However, the cytokine data conflict with that interpretation because levels did not differ between mini-muscle and normal mice (supplementary material Table S1).

One possible explanation is that the differences in sickness behavior and organ masses of mini-muscle mice *versus* others might be a consequence of the overall phenotype of the mini-muscle individuals. Mini-muscle mice differ phenotypically in many ways from mice with normal muscles (Garland et al., 2002; Swallow et al., 2005; Meek et al., 2009). For example, they have larger mass-corrected spleens and livers than normal muscle mice under baseline conditions (Garland et al., 2002; Swallow et al., 2005; Meek et al., 2009). Thus, if the inflammatory response causes an absolute increase in liver and spleen masses, then mini-muscle mice will still have larger mass-corrected organs than normal muscle mice and appear to be more sick based on organ mass alone.

However, the same logic cannot apply to sickness behavior. Sickness behavior includes many measures of activity (see body position and activities data in Table 1), and lower sickness scores are associated with higher physical activity. Previous studies have shown that mini-muscle mice sometimes run farther on wheels than normal mice (Syme et al., 2005; Hannon et al., 2008) and that HR mice in general have higher home-cage activity than C mice when housed without wheels (Malisch et al., 2009a). If the same correlation is true for mini-muscle mice, then running farther might be correlated with higher general activity in mini-muscle mice. Thus, if mini-muscle mice decrease their activity the same absolute amount as mice with normal triceps surae, they will still be more active and their higher activity could result in lower sickness behavior scores. In fact, mini-muscle mice had higher sickness behavior scores than those with normal muscles (supplementary material Fig. S1A). Although mini-muscle mice sometimes run farther on running wheels than those with normal muscles (Syme et al., 2005; Hannon

et al., 2008), this may not translate into greater activity in home cages. Thus, variation in home cage activity cannot easily explain the difference in sickness behavior scores between mini-muscle and normal-muscle mice.

Alternatively, mini-muscle mice might differ from normal mice in the density or activity of cytokine receptors, or of other molecules associated with the inflammatory response. For example, mice deficient in TNF-α receptor and mice with mutations in Toll-like receptor 4 are less sensitive to the lethal effects of large doses of LPS (Pfeffer et al., 1993; Qureshi et al., 1999). Thus, it is also possible that the receptors for inflammatory cytokines are less abundant in mini-muscle mice and they exhibit increased sickness behavior even though they do not have elevated cytokine levels.

CONCLUSIONS

When comparing circulating levels of 'stress hormones' among species and among individuals, researchers commonly assume that higher levels have greater downstream effects, such as reducing immune function (see Martin et al., 2006). Our study and that of Malisch et al. (Malisch et al., 2009b) indicate that an observation of elevated baseline circulating CORT levels is not sufficient to make accurate predictions about the suppression of immune function. Selective breeding for voluntary wheel running in mice apparently has not altered the inflammatory response, despite the more than twofold higher circulating CORT levels in HR mice. Thus, evolved increases in baseline levels of CORT may not have the same negative immunosuppressive effects as the elevation of CORT caused by stress within an individual's lifetime. These results have important implications for comparative studies examining hormonal differences among species and understanding the consequences of variation in circulating CORT levels that occur during an individual's lifetime *versus* across generations on an evolutionary timescale.

ACKNOWLEDGEMENTS

We thank K. Hunter and V. Lombardi for their help with immune techniques and lab space. We also thank members of the Garland lab for their help with the mouse colony.

FUNDING

This study was supported by the National Science Foundation [grant nos IOB-0543429 and IOS-1121273 to T.G.; IOS-0344994 to J.H.], and a grant from the Ecology, Evolution, and Conservation Biology Program at University of Nevada, Reno to C.J.D. H.S. was partially supported by a Chancellor's Postdoctoral Fellowship from University of California, Riverside.

REFERENCES

Abbott, D. H., Keverne, E. B., Bercovitch, F. B., Shively, C. A., Medoza, S. P., Saltzman, W., Snowdon, C. T., Ziegler, T. E., Banjevic, M., Garland, T., Jr et al. (2003). Are subordinates always stressed? A comparative analysis of rank differences in cortisol levels among primates. *Horm. Behav.* **43**, 67-82.

Anthony, R. M., Rutitzky, L. I., Urban, J. F., Stadecker, M. J. and Gause, W. C. (2007). Protective immune mechanisms in helminth infection. *Nat. Rev. Immunol.* **7**, 975-987.

Baumann, H. and Gauldie, J. (1994). The acute-phase response. *Immunol. Today* **15**, 74-80.

Bauss, F., Droge, W. and Mannel, D. N. (1987). Tumor-nectosis-factor mediates endotoxic effects in mice. *Infect. Immun.* **55**, 1622-1625.

Bennett, A. F. (1987). The accomplishments of ecological physiology. In *New Directions in Ecological Physiology* (ed. M. E. Feder, A. F. Bennett, W. W. Burggren and R. B. Huey), pp. 364. Cambridge: Cambridge University Press.

Beutler, B. (2004). Innate immunity: an overview. **40**, 845-859.

Beutler, B. and Cerami, A. (1989). The biology of cachctin/TNF-a primary mediator of the host response. *Annu. Rev. Immunol.* **7**, 625-655.

Bonier, F., Martin, P. R., Moore, I. T. and Wingfield, J. C. (2009). Do baseline glucocorticoids predict fitness? *Trends Ecol. Evol.* **24**, 634-642.

Chrousos, G. P. and Gold, P. W. (1992). The concepts of stress and stress system disorders: overview of physical and behavioral homeostasis. *J. Am. Med. Assoc.* **267**, 1244-1252.

Clobert, J., Garland, T., Jr and Barbault, R. (1998). The evolution of demographic tactics in lizards: a test of some hypotheses concerning life history evolution. *J. Evol. Biol.* **11**, 329-364.

Cockrem, J. F., Barrett, D. P., Candy, E. J. and Potter, M. A. (2009). Corticosterone responses in birds: Individual variation and repeatability in Adelie penguins (*Pygoscelis adeliae*) and other species, and the use of power analysis to determine sample sizes. *Gen. Comp. Endocrinol.* **163**, 158-168.

Coleman, M. A., Garland, T., Jr, Marler, C. A., Newton, S. S., Swallow, J. G. and Carter, P. A. (1998). Glucocorticoid response to forced exercise in laboratory house mice (*Mus domesticus*). *Physiol. Behav.* **63**, 279-285.

Conover, W. J. and Iman, R. L. (1981). Rank transofrmations as a bridge between parametric and nonparametric statistics. *Am. Stat.* **35**, 124-129.

Dallman, M. F., Akana, S. F., Strack, A. M., Hanson, E. S. and Sebastian, R. J. (1995). The neural network that regulates energy balance is responsive to glucocorticoids and insulin and also regulates HPA axis responsivity at a site proximal to CRF neurons. In *Stress – Basic Mechanisms and Clinical Implications*, Vol. 771 (ed. G. P. Chrousos, R. McCarty, K. Pacak, G. Cizza, E. Sternberg, P. W. Gold and R. Kvetnansky), pp. 730-742. New York: New York Academy of Sciences.

Droste, S. K., Gesing, A., Ulbricht, S., Müller, M. B., Linthorst, A. C. E. and Reul, J. M. H. M. (2003). Effects of long-term voluntary exercise on the mouse hypothalamic-pituitary-adrenocortical axis. *Endocrinology* **144**, 3012-3023.

Elin, R. J. and Wolff, S. M. (1976). Biology of endotoxin. *Annu. Rev. Med.* **27**, 127-141.

Fantuzzi, G., Hui, Z., Faggioni, R., Benigni, F., Ghezzi, P., Sipe, J. D., Shaw, A. R. and Dinarello, C. A. (1996). Effect of endotoxin in IL-1 beta-deficient mice. *J. Immunol.* **157**, 291-296.

Fuxjager, M. J., Foufopoulos, J., Diaz-Uriarte, R. and Marler, C. A. (2011). Functionally opposing effects of testosterone on two different types of parasite: implications for the immunocompetence handicap hypothesis. *Funct. Ecol.* **25**, 132-138.

Gabay, C. and Kushner, I. (1999). Mechanisms of disease: Acute-phase proteins and other systemic responses to inflammation. *N. Engl. J. Med.* **340**, 448-454.

Garland, T., Jr and Rose, M. R. (2009). *Experimental Evolution: concepts, Methods, and Applications of Selection Experiments.* Berkeley, CA: University of California Press.

Garland, T., Jr, Morgan, M. T., Swallow, J. G., Rhodes, J. S., Girard, I., Belter, J. G. and Carter, P. A. (2002). Evolution of a small-muscle polymorphism in lines of house mice selected for high activity levels. *Evolution* **56**, 1267-1275.

Garland, T., Jr, Kelly, S. A., Malisch, J. L., Kolb, E. M., Hannon, R. M., Keeney, B. K., Van Cleave, S. L. and Middleton, K. M. (2011a). How to run far: multiple solutions and sex-specific responses to selective breeding for high voluntary activity levels. *Proc. Biol. Sci.* **278**, 574-581.

Garland, T., Jr, Schutz, H., Chappell, M. A., Keeney, B. K., Meek, T. H., Copes, L. E., Acosta, W., Drenowatz, C., Maciel, R. C., van Dijk, G. et al. (2011b). The biological control of voluntary exercise, spontaneous physical activity and daily energy expenditure in relation to obesity: human and rodent perspectives. *J. Exp. Biol.* **214**, 206-229.

Gasson, J. C. (1991). Molecular physiology of granulocyte-marophage colony-stimulating factor. *Blood* **77**, 1131-1145.

Girard, I. and Garland, T., Jr (2002). Plasma corticosterone response to acute and chronic voluntary exercise in female house mice. *J. Appl. Physiol.* **92**, 1553-1561.

Gittleman, J. L. and Thompson, S. D. (1988). Energy allocation in mammalian reproduction. *Am. Zool.* **28**, 863-875.

Goymann, W. and Wingfield, J. C. (2004). Allostatic load, social status, and stress hormones – the costs of social status matter. *Anim. Behav.* **67**, 591-602.

Hannon, R. M., Kelly, S. A., Middleton, K. M., Kolb, E. M., Pomp, D. and Garland, T., Jr (2008). Phenotypic effects of the 'mini-Muscle' allele in a large HR x C57BL/6J mouse backcross. *J. Hered.* **99**, 349-354.

Hart, B. L. (1988). Biological basis of the behavior of sick animals. *Neurosci. Biobehav. Rev.* **12**, 123-137.

Harvey, S., Phillips, J. G., Rees, A. and Hall, T. R. (1984). Stress and adrenal-function *J. Exp. Zool.* **232**, 633-645.

Hau, M. (2007). Regulation of male traits by testosterone: implications for the evolution of vertebrate life histories. *BioEssays* **29**, 133-144.

Henderson, N. D. (1989). Interpreting studies that compare high-selected and low-selected line on new characters. *Behav. Genet.* **19**, 473-502.

Hoff, J. (2000). Methods of blood collection in the mouse. *Lab. Anim.* **29**, 47-53.

Houle-Leroy, P., Guderley, H., Swallow, J. G. and Garland, T., Jr (2003). Artificial selection for high activity favors mighty mini-muscles in house mice. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **284**, R433-R443.

Jacobson, L. (2005). Hypothalamic-pituitary-adrenocortical axis regulation. *Endocrinol. Metabol. Clin. North Amer.* **34**, 271-292.

Klasing, K. C. and Leshchinsky, T. V. (1999). *Functions, Costs, and Benefits of the Immune System During Development and Growth.* Johannesburg: BirdLife South Afica.

Korte, S. M., Koolhaas, J. M., Wingfield, J. C. and McEwen, B. S. (2005). The Darwinian concept of stress: benefits of allostasis and costs of allostatic load and the trade-offs in health and disease. *Neurosci. Biobehav. Rev.* **29**, 3-38.

Lee, K. A. (2006). Linking immune defenses and life history at the levels of the individual and the species. *Integr. Comp. Biol.* **46**, 1000-1015.

Lochmiller, R. L. and Deerenberg, C. (2000). Trade-offs in evolutionary immunology: just what is the cost of immunity? *Oikos* **88**, 87-98.

Maizels, R. M. and Yazdanbakhsh, M. (2003). Immune regulation by helminth parasites: cellular and molecular mechanisms. *Nat. Rev. Immunol.* **3**, 733-744.

Malisch, J. L., Saltzman, W., Gomes, F. R., Rezende, E. L., Jeske, D. R. and Garland, T., Jr (2007). Baseline and stress-induced plasma corticosterone concentrations of mice selectively bred for high voluntary wheel running. *Physiol. Biochem. Zool.* **80**, 146-156.

Malisch, J. L., Breuner, C., Gomes, F., Chappell, M. and Garland, T., Jr (2008). Circadian pattern of total and free corticosterone concentrations, corticosteroid-

binding globulin, and physical activity in mice selectively bred for high voluntary wheel-running behavior. *Gen. Comp. Endocrinol.* **156**, 210-217.
Malisch, J. L., Breuner, C. W., Kolb, E. M., Wada, H., Hannon, R. M., Chappell, M. A., Middleton, K. M. and Garland, T., Jr (2009a). Behavioral despair and home-cage activity in mice with chronically elevated baseline corticosterone concentrations. *Behav. Genet.* **39**, 192-201.
Malisch, J. L., Kelly, S. A., Bhanvadia, A., Blank, K. M., Marsik, R. L., Platzer, E. G. and Garland, T., Jr (2009b). Lines of mice with chronically elevated baseline corticosterone levels are more susceptible to a parasitic nematode infection. *Zoology* **112**, 316-324.
Martin L. B. (2009). Stress and immunity in wild vertebrates: timing is everything. *Gen. Comp. Endocrinol.* **163**, 70-76.
Martin, L. B., Weil, Z. M. and Nelson, R. J. (2006). Refining approaches and diversifying directions in ecoimmunology. *Integr. Comp. Biol.* **46**, 1030-1039.
Martin, L. B., Weil, Z. M. and Nelson, R. J. (2008). Seasonal changes in vertebrate immune activity: mediation by physiological trade-offs. *Philos. Trans. R. Soc. Lond. B* **363**, 321-339.
Matson, K. D., Cohen, A. A., Klasing, K. C., Ricklefs, R. E. and Scheuerlein, A. (2006). No simple answers for ecological immunology: relationships among immune indices at the individual level break down at the species level in waterfowl. *Proc. R. Soc. Lond. B* **273**, 815-822.
McEwen, B. and Wingfield, J. C. (2003). The concept of allostasis in biology and biomedicine. *Horm. Behav.* **43**, 2-15.
Medzhitov, R. (2007). Recognition of microorganisms and activation of the immune response. *Nature* **449**, 819-826.
Meek, T. H., Lonquich, B. P., Hannon, R. M. and Garland, T., Jr (2009). Endurance capacity of mice selectively bred for high voluntary wheel running. *J. Exp. Biol.* **212**, 2908-2917.
Moberg, G. (2000). Biological response to stress: implications for animal welfare. In *The Biology of Animal Stress: Basic Principles and Implications for Animal Welfare* (ed. G. Moberg and J. Mench), pp. 1-21. New York: CABI Publishing.
Munck, A. and Náray-Fejes-Tóth, A. (1992). The ups and downs of glucocorticoid physiology permissive and suppressive effects revisited *Mol. Cell. Endocrinol.* **90**, C1-C4.
Norris, K. and Evans, M. R. (2000). Ecological immunology: life history trade-offs and immune defense in birds. *Behav. Ecol.* **11**, 19-26.
Parkin, J. and Cohen, B. (2001). An overview of the immune system. *Lancet* **357**, 1777-1789.
Pfeffer, K., Matsuyama, T., Kundig, T. M., Wakeham, A., Kishihara, K., Shahinian, A., Wiegmann, K., Ohashi, P. S., Kronke, M. and Mak, T. W. (1993). Mice deficient for the 55 kd tumor necrosis factor receptor are resistant to endotoxic shock, yet succumb to L. monocytogenes infection. *Cell* **73**, 457-467.
Potvin, C. and Roff, D. A. (1993). Distribution-free and robust statisitcal methods: viable alternatives to parametric statisitcs. *Ecology* **74**, 1617-1628.
Pough, F. H. (1980). Advantages of ectothermy for tetrapods. *Am. Nat.* **115**, 92-112.
Qureshi, S. T., Lariviere, L., Leveque, G., Clermont, S., Moore, K. J., Gros, P. and Malo, D. (1999). Endotoxin-tolerant mice have mutations in toll-like receptor 4 (Tlr4). *J. Exp. Med.* **189**, 615-625.
Raberg, L., Grahn, M., Hasselquist, D. and Svensson, E. (1998). On the adaptive significance of stress-induced immunosuppression. *Proc. R. Soc. Lond. B* **265**, 1637-1641.
Rezende, E. L., Garland, T., Jr, Chappell, M. A., Malisch, J. L. and Gomes, F. R. (2006). Maximum aerobic performance in lines of *Mus* selected for high wheel-running activity: effects of selection, oxygen availability and the mini-muscle phenotype. *J. Exp. Biol.* **209**, 115-127.
Rhodes, J. S., Gammie, S. C. and Garland, T., Jr (2005). Neurobiology of mice selected for high voluntary wheel-running activity. *Integr. Comp. Biol.* **45**, 438-455.
Ricklefs, R. E. and Wikelski, M. (2002). The physiology/life-history nexus. *Trends Ecol. Evol.* **17**, 462-468.
Robinson, W. D., Hau, M., Klasing, K. C., Wikelski, M., Brawn, J. D., Austin, S. H., Tarwater, C. and Ricklefs, R. E. (2010). Diversification of life histories in new world birds. *Auk* **127**, 253-262.
Roff, D. A. (2002). *Life History Evolution*. Sunderland, MA: Sinauer Associates.
Romero, M. L. (2004) Physiological stress in ecology: lessons from biomedical research. *Trends. Ecol. Evol.* **19**, 249-255.
Sapolsky, R. M., Romero, L. M. and Munck, A. U. (2000). How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. *Endocr. Rev.* **21**, 55-89.
Sibly, R. M. and Calow, P. (1986). *Physiological Ecology of Animals: An Evolutionary Approach*. Oxford: Blackwell Scientific Publications.
Stearns, S. C. (1989). Trade-offs in life-history evolution. *Funct. Ecol.* **3**, 259-268.
Stearns, S. C. (1992). *The Evolution of Life Histories*. New York: Oxford University Press.
Stranahan, A. M., Lee, K. A. and Mattson, M. P. (2008). Central mechanisms of HPA axis regulation by voluntary exercise. *Neuromol. Med.* **10**, 118-127.
Suffredini, A. F., Fantuzzi, G., Badolato, R., Oppenheim, J. J. and O'Grady, N. P. (1999). New insights into the biology of the acute phase response. *J. Clin. Immunol.* **19**, 203-214.
Swallow, J. G., Carter, P. A. and Garland, T., Jr (1998). Artificial selection for increased wheel-running behavior in house mice. *Behav. Genet.* **28**, 227-237.
Swallow, J. G., Rhodes, J. S. and Garland, T., Jr (2005). Phenotypic and evolutionary plasticity of organ masses in response to voluntary exercise in house mice. *Integr. Comp. Biol.* **45**, 426-437.
Syme, D. A., Evashuk, K., Grintuch, B., Rezende, E. L. and Garland, T., Jr (2005). Contractile abilities of normal and 'mini' triceps surae muscles from mice (*Mus domesticus*) selectively bred for high voluntary wheel running. *J. Appl. Physiol.* **99**, 1308-1316.
Tieleman, B. I., Williams, J. B., Ricklefs, R. E. and Klasing, K. C. (2005). Constitutive innate immunity is a component of the pace-of-life syndrome in tropical birds. *Proc. R. Soc. Lond. B* **272**, 1715-1720.
Wahlsten, D. (1990). Insensitivity of the analysis of variance to heredity-environment interaction. *Behav. Brain Sci.* **13**, 109-161.
Wang, Q. Z., Jacobs, J., Deleo, J., Kruszyna, H., Kruszyna, R., Smith, R. and Wilcox, D. (1991). Nitric oxide hemoglobin in mice and rats in endotoxic shock. *Life Sci.* **49**, PL55-PL60.
Wingfield, J. C. and Kitaysky, A. S. (2002). Endocrine responses to unpredictable environmental events: stress or anti-stress hormones? *Integr. Comp. Biol.* **42**, 600-609.
Wingfield, J. C., Maney, D. L., Breuner, C. W., Jacobs, J. D., Lynn, S., Ramenofsky, M. and Richardson, R. D. (1998). Ecological bases of hormone-behavior interactions: the 'emergency life history stage'. *Am. Zool.* **38**, 191-206.
Xing, Z., Gauldie, J., Cox, G., Baumann, H., Jordana, M., Lei, X. F. and Achong, M. K. (1998). IL-6 is an antiinflammatory cytokine required for controlling local or systemic acute inflammatory responses. *J. Clin. Invest.* **101**, 311-320.
Zanetti, G., Heumann, D., Gerain, J., Kohler, J., Abbet, P., Barras, C., Lucas, R., Glauser, M. P. and Baumgartner, J. D. (1992). Cytokine production after intravenous or peritoneal gram-negative bacterial challenge in mice. Comparative protective efficacy of antibodies to tumor necrosis factor-alpha and to lipopolysaccharide. *J. Immunol.* **148**, 1890-1897.
Zera, A. J. and Harshman, L. G. (2001). The physiology of life history trade-offs in animals. *Annu. Rev. Ecol. Syst.* **32**, 95-126.
Zuk, M. and Stoehr, A. M. (2002). Immune defense and host life history. *Am. Nat.* **160**, S9-S22.

Figure S1.

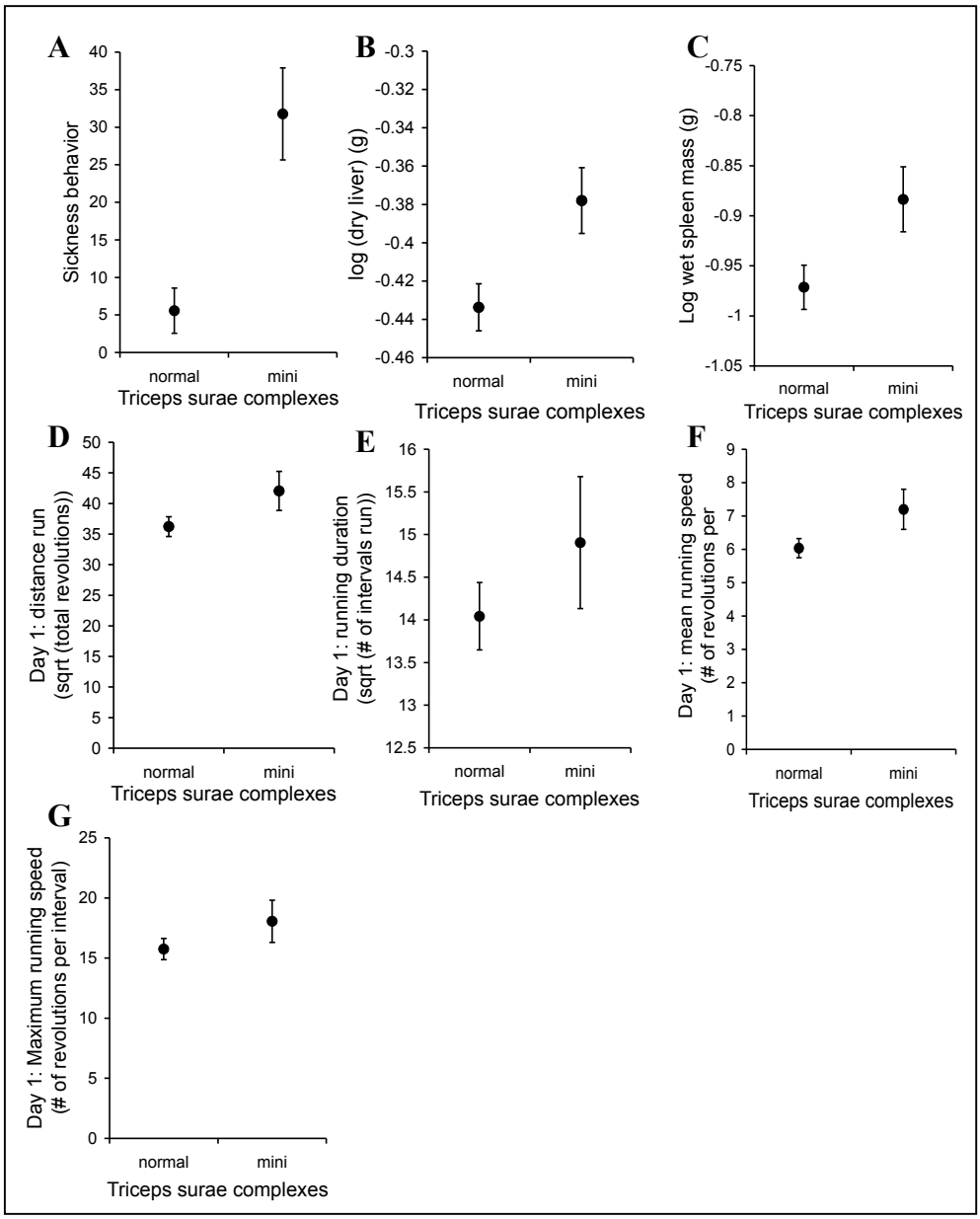
A
Sickness behavior
40
35
30
25
20
15
10
5
0
normal
mini
Triceps surae complexes
B
log (dry liver) (g)
-0.3
-0.32
-0.34
-0.36
-0.38
-0.4
-0.42
-0.44
-0.46
normal
mini
Triceps surae complexes
C
Log wet spleen mass (g)
-0.75
-0.8
-0.85
-0.9
-0.95
-1
-1.05
normal
mini
Triceps surae complexes
D
Day 1: distance run
(sqrt (total revolutions))
50
45
40
35
30
25
20
15
10
5
0
normal
mini
Triceps surae complexes
E
Day 1: running duration
(sqrt (# of intervals run))
16
15.5
15
14.5
14
13.5
13
12.5
normal
mini
Triceps surae complexes
F
Day 1: mean running speed
(# of revolutions per
9
8
7
6
5
4
3
2
1
0
normal
mini
Triceps surae complexes
G
Day 1: Maximum running speed
(# of revolutions per interval)
25
20
15
10
5
0
normal
mini
Triceps surae complexes

Table S1. Summary of response variables for the LPS experiment

Response variable	*N*	Effect	d.f. (Num, Den)	*F*	*P*
Sickness behaviour	136	Innate treatment	2, 12	47.31	<0.001
		Linetype	1, 6	3.23	0.122
		Innate treatment x linetype	2, 12	0.40	0.682
		Mini	1, 106	14.62	<0.001 (+)
log (Initial mass)	136	Innate treatment	2, 12	0.36	0.707
		Linetype	1, 6	1.49	0.268
		Innate treatment x linetype	2, 12	1.89	0.194
		Mini	1, 106	1.11	0.294 (–)
log (Final mass)	136	Innate treatment	2, 12	13.68	<0.001
		Linetype	1, 6	0.61	0.465
		Innate treatment x linetype	2, 12	0.32	0.730
		Mini	1, 106	1.10	0.298 (–)
Change in mass	136	Innate treatment	2, 12	48.74	<0.001
		Linetype	1, 6	5.33	0.060
		Innate treatment x linetype	2, 12	3.22	0.076
		Mini	1, 106	1.24	0.268 (–)
Change in haematocrit	136	Innate treatment	2, 12	26.84	<0.001
		Linetype	1, 6	1.76	0.233
		Innate treatment x linetype	2, 12	0.07	0.932
		Mini	1, 106	2.01	0.159 (+)
Haematocrit – 2 h after injection	136	Innate treatment	2, 12	35.05	<0.001
		Linetype	1, 6	0.06	0.812
		Innate treatment x linetype	2, 12	1.11	0.361
		mini	1, 106	0.79	0.375 (–)
Hematocrit – after wheels	136	Innate treatment	2, 12	0.74	0.497
		Linetype	1, 6	0.23	0.648
		Innate treatment x linetype	2, 12	0.57	0.581
		Mini	1, 106	2.17	0.144 (+)
log (Dry liver mass)	136	Innate treatment	2, 12	22.24	<0.001
		Linetype	1, 6	0.85	0.391
		Innate treatment x linetype	2, 12	0.74	0.497
		Log (final body mass)	1, 105	208.20	<0.001 (+)
		Mini	1, 105	13.12	<0.001 (+)
log (Wet spleen mass)	136	Innate treatment	2, 12	117.6	<0.001
		Linetype	1, 6	0.35	0.576
		Innate treatment x linetype	2, 12	2.89	0.094
		Log (final body mass)	1, 105	58.80	<0.001 (+)
		Mini	1, 105	8.14	0.005 (+)
Rank TNF-α	134	Innate treatment	2, 12	170.33	<0.001
		Linetype	1, 6	1.51	0.265
		Innate treatment x linetype	2, 12	0.23	0.799
		Mini	1, 104	0.57	0.452 (+)
Rank GM-CSF	134	Innate treatment	2, 12	74.97	<0.001
		Linetype	1, 6	0.17	0.693
		Innate treatment x linetype	2, 12	0.31	0.742
		Mini	1, 104	0.10	0.754 (+)
Rank IL-1β	134	Innate treatment	2, 12	99.03	<0.001
		Linetype	1, 6	1.10	0.335
		Innate treatment x linetype	2, 12	0.35	0.710
		Mini	1, 104	1.45	0.231 (–)
Rank IL-6	134	Innate treatment	2, 12	301.36	<0.001
		Linetype	1, 6	1.09	0.337
		Innate treatment x linetype	2, 12	0.47	0.637
		Mini	1, 104	0.87	0.353 (+)
Day 1, Sqrt (distance run)	135	Innate treatment	2, 12	91.21	<0.001
		Linetype	1, 6	17.76	0.006
		Innate treatment x linetype	2, 12	2.91	0.093
		Mini	1, 104	2.67	0.105 (+)
		Sqrt(freeness)	1, 104	1.19	0.279 (–)
Day 2, Sqrt (distance run)	135	Innate treatment	2, 12	21.26	0.001
		Linetype	1, 6	21.71	0.004
		Innate treatment x linetype	2, 12	0.01	0.989
		Mini	1, 104	0.01	0.930 (+)
		Sqrt(freeness)	1, 104	1.55	0.216 (–)
Day 1, Sqrt (running duration)	135	Innate treatment	2, 12	159.94	<0.001
		Linetype	1, 6	5.31	0.061
		Innate treatment x linetype	2, 12	4.87	0.028
		Mini	1, 104	1.02	0.315 (+)
		Sqrt(freeness)	1, 104	0.01	0.923 (+)
Day 2, Sqrt (running duration)	135	Innate treatment	2, 12	30.70	<0.001
		Linetype	1, 6	2.95	0.137
		Innate treatment x linetype	2, 12	0.26	0.774
		Mini	1, 104	0.01	0.930 (+)

		Sqrt(freeness)	1, 104	0.27	0.602 (–)
Day 1,	135	Innate treatment	2, 12	20.89	<0.001
Sqrt (mean speed)		Linetype	1, 6	31.58	0.001
		Innate treatment x linetype	2, 12	0.66	0.536
		Mini	1, 104	2.90	0.092 (+)
		Sqrt(freeness)	1, 104	4.87	0.030 (–)
Day 2,	135	Innate treatment	2, 12	34.67	<0.001
Sqrt (mean speed)		Linetype	1, 6	36.21	0.001
		Innate treatment x linetype	2, 12	1.73	0.219
		Mini	1, 104	2.90	0.092 (+)
		Sqrt(freeness)	1, 104	3.25	0.075 (–)
Day 1,	135	Innate treatment	2, 12	7.81	0.007
Maximum speed		Linetype	1, 6	21.69	0.004
		Innate treatment x linetype	2, 12	1.88	0.195
		Mini	1, 104	1.39	0.241 (+)
		Sqrt(freeness)	1, 104	0.03	0.854 (–)
Day 2,	135	Innate treatment	2, 12	4.40	0.037
Maximum speed		Linetype	1, 6	36.21	<0.001
		Innate treatment x linetype	2, 12	0.16	0.853
		Mini	1, 104	0.09	0.767 (+)
		Sqrt(freeness)	1, 104	2.00	0.161 (–)

Statistics were run in SAS using Proc Mixed. Fixed effects that are significant at α=0.05 are bold**.** log-transformed body mass was included as a covariate in the analyses for organ masses. Wheel freeness was included in analyses for wheel running variables. Sign after *P*-values for mini-muscle indicates direction of mini-muscles effect (e.g. + indicates that the mini-muscle mice have a higher mean than normal-muscle mice).

N, sample size; d.f., degrees of freedom; mini, dummy variable that indicates whether the mice have mini-muscle; sqrt, square-root transformed.

Table S2. Results from the analysis examining how rank-TNFα varied within each LPS treatment

LPS treatment	*N*	Effect	d.f. (Num, Den)	*F*	*P*
LPS	45	Linetype	1,6	1.26	0.304
		Mini	1,31	0.11	0.742 (–)
Sham	44	Linetype	1,5	0.04	0.846
		Mini	1,31	3.00	0.093 (+)
Baseline	45	Linetype	1,6	1.01	0.353
		Mini	1,31	0.67	0.418 (–)

Statistics were run in SAS using Proc Mixed. Sign after *P*-values for mini-muscle indicates direction of mini-muscle effect (e.g. + indicates that the mini-muscle mice have a higher mean than normal muscle mice).

N, sample size; d.f., degrees of freedom; mini, dummy variable that indicates whether the mice have mini-muscle.