The Journal of Experimental Biology 209, 115-127
Published by The Company of Biologists 2006
doi:10.1242/jeb.01883

Maximum aerobic performance in lines of *Mus* selected for high wheel-running activity: effects of selection, oxygen availability and the mini-muscle phenotype

Enrico L. Rezende*, Theodore Garland, Jr, Mark A. Chappell, Jessica L. Malisch and Fernando R. Gomes

Department of Biology, University of California, Riverside, CA 92521, USA

*Author for correspondence at present address: Integrative Ecology Group, Estación Biológica de Doñana, CSIC, Apdo. 1056, E-41080 Seville, Spain (e-mail: enrico.rezende@ebd.csic.es)

Accepted 10 September 2005

Summary

We compared maximum aerobic capacity during forced exercise ($\dot{V}_{O2max}$) in hypoxia (P_{O_2}=14% O_2), normoxia (21%) and hyperoxia (30%) of lines of house mice selectively bred for high voluntary wheel running (S lines) with their four unselected control (C) lines. We also tested for pleiotropic effects of the 'mighty mini-muscle' allele, a Mendelian recessive that causes a 50% reduction in hind limb muscle but a doubling of mass-specific aerobic enzyme activity, among other pleiotropic effects. $\dot{V}_{O2max}$ of female mice was measured during forced exercise on a motorized treadmill enclosed in a metabolic chamber that allowed altered P_{O_2}. Individual variation in $\dot{V}_{O2max}$ was highly repeatable within each P_{O_2}, and values were also significantly correlated across P_{O_2}. Analysis of covariance showed that S mice had higher body-mass-adjusted $\dot{V}_{O2max}$ than C at all P_{O_2}, ranging from +10.7% in hypoxia to +20.8% in hyperoxia. $\dot{V}_{O2max}$ of S lines increased practically linearly with P_{O_2}, whereas that of C lines plateaued from normoxia to hyperoxia, and respiratory exchange ratio (=CO_2 production/$\dot{V}_{O2max}$) was lower for S lines. These results suggest that the physiological underpinnings of $\dot{V}_{O2max}$ differ between the S and C lines. Apparently, at least in S lines, peripheral tissues may sustain higher rates of oxidative metabolism if central organs provide more O_2. Although the existence of central limitations in S lines cannot be excluded based solely on the present data, we have previously reported that both S and C lines can attain considerably higher $\dot{V}_{O2max}$ during cold exposure in a He–O_2 atmosphere, suggesting that limitations on $\dot{V}_{O2max}$ depend on interactions between the central and peripheral organs involved. In addition, mini-muscle individuals had higher $\dot{V}_{O2max}$ than did those with normal muscles, suggesting that the former might have higher hypoxia tolerance. This would imply that the mini-muscle phenotype could be a good model to test how exercise performance and hypoxia tolerance could evolve in a correlated fashion, as previous researchers have suggested.

Key words: artificial selection, central limitation, exercise, hypoxia, hyperoxia, maximum metabolic rate, oxygen availability, peripheral limitation, respiratory exchange ratio, symmorphosis.

Introduction

The vertebrate O_2 transport and delivery system can be considered to be composed of four steps in series: ventilatory convection, alveolar–capillary diffusion, blood convection, and tissue capillary-to-cell diffusion of O_2 (Taylor and Weibel, 1981; Wagner, 1996). How whole-animal maximum aerobic capacity ($\dot{V}_{O2max}$) should respond to changes in ambient partial pressure of O_2 (P_{O_2}) depends on which of these steps is limiting and on mitochondrial oxidative capacity in peripheral tissues. For instance, hypoxia should cause a reduction in $\dot{V}_{O2max}$ unless both O_2 delivery and mitochondrial oxidative capacity were in excess. In this context, Lindstedt et al. (1988) have shown that when O_2 delivery is reduced by anemia or hypoxia, $\dot{V}_{O2max}$ declines in direct proportion to O_2 delivery, whereas O_2 extraction from the blood remains relatively constant (~90%). Effects of hyperoxia on $\dot{V}_{O2max}$ are not as straightforward, because limitations at different levels in the O_2 cascade may lead to different experimental outcomes (Richardson et al., 1999; Lindstedt and Conley, 2001; Noakes et al., 2001). Increased $\dot{V}_{O2max}$ in hyperoxia may suggest that O_2 uptake and delivery systems are more relevant for $\dot{V}_{O2max}$ in normoxia than mitochondrial oxidative capacity (i.e. central limitation), whereas potential constraints at the mitochondrial level in the muscles might be the case if $\dot{V}_{O2max}$ remains unchanged at higher P_{O_2} (i.e. peripheral limitation). [However, limits in cardiac output or blood O_2 saturation in the latter example could not be ruled out, as convective steps in the O_2 cascade (e.g. O_2 movement into the lungs, blood O_2 transport, muscle blood flow) might impose a ceiling in $\dot{V}_{O2max}$ regardless of the higher diffusion consequent of the increase O_2 gradient.]

Alternatively, it could be the case that all steps in the O_2 cascade are virtually identical in capacity, such that none, by itself, would set the limit for maximum organismal performance. Indeed, symmorphosis (Taylor and Weibel, 1981) has been proposed as a response at many levels of biological organization to the '*powerful optimizing process*' (Alexander, 1989, p. 1200) attributed to natural selection. Whether animals are symmorphotic presently remains a matter of debate among physiologists (Hammond and Diamond, 1997; Bacigalupe and Bozinovic, 2002), and several authors have argued against its biological relevance in the light of evolutionary theory and various empirical results (Garland and Huey, 1987; Garland, 1998; Gordon, 1998; brief review in Suarez and Darveau, 2005). Almost all existing studies aimed at testing symmorphosis have been interspecific comparisons, so experimental evolution can offer a novel approach to the study of such hypotheses about the correlated evolution of complex phenotypes (Garland, 2001, 2003; Swallow and Garland, 2005; Garland and Carter, 1994).

Artificial selection for high voluntary wheel running in four replicate lines of house mice (Swallow et al., 1998a) provides an opportunity to examine interactions between aerobic limits and locomotor activity. Existing evidence indicates that normoxic $\dot{V}_{O2max}$ has increased, especially in males, coincident with the evolution of higher activity levels in the selected (S) lines as compared with their random bred control (C) lines (Swallow et al., 1998b; Rezende et al., 2005, 2006). Therefore, selection might have affected hypoxia tolerance at the muscle level because these lines run voluntarily at speeds approaching their maximum aerobic speeds for at least some minutes per night (Girard et al., 2001; Rezende et al., 2005, 2006).

An unexpected part of the response to selection has been the evolution of the 'mighty mini-muscle' phenotype, which represents homozygotes for a Mendelian recessive allele that halves gastrocnemius muscle mass (Garland et al., 2002; Belter et al., 2004) while doubling per gram aerobic capacity (Houle-Leroy et al., 2003). This phenotype occurs in only two of the four selected lines and has now gone to fixation in one of them. Apparently, the other two selected lines lost the allele, which was initially rare, by random genetic drift during the early generations of the experiment (Garland et al., 2002). In any case, the increase in frequency of this gene of major phenotypic effect represents an important component of the overall response to selective breeding. Moreover, because it has significant effects on a variety of other traits in addition to gastrocnemius muscle mass (Garland et al., 2002; Houle-Leroy et al., 2003; Swallow et al., 2005; Syme et al., 2005; Kelly et al., in press), it should be accounted for in statistical analyses. For example, the soleus muscle of mini-muscle individuals is actually larger than in normal individuals (Syme et al., 2005). And, considering that mini-muscle individuals often run both more and faster on wheels as compared with normal-muscled individuals within the same line (Syme et al., 2005; Kelly et al., in press), one might expect that the phenotype could also affect $\dot{V}_{O2max}$.

The purpose of this study was to determine the effects of different atmospheric P_{O_2} on $\dot{V}_{O2max}$ during forced exercise in the S and C lines of mice and to test for pleiotropic effects of the mini-muscle phenotype. Results can be important for several reasons. First, they may provide insight about the physiology underlying $\dot{V}_{O2max}$ and the relative importance of O_2 uptake and delivery systems at the whole-individual level. Second, given that S lines have elevated $\dot{V}_{O2max}$ as compared with C lines (Swallow et al., 1998b; Rezende et al., 2005, 2006), hypoxia and hyperoxia may affect S and C lines differently, which would suggest differences at lower physiological levels in the O_2 cascade, as has been found in lines of rats selected for high *vs* low treadmill performance (e.g. Henderson et al., 2002; Howlett et al., 2003). Such results would be relevant for symmorphosis. Specifically, if S lines have evolved in a symmorphotic fashion, one would expect effects of P_{O_2} on $\dot{V}_{O2max}$ to be independent of selection history, and differences between S and C would remain constant across different P_{O_2}. Conversely, if S and C lines respond differently to changes in P_{O_2}, this would suggest that traits subordinate to $\dot{V}_{O2max}$ have evolved in a 'non-symmorphotic' fashion in response to selection. Third, effects of hyperoxia on $\dot{V}_{O2max}$ and running performance could reveal whether aerobic constraints may be limiting to the evolution of even higher levels of wheel running in the S lines. An increased performance by S mice in a hyperoxic atmosphere would suggest that they might run even more (faster) on wheels if O_2 availability at the tissue level were higher. Finally, results may help elucidate why the mini-muscle phenotype has been favored by the selection regimen (Garland et al., 2002; Houle-Leroy et al., 2003; Swallow et al., 2005; Syme et al., 2005; Kelly et al., in press).

Materials and methods

Animals and experimental protocol

We used 59 females from generation 36 post-selection of a selection experiment for high voluntary wheel running, as described in detail elsewhere (Swallow et al., 1998a; Garland, 2003). Briefly, voluntary running performance was measured on wheels every generation in eight independent lines of laboratory mice (~10 families per line each generation). After founding of the initial colony, four of the lines were randomly assigned to the selection treatment (lab designations 3, 6, 7 and 8) and the other four were treated as controls (lines 1, 2, 4 and 5). In the four selected (S) lines, the best male and female runners within each family were used as breeders for the next generation (and second-highest runners as additional extra breeders, if necessary). The selection criterion was running distance during the last 2 days of a 6-day trial. In the four control (C) lines, breeders were randomly chosen at weaning (21 days of age). Selection was interrupted for four generations (32–35) to transfer the colony from the University of Wisconsin-Madison to California. For generation 36, a sample of which was studied here, the average difference between S and C mice in wheel running on days 5+6 of the 6-day trial

was 2.37-fold (averaged over all eight lines and both sexes; T. Garland, Jr., unpublished results; $P_{selection}<0.001$).

After breeders for the next generation were selected, we obtained our individual females from different families from within each line. We used six females per line, except for line 6, which is polymorphic for the mini-muscle allele mentioned above (sibs were sometimes used in this line; see Statistics), where we used 17 females. Note that line 3 is now apparently fixed for the mini-muscle allele (i.e. all individuals have the phenotype; Syme et al., 2005; T. Garland, Jr., unpublished results). Prior to choosing these individuals for study, we excluded the lowest runner within each family to compensate for the bias caused by having chosen the highest runners to become breeders during the selection protocol (see previous paragraph). Mice were housed four per cage from weaning to selection, with food and water *ad libitum*, then measured for 6 days on wheels, beginning at around 62 days of age (range 54–70 days), and subsequently were maintained in individual cages.

At approximately 95 days of age (79–110 days), individuals were randomly assigned to five measurement batches of 12 mice each (although within each line, older individuals were assigned to initial batches to minimize differences in age). Each batch was used for treadmill measurements at three different P_{O_2} (see below) for six consecutive days – measurements were performed twice on consecutive days at each P_{O_2}. To control for training or 'acclimating' effects, batches were measured in each atmosphere in different sequences, and individuals were measured in random order within each batch. All trials were performed between 09.00 and 12.00 h. On day 7, mice were sacrificed and dissected, and individuals with the mini-muscle phenotype were identified by inspection of graphs of muscle mass *versus* body mass (see Garland et al., 2002; Houle-Leroy et al., 2003; Belter et al., 2004; Syme et al., 2005).

Treadmill measurements in hypoxia, normoxia and hyperoxia

Maximum aerobic capacity during forced exercise ($\dot{V}_{O_2max}$) was estimated with open-flow respirometry by running mice in an enclosed motorized treadmill, as described previously (Hayes and Chappell, 1990; Chappell et al., 2003; Rezende et al., 2005). The treadmill had an inclination of 25°, which maximizes the $\dot{V}_{O_2max}$ values obtained in mice (Kemi et al., 2002) and resulted in significantly higher $\dot{V}_{O_2max}$ in males from generation 33 when compared with males tested without slopes (E.L.R., T.G., F.R.G. and M.A.C., unpublished data). Treadmill tests were performed at room temperature (22–25°C). We used positive-pressure, flow-through respirometry to determine rates of oxygen consumption ($\dot{V}_{O_2}$) and carbon dioxide production ($\dot{V}_{CO_2}$). Airflow was regulated with upstream mass flow controllers (Applied Materials, Tylan, Billerica, MA, USA), with a flow rate of 2100 ml min^{-1}, that maintained excurrent O_2 concentrations within 0.5% of incurrent gas. Approximately 100 ml min^{-1} of excurrent air was subsampled and analyzed for CO_2 (which was then scrubbed) and O_2. Oxygen concentration was measured with an Ametek/Applied Electrochemistry S-3A analyzer (Pittsburgh, PA, USA), and CO_2 was measured with LiCor 6251. Data from gas analyzers and other instruments were recorded on Macintosh computers equipped with National Instruments A-D converters and 'Labhelper' software (Warthog Systems, www.warthog.ucr.edu).

To maintain different atmospheric O_2 concentrations, flow controllers were supplied with room air or connected to cylinders with different mixtures of oxygen and nitrogen creating hypoxia (14% O_2) and hyperoxia (30% O_2). Reference readings of incurrent gas were obtained at the start and end of measurements. References from cylinders were always checked against room air before measurements and remained precise (±0.1%) throughout the experiment. Mice were placed inside the treadmill's working section (6 cm wide, 7 cm high, 13.5 cm long), which was flushed for ~1 min to attain stable O_2 concentration during hypoxia and hyperoxia (for consistency, this was also applied in normoxic trials). Mice were allowed a 1–2 min adjustment period and were then run at increasing speeds starting at 0.15–0.2 m s^{-1} and raised in step increments of ~0.15 m s^{-1} every 45 s. A trial was terminated when mice could no longer keep pace with the treadmill and $\dot{V}_{O_2}$ did not increase with increasing tread speed. All animals failed to maintain position at the highest tread speeds implemented. Trial quality was also assessed using a subjective scale (five categories, from 'poor' to 'excellent'; following Swallow et al., 1998b; Rezende et al., 2005). In this sample of mice, no individual was scored as 'poor' and hence none were excluded from analyses. Preliminary analyses using trial quality as an additional covariate indicated that it was never statistically significant, so it was not included in final models.

We applied the 'instantaneous' transformation (Bartholomew et al., 1981) to resolve rapid changes in metabolism. Effective volume of the treadmill was 903 ml (Chappell et al., 2004). We calculated $\dot{V}_{O_2}$ (ml min^{-1}) as in Eqn 1, after scrubbing subsampled air of water vapor and CO_2 (Drierite and soda lime, respectively):

$$\dot{V}_{O_2} = \dot{V}\,(F_{I_{O_2}} - F_{E_{O_2}})/(1 - F_{E_{O_2}}), \quad (1)$$

where $\dot{V}$ is flow rate (ml min^{-1} STP; standard temperature and pressure), and $F_{I_{O_2}}$ and $F_{E_{O_2}}$ are the fractional O_2 concentrations in incurrent and excurrent air, respectively ($F_{I_{O_2}}$ was 0.3002, 0.2095 and 0.1405 for hyperoxia, normoxia and hypoxia; $F_{E_{O_2}}$ was always within 0.5% of $F_{I_{O_2}}$). $\dot{V}_{CO_2}$ (ml min^{-1}) was calculated as:

$$\dot{V}_{CO_2} = \{\dot{V}\,[(F_{E_{CO_2}} - F_{I_{CO_2}}) - F_{E_{CO_2}}]\,\dot{V}_{O_2}\} / (1 - F_{E_{CO_2}}), \quad (2)$$

where $F_{I_{CO_2}}$ and $F_{E_{CO_2}}$ are the fractional CO_2 concentrations of incurrent and excurrent gas, respectively ($F_{I_{CO_2}}$ was ~0.0004 in normoxia and 0 in hypoxia and hyperoxia). We determined $\dot{V}_{O_2max}$ and $\dot{V}_{CO_2max}$ as the highest 60-s continuous average in the trial, and calculated the respiratory exchange ratio (RER=$\dot{V}_{CO_2}/\dot{V}_{O_2}$) at $\dot{V}_{O_2max}$ and $\dot{V}_{CO_2max}$. Maximum running speed attained in each trial was measured with a calibrated tachometer attached to the treadmill's motor drive. All

calculations and corrections on $\dot{V}_{O_2}$ files were performed using LabAnalyst software (Warthog Systems).

Statistics

Analyses were performed with SPSS for Windows and SAS PROC MIXED (version 8, SAS Institute, Cary, NC, USA, 1996). Repeatability between traits, measured on two consecutive days, was estimated with Pearson product-moment correlations (Hayes and Jenkins, 1997) on residuals from nested analysis of covariance (ANCOVA), correcting for selection history (line type and line), mass, batch and age (see below). To check for day-to-day 'training' effects, we used paired *t*-tests comparing trials 1 and 2.

As shown below, individual performance in the first and second trials within a given P_{O_2} was highly repeatable, consistent with maximal effort. Because we were interested in maximum performance, we selected the higher of the two $\dot{V}_{O_2max}$ recorded at each P_{O_2} for most analyses, consistent with previous studies (e.g. Swallow et al., 1998b; Rezende et al., 2005). Differences between line types (S *versus* C) were assessed with type III sums of squares, using linear mixed models. Line type was the grouping variable, replicate lines (N=8) were nested within line type as a random factor, mass and age were covariates, and batch was a cofactor. Line effects were determined by comparing maximum-likelihood (ML) estimates with and without lines in the model (the difference between ML values follows a χ^2 distribution with 1 d.f.). Least-squares means adjusted to a common age and body mass were calculated to estimate the difference between S and C lines. Regular ANCOVAs were also performed for S and C separately, including line as a random factor (four lines) and using the same covariates of the nested ANCOVA. Discriminating effects of the mini-muscle allele from other effects is not straightforward (see Discussion); hence, we employed two different approaches. First, nested ANCOVAs were performed with an additional dummy variable equal to 0 for normal phenotypes and 1 for mini-muscles (N=48 and 11, respectively). Second, we compared individuals with and without the mini-muscle phenotype within line 6, employing regular ANCOVA with mass and age as covariates.

We assessed how $\dot{V}_{O_2}$ changed as a function of O_2 concentration and whether these changes differed due to selection history, employing general linear models (GLM) for repeated measures (SPSS for Windows). Individuals were experimental units, partial pressure (hypoxia, normoxia and hyperoxia) was the within-subjects factor, and selection history and lines were included as between-subjects factors (the latter, in this case, was treated as a fixed factor). Significance of selection history (S *versus* C) was estimated by dividing mean square error (MSE) of the model including only line type as the fixed factor by MSE from the model with line × line type (i.e. same d.f. obtained in the nested ANOVA). To determine how variables differed between trials, contrasts (differences between successive values for each individual) were compared employing multivariate ANOVAs (test of within-subject contrasts). We also tested for individual effects *across* different P_{O_2} with an F ratio between the MSEs between- and within-individuals (=$MSE_{between}$/MSE_{within}), obtained from a repeated-measures ANOVA with individuals as the grouping factors. When there are only two repeated measures, this test is analogous to Pearson's correlations and can also be used to test for 'consistency' between measurements, i.e. 'repeatability' (Hayes and Jenkins, 1997; Zar, 1999; Rezende et al., 2005).

All P values shown correspond to two-tailed tests, although in many cases one-tailed tests would be appropriate. For example, in all measures of repeatability of individual differences, the alternative hypothesis is for a positive correlation, not just a correlation different from zero (unless fatigue effects were large and carried over for 24 h). Similarly, for $\dot{V}_{O_2max}$, the expectation from both first principles and previous studies (Swallow et al., 1998b; Rezende et al., 2005) is that mice from S lines will have higher values than those from C lines. On the other hand, multiple tests are performed that compare the same set of animals for various traits, which would suggest some sort of correction for multiple comparisons (e.g. see Rice, 1989; Curran-Everett, 2000). Thus, we report test statistics, d.f. and P values for all statistical tests; independent conclusions may be drawn from these as desired.

Results

Maximum rates of O_2 consumption ($\dot{V}_{O_2max}$) and CO_2 production ($\dot{V}_{CO_2max}$) were highly repeatable in hypoxia, normoxia and hyperoxia, regardless of selection history (Table 1). Respiratory exchange ratios obtained at $\dot{V}_{O_2max}$ ($RER_{\dot{V}_{O_2max}}$) were repeatable only in hypoxia and normoxia (P<0.02 in all cases), whereas $RER_{\dot{V}_{CO_2max}}$ were not repeatable in most analyses. Maximum running speeds (treadmill with 25° slope) in S and C lines were always significantly repeatable between days (Table 1). Accordingly, effects of individual variation across P_{O_2} remained statistically significant for $\dot{V}_{O_2max}$, $\dot{V}_{CO_2max}$ and maximum running speeds but not for either of the RER indexes (see below).

Training effects were negligible for $\dot{V}_{O_2max}$, which was 2.2% lower in the first trial in hypoxia (two-tailed P=0.023), and the same was true for $\dot{V}_{CO_2max}$ (P=0.003). No differences between trials 1 and 2 were observed in maximum running speeds at any P_{O_2} (P>0.05 in all cases).

Selection, body mass and mini-muscle effects

Mean (unadjusted) body mass was 26.6±0.3 g (mean ± s.e.m.) for S lines and 30.4±0.5 g for C lines. Unadjusted individual $\dot{V}_{O_2max}$ was 4.888±0.083 ml O_2 min^{-1} and 4.614±0.079 ml O_2 min^{-1} for S and C lines in hypoxia (i.e. 0.184±0.003 ml O_2 min^{-1} g^{-1} and 0.152±0.003 ml O_2 min^{-1} g^{-1} on a per gram basis), 5.662±0.124 ml O_2 min^{-1} and 5.225±0.130 ml O_2 min^{-1} in normoxia (0.212±0.005 ml O_2 min^{-1} g^{-1} and 0.171±0.004 ml O_2 min^{-1} g^{-1}) and 6.104±0.138 ml O_2 min^{-1} and 5.481±0.151 ml O_2 min^{-1} in hyperoxia (0.229±0.005 ml O_2 min^{-1} g^{-1} and 0.182±0.005 ml O_2 min^{-1} g^{-1}).

Table 1. *Repeatability analyses estimating individual consistency of* $\dot{V}_{O2max}$, $\dot{V}_{CO2max}$, *respiratory exchange ratios at* $\dot{V}_{O2max}$ *and at* $\dot{V}_{CO2max}$, *and maximum running speed on the treadmill between days 1 and 2 at each partial pressure of* O_2

	Hypoxia (14% O_2)		Normoxia (21% O_2)		Hyperoxia (30% O_2)	
	r	*P*	*r*	*P*	*r*	*P*
Pooled (*N*=59)						
$\dot{V}_{O2max}$	0.607	**<0.0001**	0.711	**<0.001**	0.587	**<0.0001**
$\dot{V}_{CO2max}$	0.733	**<0.0001**	0.810	**<0.0001**	0.797	**<0.0001**
$RER_{\dot{V}_{O_2max}}$	0.496	**<0.0001**	0.346	**0.004**	0.041	0.379
$RER_{\dot{V}_{CO_2max}}$	0.280	**0.016**	0.057	0.333	0.117	0.188
Maximum speed	0.758	**<0.0001**	0.728[a]	**<0.0001**	0.862[b]	**<0.0001**
Selected (*N*=35)						
$\dot{V}_{O2max}$	0.633	**<0.0001**	0.634	**<0.0001**	0.719	**<0.0001**
$\dot{V}_{CO2max}$	0.733	**<0.0001**	0.824	**<0.0001**	0.840	**<0.0001**
$RER_{\dot{V}_{O_2max}}$	0.443	**0.004**	0.356	**0.018**	0.064	0.356
$RER_{\dot{V}_{CO_2max}}$	0.232	0.090	0.354	**0.018**	0.172	0.162
Maximum speed	0.749	**<0.0001**	0.780[a]	**<0.0001**	0.800[b]	**<0.0001**
Control (*N*=24)						
$\dot{V}_{O2max}$	0.627	**<0.001**	0.720	**<0.0001**	0.476	**0.009**
$\dot{V}_{CO2max}$	0.750	**<0.0001**	0.759	**<0.0001**	0.763	**<0.0001**
$RER_{\dot{V}_{O_2max}}$	0.497	**0.007**	0.476	**0.009**	0.139	0.259
$RER_{\dot{V}_{CO_2max}}$	0.217	0.154	–0.300	n.a.	0.190	0.186
Maximum speed	0.723	**<0.0001**	0.679	**<0.0001**	0.683	**<0.0001**

[a]Two individuals excluded from analyses because maximum speed unavailable.

[b]One individual excluded from analyses because maximum speed unavailable.

Pearson product-moment correlations were computed for residuals from nested ANCOVAs accounting for selection history (line types and/or lines), mass, age and batch effects. *P*-values are for one-tailed hypotheses, and values in bold indicate statistically significant values at $P<0.05$. n.a., direction of correlation opposite to directional hypothesis, so *P* value not reported.

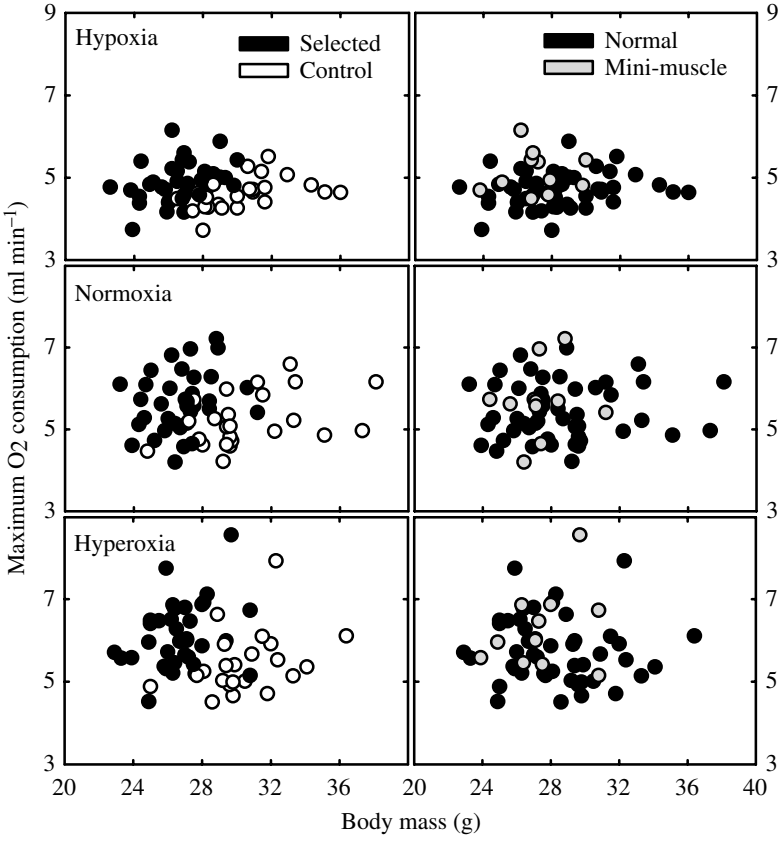

Fig. 1. Maximum aerobic capacity ($\dot{V}_{O2max}$) measured during forced exercise in a hypoxic (14% O_2), normoxic (21% O_2) and hyperoxic atmosphere (30% O_2) for 59 mice from lines selected for high voluntary wheel running (S) and their non-selected control lines (C). Each point represents the highest 60-s average obtained in two measurements on the treadmill performed on consecutive days (see Materials and methods). (Left panels) Aerobic capacity in relation to selection history – i.e. S (closed circles) *versus* C (open circles). Adjusted means for S and C in each P_{O_2} are listed in Table 2. (Right panels) The same graph, but highlighting individuals with the mini-muscle phenotype (all in S lines) *versus* normal phenotypes. Effects of size, selection history and mini-muscle are summarized in Table 2.

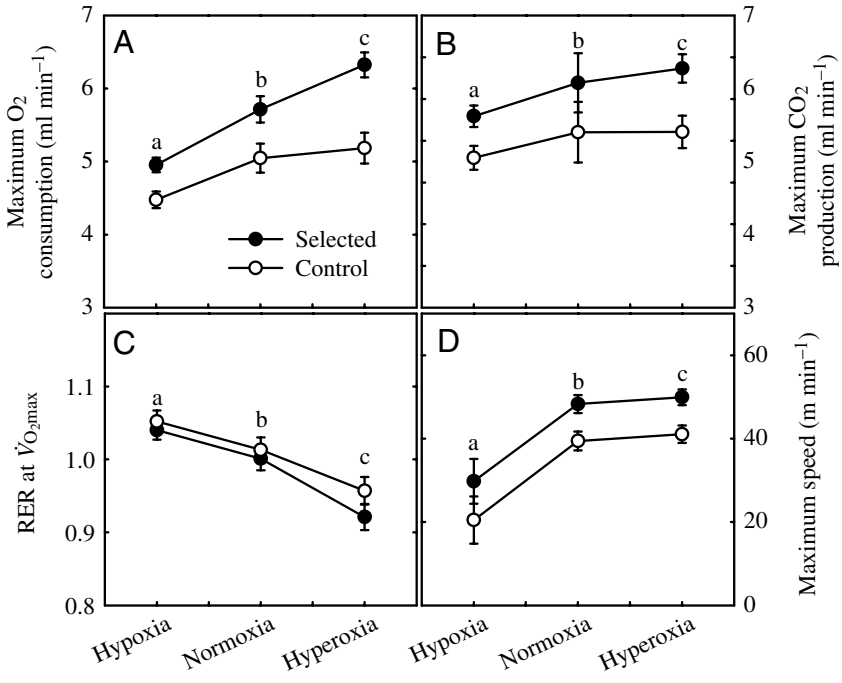

Fig. 2. (A) Maximum aerobic capacity ($\dot{V}_{O_2max}$), (B) carbon dioxide production, (C) respiratory exchange ratio at $\dot{V}_{O_2max}$ (RER=$\dot{V}_{CO_2}/\dot{V}_{O_2max}$) and (D) maximum running speeds, measured in S and C mice (N=35 and 24, respectively) during forced exercise on a treadmill with a hypoxic (14% O_2), normoxic (21% O_2) or hyperoxic atmosphere (30% O_2). Values are shown as adjusted means ± s.e.m. Statistical differences between hypoxic, normoxic and hyperoxic trials are shown in lowercase letters; effects of selection history for each trait at a given P_{O_2} are shown in Table 2; overall effects of selection on the entire dataset (i.e. across P_{O_2}) are listed in Table 4.

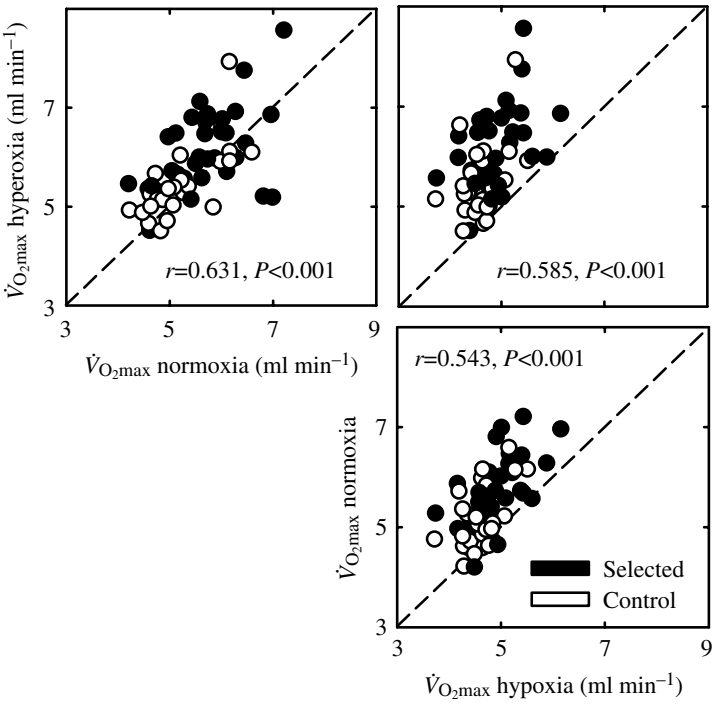

Fig. 3. Individual consistency of maximum aerobic performance ($\dot{V}_{O_2max}$) measured across different P_{O_2}. Points represent the highest 60-s averages for each individual obtained from either of two treadmill trials each in hypoxia (14% O_2), normoxia (21% O_2) and hyperoxia (30% O_2). The broken line represents equality (x=y) in each plot. Pearson product-moment correlations are also reported for residuals calculated from nested ANCOVAs performed separately for each P_{O_2}, including mass and age as covariates and batch as a cofactor.

According to the nested ANCOVA, mice from S lines were significantly smaller in body mass than C mice (adjusted means of 27.6±1.5 g and 31.5±1.5 g, respectively, $P_{selection}$=0.0038; Fig. 1), but there were no differences between normal and mini-muscle phenotypes (full model nested ANCOVA, P_{mini}=0.720). In addition, according to regular ANCOVA within line 6 with age as covariate, body mass did not differ between mini-muscle and normal mice ($F_{1,14}$=0.088, P=0.771).

After accounting for variation in body mass, S lines achieved higher $\dot{V}_{O_2max}$ and $\dot{V}_{CO_2max}$ than C mice (Fig. 2), with the greatest difference in hyperoxia (Table 2). In the full nested model, individuals with mini-muscles had statistically higher $\dot{V}_{O_2max}$ and $\dot{V}_{CO_2max}$ only in hypoxic conditions (Table 2; Fig. 1). Considering only individuals within selection line 6, those with the mini-muscle phenotype achieved statistically higher $\dot{V}_{O_2max}$ and $\dot{V}_{CO_2max}$ in hypoxia (Table 3) but not in normoxia or hyperoxia (P>0.21 in all cases; Table 3).

Maximum running speeds were significantly higher in S mice at all O_2 concentrations (Table 2; Fig. 2). Regardless of atmospheric P_{O_2}, $\dot{V}_{O_2max}$ and $\dot{V}_{CO_2max}$ increased significantly with body mass, whereas maximum running speeds did not (Table 2). Speeds were not affected by mini-muscle status in the full data set (Table 2) or within line 6 (Table 3).

Aerobic performance at different P_{O_2}

Repeated-measures analyses show that increased P_{O_2} led to increased $\dot{V}_{O_2max}$ and $\dot{V}_{CO_2max}$ in all lines,

Table 2. *Adjusted means ± s.e.m. for S and C female mice, with line types tested over replicate lines nested within line types*

	Selected	Control	S/C	$P_{selection}$	$P_{body\ mass}$	P_{mini}
Hypoxia (14% O_2)						
$\dot{V}_{O_2max}$ (ml min^{-1})	4.911±0.092	4.549±0.110	1.080	0.0687	**0.0297**	**0.0311** (+)
$\dot{V}_{O_2max}$ (ml min^{-1})	4.955±0.100	4.477±0.114	1.107	**0.0291**	**0.0231**	
$\dot{V}_{CO_2}$ @ $\dot{V}_{O_2max}$[a]	5.127±0.126	4.746±0.147	1.080	0.1232	**0.0077**	**0.0225** (+)
$\dot{V}_{CO_2}$ @ $\dot{V}_{O_2max}$[a]	5.205±0.132	4.648±0.149	1.120	**0.0435**	**0.0061**	
$\dot{V}_{CO_2max}$ (ml min^{-1})	5.240±0.123	4.875±0.139	1.075	0.1279	**0.0022**	**0.0388** (+)
$\dot{V}_{CO_2max}$ (ml min^{-1})	5.294±0.129	4.795±0.143	1.104	0.0547	**0.0020**	
$RER_{\dot{V}_{O_2max}}$[b]	1.036±0.014	1.057±0.016	0.980	0.3001		0.2363
$RER_{\dot{V}_{O_2max}}$[b]	1.040±0.013	1.052±0.015	0.988	0.5035		
$RER_{\dot{V}_{CO_2max}}$[b]	1.076±0.018	1.117±0.019	0.963	**0.0314**		0.5711
$RER_{\dot{V}_{CO_2max}}$[b]	1.077±0.017	1.115±0.018	0.966	**0.0305**		
Maximum speeds (m min^{-1})	29.77±5.37	20.47±5.67	1.454	**0.0172**	0.4276	0.3941
Maximum speeds (m min^{-1})	30.04±5.34	21.49±5.55	1.398	**0.0188**	0.4616	
Normoxia (21% O_2)						
$\dot{V}_{O_2max}$ (ml min^{-1})	5.709±0.190	5.052±0.211	1.130	0.0777	**0.0419**	0.9193
$\dot{V}_{O_2max}$ (ml min^{-1})	5.715±0.180	5.047±0.198	1.132	0.0588	**0.0397**	
$\dot{V}_{CO_2}$ @ $\dot{V}_{O_2max}$[a]	5.794±0.171	5.002±0.194	1.158	**0.0317**	**0.0046**	0.3793
$\dot{V}_{CO_2}$ @ $\dot{V}_{O_2max}$[a]	5.759±0.167	5.050±0.186	1.140	**0.0387**	**0.0053**	
$\dot{V}_{CO_2max}$ (ml min^{-1})	5.637±0.407	4.964±0.420	1.136	0.0635	**0.0019**	0.3592
$\dot{V}_{CO_2max}$ (ml min^{-1})	5.694±0.354	5.102±0.363	1.116	0.0806	**0.0022**	
$RER_{\dot{V}_{O_2max}}$[b]	1.005±0.016	1.007±0.018	0.998	0.9093		0.1857
$RER_{\dot{V}_{O_2max}}$[b]	1.001±0.016	1.013±0.017	0.988	0.5168		
$RER_{\dot{V}_{CO_2max}}$[b]	1.049±0.011	1.097±0.014	0.956	**0.0465**		0.3750
$RER_{\dot{V}_{CO_2max}}$[b]	1.047±0.011	1.101±0.013	0.951	**0.0206**		
Maximum speeds (m min^{-1})	48.27±2.16	39.41±2.27	1.225	**0.0388**	0.4962	0.4377
Maximum speeds (m min^{-1})	47.94±2.17	39.80±2.26	1.205	**0.0481**	0.4815	
Hyperoxia (30% O_2)						
$\dot{V}_{O_2max}$ (ml min^{-1})	6.241±0.261	5.222±0.281	1.195	**0.0164**	**0.0065**	0.6228
$\dot{V}_{O_2max}$ (ml min^{-1})	6.270±0.250	5.189±0.267	1.208	**0.0074**	**0.0040**	
$\dot{V}_{CO_2}$ @ $\dot{V}_{O_2max}$[a]	5.718±0.181	5.004±0.211	1.143	0.0620	**0.0117**	0.4602
$\dot{V}_{CO_2}$ @ $\dot{V}_{O_2max}$[a]	5.759±0.166	4.954±0.193	1.162	**0.0279**	**0.0063**	
$\dot{V}_{CO_2max}$ (ml min^{-1})	5.827±0.183	5.158±0.210	1.130	0.0740	**0.0126**	0.4266
$\dot{V}_{CO_2max}$ (ml min^{-1})	5.867±0.170	5.106±0.194	1.149	**0.0349**	**0.0085**	
$RER_{\dot{V}_{O_2max}}$[b]	0.918±0.018	0.961±0.020	0.955	0.0580		0.3871
$RER_{\dot{V}_{O_2max}}$[b]	0.921±0.018	0.957±0.019	0.962	0.0729		
$RER_{\dot{V}_{CO_2max}}$[b]	0.974±0.025	1.005±0.026	0.969	0.1852		0.9342
$RER_{\dot{V}_{CO_2max}}$[b]	0.974±0.024	1.006±0.026	0.968	0.1431		
Maximum speeds (m min^{-1})	49.91±1.91	41.06±2.10	1.216	**0.0280**	0.4987	0.3355
Maximum speeds (m min^{-1})	49.51±1.96	41.57±2.11	1.191	**0.0400**	0.4501	

[a]When we ran the models with $\dot{V}_{O_2max}$ as an additional covariate, selection and mass effects were never significant (P>0.05).

[b]Body mass was not included in the model.

Lines and measurement batch were included as random factors, and mass, age and mini-muscle were included as covariates. The two rows for each trait represent the results of analyses performed with and without a dummy variable accounting for the mini-muscle phenotype. Adjusted means were calculated from SAS PROC MIXED for a female of 28.2 g (±0.1 g between measurements) and 95 days of age. Values are for two-tailed tests, and those in bold indicate statistical significant effects (P<0.05). When the mini-muscle effect was significant, we also reported the sign of the partial regression from the mixed model in parentheses.

regardless of selection history (Fig. 2; Table 4). Similarly, within-subject contrasts – i.e. testing whether the mean at a given P_{O_2} is significantly different from the mean obtained at the previous level – show that $\dot{V}_{O_2max}$ mean values were always significantly higher at higher P_{O_2}, regardless of selection history ($F_{1,51}$=74.24, P<0.0001 for hypoxia *versus* normoxia, and $F_{1,51}$=57.35, P<0.0001 for hyperoxia *versus* normoxia; Fig. 2A). The same was true for $\dot{V}_{CO_2max}$ (P<0.008 in both cases; Fig. 2B). Selection history had a marginal effect on $\dot{V}_{O_2max}$ (one-tailed P=0.057; Table 4), supporting the conclusion that S mice

Table 3. *Regular ANCOVA (no nested factors) performed for line 6 individuals to check mini-muscle effects*

	Normal (*N*=12)	Mini-muscle (*N*=5)	Mini/normal	$P_{\text{body mass}}$	P_{mini}
Hypoxia					
$\dot{V}_{O_2max}$ (ml min^{-1})	4.925±0.087	5.284±0.135	1.073	**0.012**	**0.045**
$\dot{V}_{CO_2max}$ (ml min^{-1})	5.226±0.111	5.679±0.173	1.087	**0.004**	**0.048**
Maximum speed (m min^{-1})	46.5±1.5	44.3±2.4	0.953	0.452	0.452
Normoxia					
$\dot{V}_{O_2max}$ (ml min^{-1})	5.853±0.176	6.073±0.274	1.038	**0.043**	0.515
$\dot{V}_{CO_2max}$ (ml min^{-1})	6.100±0.141	5.965±0.219	0.978	**0.004**	0.614
Maximum speed (m min^{-1})	52.2±1.6	51.5±2.5	0.987	0.467	0.814
Hyperoxia					
$\dot{V}_{O_2max}$ (ml min^{-1})	6.140±0.169	6.554±0.262	1.067	**0.002**	0.208
$\dot{V}_{CO_2max}$ (ml min^{-1})	5.823±0.195	6.185±0.302	1.062	**0.013**	0.333
Maximum speed (m min^{-1})	54.2±1.3	53.0±2.0	0.978	0.965	0.627

P-values represent two-tailed tests. Age and mass were covariates in the model. Adjusted means ± s.e.m. were calculated for a 27.0 g female.

attain higher $\dot{V}_{O_2max}$ at any given P_{O_2} (Fig. 2A). This was not the case with $\dot{V}_{CO_2max}$ (one-tailed P=0.237).

Changes in P_{O_2} had a less dramatic effect on $\dot{V}_{CO_2max}$ than on $\dot{V}_{O_2max}$ (Fig. 2). Accordingly, RER at $\dot{V}_{O_2max}$ and $\dot{V}_{CO_2max}$ changed significantly between ambient P_{O_2}. In addition, both RER indexes were significantly lower for S mice when mass effects were not controlled (Tables 2, 4). Contrast analyses from repeated-measures ANCOVA found significant differences in RER between hypoxia *versus* normoxia, and normoxia *versus* hyperoxia (P<0.0001 in both cases). In addition, maximum running speeds on the treadmill also changed significantly with P_{O_2}, being 13% higher in normoxia than in hypoxia ($F_{1,57}$=54.78, P<0.001) and 4.2% higher in hyperoxia than in normoxia ($F_{1,57}$=48.73, P<0.001; Fig. 2D). Line type (S *versus* C) was statistically significant as between-subject effects ($F_{1,6}$=76.43, P<0.001), indicating that S mice achieved higher speeds. Comparisons among contrasts were not performed separately for S and C because the line type × trial interaction was never statistically significant (i.e. responses across trials did not differ between S and C).

Values of $\dot{V}_{O_2max}$ at different P_{O_2} were significantly correlated at the level of individual variation after accounting for selection history, mass, age and batch (Fig. 3). Correlations performed separately for S and C, employing residuals from regular ANCOVA with lines as grouping factors, provide identical results: individual consistency in $\dot{V}_{O_2max}$ across P_{O_2} was statistically significant regardless of selection history (r>0.352, P<0.046 in all cases; Fig. 3).

Discussion

Previous studies of these same lines at generations 10, 32 and 34 also indicated that S mice have higher average aerobic capacities than C mice in normoxia (respectively, +6% for males in Swallow et al., 1998b; +18% for males in Rezende et al., 2006; +7% for females in Rezende et al., 2005). Although our *a priori* expectation was that mice from the S lines would evolve higher $\dot{V}_{O_2max}$, a small response, especially at earlier generations, is consistent with a study of the base population, which found limited evidence for narrow-sense heritability in this trait (Dohm et al., 2001). In the present study of females from generation 36, controlling statistically for body mass but not for the mini-muscle phenotype, S lines attained $\dot{V}_{O_2max}$ averaging 13.2% higher than C in normoxia (one-tailed P=0.0294; Table 2; Fig. 2A), although the values are within the range of those previously reported (Swallow et al., 1998b; Rezende et al., 2005,2006). Accordingly, we also detected a marginal effect of line type on $\dot{V}_{O_2max}$ in the repeated-measures ANOVA (one-tailed P=0.057; Table 4), supporting the conclusion that S females attained higher $\dot{V}_{O_2max}$ at all P_{O_2} as compared with C mice (Fig. 3).

In hypoxia, S lines attained $\dot{V}_{O_2max}$ averaging 10.7% (P=0.0291) higher than C, whereas in hyperoxia the difference was 20.8% (P=0.0074; Table 2; Fig. 2A). The repeated-measures analysis was unable to detect a significant line type × P_{O_2} interaction (two-tailed P=0.103; Table 4), although Table 2 and Fig. 2 suggest that S and C lines actually do respond differently when exposed to variation in P_{O_2}, particularly between normoxia and hyperoxia. Moreover, quadratic effects in the repeated-measure analysis bordered significance in C lines ($F_{1,20}$=4.08, two-tailed P=0.057) but not in S lines ($F_{1,31}$=1.16, P=0.289), in spite of the higher sample size in the latter group. That is, within the range of P_{O_2} studied, $\dot{V}_{O_2max}$ seems to increase essentially linearly in S mice but plateaus at normoxic levels in C mice (Fig. 2A).

With respect to maximal CO_2 production, Astorino and Robergs (2003, p. 13) stated that '*an enhanced* $\dot{V}_{CO_2}$ *may represent greater oxidative ATP production, yet few studies have demonstrated significant increases in* $\dot{V}_{CO_2}$ *in hyperoxia*'. Contrasts analyses show that S mice attained significantly higher $\dot{V}_{CO_2max}$ in hyperoxia *versus* normoxia ($F_{1,31}$=7.45, P=0.010), but C mice did not ($F_{1,20}$=1.58, P=0.223; Fig. 2B), suggesting that oxidative metabolism increased significantly in S lines during trials in hyperoxia. Although $\dot{V}_{CO_2max}$ could

Table 4. *F-values obtained from tests of between- and within-subjects effects from repeated-measures ANOVA on $\dot{V}_{O2max}$ attained during hypoxia, normoxia and hyperoxia*

		d.f.	*F*	Two-tailed *P*
$\dot{V}_{O_2max}$	Line type	1,6	3.405	0.114
	P_{O_2}	2,102	**63.116**	**<0.0001**
	P_{O_2} × line type	2,12	2.766	0.103
	P_{O_2} × line (line type)	12,102	0.628	0.814
	Individual	51,102	1.161	0.259
	Individual[a]	58,116	**1.492**	**0.035**
$\dot{V}_{CO_2max}$	Line type	1,6	0.581	0.475
	P_{O_2}	2,102	**27.546**	**<0.0001**
	P_{O_2} × line type	2,12	1.324	0.302
	P_{O_2} × line (line type)	12,102	0.776	0.673
	Individual	51,102	**2.373**	**0.0001**
	Individual[a]	58,116	**2.646**	**<0.0001**
$RER_{\dot{V}_{O_2max}}$	Line type	1,6	**15.204**	**0.008**
	P_{O_2}	2,102	**44.668**	**<0.0001**
	P_{O_2} × line type	2,12	0.005	0.995
	P_{O_2} × line (line type)	12,102	1.179	0.318
	Individual	51,102	1.242	0.177
	Individual[a]	58,116	1.116	0.305
$RER_{\dot{V}_{CO_2max}}$	Line type	1,6	**17.877**	**0.006**
	P_{O_2}	2,102	**35.525**	**<0.0001**
	P_{O_2} × line type	2,12	0.881	0.439
	P_{O_2} × line (line type)	12,102	0.271	0.992
	Individual	51,102	0.464	0.998
	Individual[a]	58,116	0.572	0.990
Maximum speed	Line type	1,6	**12.217**	**0.013**
	P_{O_2}	2,98	**51.425**	**<0.0001**
	P_{O_2} × line type	2,12	1.004	0.395
	P_{O_2} × line (line type)	12,98	1.243	0.265
	Individual	49,98	**2.612**	**<0.0001**
	Individual[a]	56,112	**4.800**	**<0.0001**

Analyses were performed including line type and line as factors, and line type effects were tested over the model with line × line type interaction as the grouping factor. Effects of individual variation were also estimated employing a different model with individuals as the grouping factor and P_{O_2} as the within-subject effect ([a]), being analogous to testing for repeatability across different P_{O_2} (see Materials and methods). Mass and age were not included as covariates. Values in bold indicate statistical significance at $P<0.05$ (pooled $N=59$).

increase because of higher muscle fiber recruitment in hyperoxia, or increased oxidative metabolism, the significant reductions observed in RER in hyperoxia (Fig. 2C) support the latter alternative (see below). Consistent with this possibility, Houle-Leroy et al. (2000) studied both sexes at generation 14 and found that S mice housed without wheel access exhibited a trend for higher mixed hind limb muscle aerobic capacities, as indicated by higher levels of mitochondrial (cytochrome *c* oxidase, carnitine palmitoyltransferase, citrate synthase, pyruvate dehydrogenase) and glycolytic (hexokinase, phosphofructokinase) enzymes, with lower anaerobic capacities, as indicated by lactate dehydrogenase (especially in males). [In males from generation 10, when the S *versus* C differential in wheel running was substantially less, Zhan et al. (1999) found no evidence of elevated succinate dehydrogenase activity in the medial gastrocnemius muscles of S mice.] Interestingly, $\dot{V}_{CO_2max}$ is related to some types of locomotor endurance at the level of individual variation in lizards (Garland, 1984).

Aerobic capacity and the mini-muscle phenotype

The mini-muscle allele is still segregating within line 6 but is now apparently fixed in line 3 (see also Syme et al., 2005). We therefore increased sample size within line 6 (see Materials and methods) to allow within-line comparisons. Whether any effects within line 6 would be the same as in line 3 are unknown and difficult to study because these lines probably differ in allele frequencies at many other loci, leading to possible effects of different genetic backgrounds and/or epistatic interactions. Nevertheless, analyses of the entire data set (all eight lines) and just within the polymorphic S line 6 were consistent: individuals with the mini-muscle phenotype had significantly higher $\dot{V}_{O_2max}$ and $\dot{V}_{CO_2max}$ under hypoxic conditions but not in normoxia or hyperoxia (Tables 2, 3; Fig. 1). Although previous studies have described a twofold increase in mass-specific oxidative capacity in hind limb muscles of individuals with the mini-muscle phenotype (Houle-Leroy et al., 2003), and higher wheel-running distances and speeds in some samples (Syme et al., 2005; Kelly et al., in press), this is the first report of significant effects of the mini-muscle phenotype on aerobic capacity at the whole-organism level.

Although the most straightforward explanation for elevated $\dot{V}_{O_2max}$ would be increased mass-specific aerobic capacity (*in vitro* catalytic rates of oxidative enzymes; Houle-Leroy et al., 2003), several lines of evidence suggest that additional factors may be involved. First, individuals with the mini-muscle phenotype have gastrocnemius almost 50% lighter than normal, compensating for the twofold increase in mass-specific oxidative capacity (Garland et al., 2002; Houle-Leroy et al., 2003; Syme et al., 2005). Second, one might expect that any favorable effect of the mini-muscle allele would be more pronounced under those conditions in which it was selected (normoxia). By contrast, differences between mini-muscle and normal individuals were somewhat inconsistent across different P_{O_2} (i.e. genotype-by-environment interaction; Tables 2, 3). Although this does not exclude the possibility that higher cellular oxidative capacities could increase O_2 gradient from capillaries to muscles, increasing overall O_2 extraction from the air and ultimately $\dot{V}_{O_2max}$ under hypoxic conditions, that scenario assumes that effects of hypoxia on additional steps of the O_2 cascade are negligible, which might not be the case (see below). Third, pleiotropic effects of the mini-muscle allele on other steps in the O_2 cascade are possible (see also Belter et al., 2004; Swallow et al., 2005).

Individuals with the mini-muscle must provide the same amount of O_2 to their muscles as normal mice but are constrained by smaller overall muscle cross-sections, and potentially fewer capillaries, to transport blood into these muscles. Mice with the mini-muscle phenotype have significantly larger ventricle mass than normal (Garland et al., 2002; Swallow et al., 2005; unpublished result in Gomes et al., 2004), which might be indicative of compensations at other levels, such as higher cardiac outputs. Another possibility is that the capillary/fiber ratio has increased in the mini-muscle, as has been described in muscles with high oxidative capacity (Hepple and Vogell, 2004; and references therein), which could facilitate the O_2 flux to muscles and decrease the necessity of increased blood pressures.

Factors limiting $\dot{V}_{O2max}$

Factors limiting maximum O_2 uptake may be broadly broken into four categories: (1) pulmonary diffusing capacity, (2) maximum cardiac output, (3) oxygen transport in the blood and (4) skeletal muscle characteristics (Wagner, 1996; Richardson et al., 1999; Bassett and Howley, 2000; see Hochachka, 2003, for details of O_2 transport within cells). According to Fick's law of diffusion and the Fick principle, the relative importance of each of these factors in determining $\dot{V}_{O2max}$ may vary at different atmospheric P_{O_2}. Hence, the interaction between environmental P_{O_2}, selection history and the presence or absence of the mini-muscle phenotype and their effects on $\dot{V}_{O2max}$ must be considered in a holistic manner.

Our results show that neither S nor C lines can compensate for the lower P_{O_2} gradient during hypoxic trials (Fig. 2). Assuming a mixed venous P_{O_2} of 21 and 34 Torr (1 Torr=133.3 Pa) during $\dot{V}_{O2max}$ in hypoxia and normoxia (based on measurements performed in rats by Henderson et al., 2002; their inspired P_{O_2} in hypoxia was ~30 Torr lower than ours, however) and everything else being equal, one would predict a decrease of 32% in $\dot{V}_{O2max}$ in hypoxia compared with normoxia, contrasting with 12.9±1.6% and 10.9±1.9% for S and C, respectively (Table 2; Fig. 2). Although lower diffusion rates are probably the main causal factor, several physiological processes, such as reduced cardiac output, might be contributing to lower $\dot{V}_{O2max}$ in hypoxia (e.g. Gonzalez et al., 1998; Calbet et al., 2003).

The mean increase of 7.2% in $\dot{V}_{O2max}$ in hyperoxia *versus* normoxia is considerably lower than the 55.1% increase predicted if venous P_{O_2} was assumed to be identical to values obtained in normoxia (above), supporting the hypothesis that muscle aerobic capacity might constrain higher $\dot{V}_{O2max}$ in hyperoxia. Accordingly, previous studies have reported relatively minor increases in $\dot{V}_{O2max}$ with increased O_2 delivery (Spriet et al., 1986; Lindstedt et al., 1988; Richardson et al., 1999; Lindstedt and Conley, 2001), which suggests that O_2 delivery was enhanced beyond the capacity of mitochondria to metabolize the O_2 available. Although it is possible that mitochondrial capacity might be reaching its limit in 30% O_2 (e.g. Richardson et al., 1999 in humans at 100% P_{O_2}), two lines of evidence suggest that $\dot{V}_{O2max}$ in S and C lines under normal conditions (i.e. normoxia) is constrained by pulmonary and cardiovascular systems. First, RER decreased significantly as P_{O_2} increased (Fig. 2C), which might be due to a higher O_2 transport, facilitated by higher O_2 concentration at the alveolar level. Second, both S and C mice had significantly higher $\dot{V}_{O2max}$ in hyperoxia compared with normoxia (see Results), emphasizing that peripheral muscles can increase aerobic metabolism when more intracellular O_2 is available.

The differential response of $\dot{V}_{O2max}$ between S and C mice during hyperoxia suggests that selection for high wheel running affected differentially some of the components involved in the O_2 cascade. Polynomial analyses suggest that C animals could not increase $\dot{V}_{O2max}$ in hyperoxia to the extent that S mice did (Fig. 2A), although the P_{O_2} × line type interaction was not statistically significant according to a two-tailed hypothesis (Table 4). Nevertheless, ANOVAs and ANCOVAs have relatively low power to detect interactions (e.g. Wahlsten, 1990; Houle-Leroy et al., 2000), and it is reasonable to suggest that $\dot{V}_{O2max}$ responses to hyperoxia differ between S and C lines (Table 4; P_{O_2} × line type $F_{2,12}$=2.77, one-tailed P=0.051).

If so, then two non-exclusive hypotheses are possible, assuming that O_2 transport within the cell and mitochondrial oxidative capacity do not limit cellular metabolism (see above) or differ between line types (e.g. Houle-Leroy et al., 2000). First, O_2 extraction in lungs and/or muscles might have increased in S lines. Henderson et al. (2002) showed that rats selected for high treadmill endurance during forced exercise evolved higher $\dot{V}_{O2max}$ despite minimal changes in convective O_2 transport (cardiac output was significantly different between lines, however; Hussain et al., 2001). A higher O_2 extraction was attributed to greater muscle fiber capillarization (smaller muscle fibers, same number of capillaries) in the high endurance performers (see also Howlett et al., 2003). If the same pattern is true in our lines, this could explain why S lines have consistently higher $\dot{V}_{O2max}$ than C mice regardless of P_{O_2}.

If differences between S and C were simply caused by unequal O_2 extraction, then one would expect the relationship between P_{O_2} and $\dot{V}_{O2max}$ to be linear for both S and C, but with different slopes. The presence of a significant quadratic component in C lines (Fig. 2A) suggests that other factors might be involved. Therefore, a second alternative is that O_2 convection (either due to ventilatory convection or cardiac output) limits higher $\dot{V}_{O2max}$ in C lines – but not in S lines – in normoxia and hyperoxia. One possibility is that hyperoxia might be preventing hemoglobin desaturation at high exercise levels in S lines, as reported for humans (Nummela et al., 2002). Breathing frequency was significantly higher in S lines during $\dot{V}_{O2max}$ in a He–O_2 atmosphere (Rezende et al., 2005), although it is not clear if that would be the case during forced exercise.

Alternatively, S mice might have increased cardiac output. There is general agreement that higher rates of blood flow result in an improvement in $\dot{V}_{O2max}$ (e.g. Saltin and Strange, 1992), and we have recently found that S mice have significantly larger ventricles after accounting for mini-muscle

effects (Gomes et al., 2004; differential training effects between S and C may be a factor because mice were measured after 6 days of wheel access). Although severe hyperoxia (100% O_2) causes reduction of microvascular flow due to vasoconstriction in resting hamsters (Tsai et al., 2003), we do not know if any such effects might occur at 30% O_2 during strenuous exercise in mice, or if increased cardiac outputs and presumably higher blood pressures could prevent such maldistribution of perfusion in the microcirculation. Nevertheless, significantly lower RER in S lines (Table 4) suggests that selection for high activity has led to more efficient O_2 transport to muscles.

Inspection of the values in Table 3 shows that mice with mini-muscles always tended to have higher $\dot{V}_{O2max}$ than those with the normal phenotype (see also Fig. 1). Although larger hearts and higher mass-specific muscle aerobic capacity suggest that increases in cardiac output, blood pressure and P_{O_2} gradient at the capillary level might potentially explain (at least partially) the differences in aerobic capacity between phenotypes, further studies are required to address this question. Indeed, many of the factors that could explain S *versus* C differences could also explain differences between mini *versus* normal phenotypes (e.g. greater muscle fiber capillarization; as in Howlett et al., 2003).

Finally, some studies have proposed that central nervous system limitations might constrain $\dot{V}_{O2max}$ in hyperoxia (see Astorino and Robergs, 2003). Neurobiological differences between S and C lines have been described, especially regarding dopaminergic function (review in Rhodes et al., 2005), and dopaminergic function is known to be involved in respiratory control in mice (e.g. Huey et al., 2003). The higher variability in $\dot{V}_{CO2max}$ estimates in normoxia in both S and C lines (Fig. 2B) might be related to changes in ventilatory regulation in this treatment, associated with the presence of atmospheric CO_2 in the incurrent gas – i.e. CO_2 was not removed in normoxia, and baseline measurements were used to obtain corrected values of $\dot{V}_{CO_2}$ (see Materials and methods). This possibility seems very unlikely, however, given the very small fraction of CO_2 inspired compared to lung P_{CO_2}. It is also unclear whether ventilatory differences between S and C lines (above), possibly associated with changes in dopamine function *per se*, could explain the 'plateau' in $\dot{V}_{O2max}$ observed in C lines in hyperoxia (Fig. 2A).

Concluding remarks

Our results support previous indications that selection for high voluntary wheel running has caused increased aerobic capacity (Swallow et al., 1998b; Rezende et al., 2005, in press). Although we have previously reported correlated responses at the level of peripheral tissues (e.g. increased frequency of the mini-muscle allele leading to more oxidative gastrocnemius), the main difference between S and C lines across different P_{O_2} seems to be associated with central factors in the O_2 cascade. Our results also show that individual variation in aerobic performance can be independent of P_{O_2}; that is, individuals with high aerobic performances in normoxia consistently achieve higher $\dot{V}_{O2max}$ in hypoxia and hyperoxia (Table 4). In addition, selection history seems to be the major factor determining how individuals will perform at different P_{O_2}.

Because mice attained higher $\dot{V}_{O2max}$ in hyperoxia regardless of selection history, our study shows that $\dot{V}_{O2max}$ in normoxia seems to be centrally limited in these lines. However, in a previous study (Rezende et al., 2005), we used cold exposure in combination with a He–O_2 atmosphere to demonstrate conclusively that lungs and heart can provide more O_2 than necessary to attain $\dot{V}_{O2max}$ on the treadmill, a result that indicates that central limitations do not seem to be applicable for $\dot{V}_{O2max}$ during forced exercise. Taken together, these results suggest that whole-organism $\dot{V}_{O2max}$ is probably dependent on the interaction of several factors at many different levels in the O_2 cascade, as has been suggested in recent studies of metabolic regulation (e.g. Lindstedt and Conley, 2001; Darveau et al., 2002; Suarez and Darveau, 2005). Nevertheless, the differential response to O_2 availability between S and C lines suggests that subordinate traits have evolved somewhat independently and thus that S lines have not evolved in a strictly symmorphotic fashion.

We have also observed that mice with the mini-muscle phenotype achieved significantly higher $\dot{V}_{O2max}$ than normal only under hypoxic conditions, which indicates a genotype-by-environment interaction. This interesting result seems consistent with the idea that phenotypes favorable for locomotor endurance might also evolve as a correlated response to selection for hypoxia tolerance (Hochachka et al., 1998). Further, it would seem that a selection experiment designed to increase hypoxia tolerance could be a novel way to explore such hypotheses about correlated evolution of complex phenotypes (see Swallow and Garland, 2005 and references therein). Finally, the mini-muscle phenotype may prove an interesting model to elucidate genetic components underlying increased hypoxia tolerance and its evolution (Powell, 2002). Although the reasons why the mini-muscle phenotype achieves higher $\dot{V}_{O2max}$ only in hypoxia are yet unknown, we hypothesize that higher cardiac outputs and muscle vascularization, as well as increased P_{O_2} gradients due to enhanced muscle O_2 extraction, are involved.

List of abbreviations

$\dot{V}_{O_2}$	oxygen consumption (ml O_2 min^{-1})
$\dot{V}_{O2max}$	maximum oxygen consumption observed during forced exercise (highest 1-min average)
$\dot{V}_{CO_2}$	CO_2 production (ml CO_2 min^{-1})
$\dot{V}_{CO2max}$	maximum CO_2 production during forced exercise (highest 1-min average)
$RER_{\dot{V}O2max}$	respiratory exchange ratio ($=\dot{V}_{CO_2}/\dot{V}_{O_2}$) during $\dot{V}_{O2max}$
$RER_{\dot{V}CO2max}$	respiratory exchange ratio during $\dot{V}_{CO2max}$
P_{O_2}	oxygen availability (14, 21 and 30% O_2)

We thank L. Karpinski, J. Sinclair and several undergraduates for their help with the mouse colony, and J. Urrutia for constructing the treadmill. All animal procedures are in compliance with the UCR Institutional Animal Care and Use Committee and US laws. This study was supported by NSF IBN-0111604 (to K. A. Hammond and M.A.C.) and NSF IBN-0212567 (T.G.).

References

Alexander, R. McN. (1989). Optimization and gaits in the locomotion of vertebrates. *Physiol. Rev.* **69**, 1199-1227.

Astorino, T. A. and Robergs, R. A. (2003). Effect of hyperoxia on maximal oxygen uptake, blood acid-base balance, and limitations to exercise tolerance. *J. Exerc. Physiol.* **6**, 8-20.

Bacigalupe, L. D. and Bozinovic, F. (2002). Design, limitations and sustained metabolic rate: lessons from small mammals. *J. Exp. Biol.* **205**, 2963-2970.

Bartholomew, G. A., Vleck, D. and Vleck, C. M. (1981). Instantaneous measurements of oxygen consumption during preflight warm-up and postflight cooling in sphinged and saturnid moths. *J. Exp. Biol.* **90**, 17-32.

Bassett, D. R., Jr and Howley, E. T. (2000). Limiting factors of maximum oxygen uptake and determinants of endurance performance. *Med. Sci. Sports Exer.* **32**, 70-84.

Belter, J. G., Carey, H. V. and Garland, T., Jr (2004). Effects of voluntary exercise and genetic selection for high activity levels on HSP72 expression in house mice. *J. Appl. Physiol.* **96**, 1270-1276.

Calbet, J. A. L., Boushel, R., Rådegran, G., Søndergaard, H., Wagner, P. D. and Saltin, B. (2003). Determinants of maximal oxygen uptake in severe acute hypoxia. *Am. J. Physiol.* **284**, R291-R303.

Chappell, M. A., Rezende, E. L. and Hammond, K. A. (2003). Age and aerobic performance in deer mice. *J. Exp. Biol.* **206**, 1221-1231.

Chappell, M. A., Garland, T., Rezende, E. L. and Gomes, F. R. (2004). Voluntary running in deer mice: speed, distance, energy costs and temperature effects. *J. Exp. Biol.* **207**, 3839-3854.

Curran-Everett, D. (2000). Multiple comparisons: philosophies and illustrations. *Am. J. Physiol.* **279**, R1-R8.

Darveau, C.-A., Suarez, R. K., Andrews, R. D. and Hochachka, P. W. (2002). Allometric cascade as a unifying principle of body mass effects on metabolism. *Nature* **417**, 166-170.

Dohm, M. R., Hayes, J. P. and Garland, T., Jr (2001). The quantitative genetics of maximal and basal metabolic rates of oxygen consumption in mice. *Genetics* **159**, 267-277.

Garland, T., Jr (1984). Physiological correlates of locomotory performance in a lizard: an allometric approach. *Am. J. Physiol.* **247**, R806-R815.

Garland, T., Jr (1998). Conceptual and methodological issues in testing the predictions of symmorphosis. In *Principles of Animal Design. The Optimization and Symmorphosis Debate* (ed. E. R. Weibel, C. R. Taylor and L. Bolis), pp. 40-47. Cambridge: Cambridge University Press.

Garland, T., Jr (2001). Phylogenetic comparison and artificial selection: two approaches in evolutionary physiology. In *Hypoxia: From Genes to the Bedside. Advances in Experimental Biology and Medicine*, vol. 502 (ed. R. C. Roach, P. D. Wagner, and P. H. Hackett), pp. 107-132. New York: Kluwer Academic/Plenum Publishers.

Garland, T., Jr (2003). Selection experiments: an under-utilized tool in biomechanics and organismal biology. In *Vertebrate Biomechanics and Evolution* (ed. V. L. Bels, J.-P. Gasc and A. Casinos), pp. 23-56. Oxford: UK BIOS Scientific Publishers.

Garland, T., Jr and Carter, P. A. (1994). Evolutionary physiology. *Annu. Rev. Physiol.* **56**, 579-621.

Garland, T., Jr and Huey, R. H. (1987). Testing symmorphosis: does structure match functional requirements? *Evolution* **41**, 1404-1409.

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.

Girard, I., McAleer, M. C., Rhodes, J. S. and Garland, T., Jr (2001). Selection for high voluntary wheel-running increases speed and intermittency in house mice (*Mus domesticus*). *J. Exp. Biol.* **204**, 4311-4320.

Gomes, F. R., Rezende, E. L., Bunkers, J. L. and Garland, T., Jr (2004). Organ masses and carbohydrate metabolism of mice artificially selected for high voluntary wheel running. *Integr. Comp. Biol.* **43**, 912.

Gonzalez, N. C., Clancy, R. L., Moue, Y. and Richalet, J.-P. (1998). Increasing maximal heart size increases maximal O_2 uptake in rats acclimatized to simulated altitude. *J. Appl. Physiol.* **84**, 164-168.

Gordon, M. S. (1998). Evolution of optimal systems. In *Principles of Animal Design. The Optimization and Symmorphosis Debate* (ed. E. R. Weibel, C. R. Taylor and L. Bolis), pp. 37-39. Cambridge: Cambridge University Press.

Hammond, K. A. and Diamond, J. (1997). Maximal sustained energy budgets in humans and animals. *Nature* **386**, 457-462.

Hayes, J. P. and Chappell, M. A. (1990). Individual consistency of maximal oxygen consumption in deer mice. *Funct. Ecol.* **4**, 495-503.

Hayes, J. P. and Jenkins, S. H. (1997). Individual variation in mammals. *J. Mamm.* **78**, 274-293.

Henderson, K. K., Wagner, H., Favret, F., Britton, S. L., Koch, L. G., Wagner, P. D. and Gonzalez, N. C. (2002). Determinants of maximal O_2 uptake in rats selectively bred for endurance running capacity. *J. Appl. Physiol.* **93**, 1265-1274.

Hepple, R. S. and Vogell, J. E. (2004). Anatomic capillarization is maintained in relative excess of fiber oxidative capacity in some skeletal muscles of late middle-aged rats. *J. Appl. Physiol.* **96**, 2257-2264.

Hochachka, P. W. (2003). Intracellular convection, homeostasis and metabolic regulation. *J. Exp. Biol.* **206**, 2001-2009.

Hochachka, P. W., Gunga, H. C. and Kirsch, K. (1998). Our ancestral physiological phenotype: an adaptation for hypoxia tolerance and for endurance performance? *Proc. Natl. Acad. Sci. USA* **95**, 1915-1920.

Houle-Leroy, P., Garland, T., Swallow, J. G. and Guderley, H. (2000). Effects of voluntary activity and genetic selection on muscle metabolic capacitites in house mice *Mus domesticus*. *J. Appl. Physiol.* **89**, 1608-1616.

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.* **284**, R433-R443.

Howlett, R. A., Gonzalez, N. C., Wagner, H. E., Fu, Z., Britton, S. L., Koch, L. G. and Wagner, P. D. (2003). Skeletal muscle capillarity and enzyme activity in rats selectively bred for running endurance. *J. Appl. Physiol.* **94**, 1682-1688.

Huey, K. A., Szewczak, J. M. and Powell, F. L. (2003). Dopaminergic mechanisms of neural plasticity in respiratory control: transgenic approaches. *Respir. Physiol. Neurobiol.* **135**, 133-144.

Hussain, S. O., Barbato, J. C., Koch, L. G., Metting, P. J. and Britton, S. L. (2001). Cardiac function in rats selectively bred for low- and high-capacity running. *Am. J. Physiol.* **281**, R1787-R1791.

Jones, J. H. and Lindstedt, S. L. (1993). Limits to maximal performance. *Annu. Rev. Physiol.* **55**, 547-569.

Kelly, S. A., Czech, P. P., Wight, J. T., Blank, K. M. and Garland, T., Jr (in press). Experimental evolution and phenotypic plasticity of hindlimb bones in high-activity house mice. *J. Morph.* in press.

Kemi, O. J., Loennechen, J. P., Wisløff, U. and Ellingsen, Ø. (2002). Intensity-controlled treadmill running in mice: cardiac and skeletal muscle hypertrophy. *J. Appl. Physiol.* **93**, 1301-1309.

Koch, L. G. and Britton, S. L. (2001). Artificial selection for intrinsic aerobic endurance running capacity in rats. *Physiol. Genomics* **5**, 45-52.

Lindstedt, S. L. and Conley, K. E. (2001). Human aerobic performance: too much ado about limits to $\dot{V}_{O_2}$. *J. Exp. Biol.* **204**, 3195-3199.

Lindstedt, S. L., Wells, D. J., Jones, J. H., Hoppeler, H. and Thronson, H. A., Jr (1988). Limitations to aerobic performance in mammals: interaction of structure and demand. *Int. J. Sports Med.* **9**, 210-217.

Noakes, T. D., Peltonen, J. E. and Rusko, H. K. (2001). Evidence that a central governor regulates exercise performance during acute hypoxia and hyperoxia. *J. Exp. Biol.* **204**, 3225-3234.

Nummela, A., Hämäläine, I. and Rusko, H. (2002). Effect of hyperoxia on metabolic responses and recovery in intermittent exercise. *Scand. J. Med. Sci. Sports* **12**, 309-315.

Powell, F. L. (2002). Functional genomics and the comparative physiology of hypoxia. *Annu. Rev. Physiol.* **65**, 203-230.

Rezende, E. L., Chappell, M. A., Gomes, F. R., Malisch, J. L. and Garland, T., Jr (2005). Maximal metabolic rates during voluntary exercise, forced exercise, and cold exposure in house mice selectively bred for high wheel running. *J. Exp. Biol.* **208**, 2447-2458.

Rezende, E. L., Kelly, S. A., Gomes, F. R., Chappell, M. A. and Garland, T., Jr (2006). Effects of size, sex, and voluntary running speeds on costs of locomotion in lines of laboratory mice selectively bred for high wheel-running activity. *Physiol. Biochem. Zool.* **79**, in press.

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.

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution* **43**, 223-225.

Richardson, R. S., Leigh, J. S., Wagner, P. D. and Noyszewski, E. A. (1999). Cellular PO_2 as a determinant of maximal mitochondrial O_2 consumption in trained human skeletal muscle. *J. Appl. Physiol.* **87**, 325-331.

Saltin, B. and Strange, S. (1992). Maximal O_2 uptake: 'old' and 'new' arguments for a cardiovascular limitation. *Med. Sci. Sports Exerc.* **24**, 30-37.

Spriet, L. L., Gledhill, N., Froese, A. B. and Wilkes, D. L. (1986). Effect of graded erythrocythemia on cardiovascular and metabolic responses to exercise. *J. Appl. Physiol.* **61**, 1942-1948.

Suarez, R. K. and Darveau, C. A. (2005). Multi-level regulation and metabolic scaling. *J. Exp. Biol.* **208**, 1627-1634.

Swallow, J. G. and Garland, T., Jr (2005). Selection experiments as a tool in evolutionary and comparative physiology: insights into complex traits–An introduction to the symposium. *Integr. Comp. Biol.* **45**, 387-390.

Swallow, J. G., Carter, P. A. and Garland, T., Jr (1998a). Artificial selection for increased wheel-running behavior in house mice. *Behav. Genet.* **28**, 227-237.

Swallow, J. G., Garland, T., Carter, P. A., Zhan, W. Z. and Sieck, G. C. (1998b). Effects of voluntary activity and genetic selection on aerobic capacity in house mice (*Mus domesticus*). *J. Appl. Physiol.* **84**, 69-76.

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.

Taylor, C. R. and Weibel, E. R. (1981). Design of the mammalian respiratory system I: problem and strategy. *Respir. Physiol.* **44**, 1-10.

Tsai, A. G., Cabrales, P., Winslow, R. M. and Intaglietta, M. (2003). Microvascular oxygen distributionin awake hamster window chamber model during hyperoxia. *Am. J. Physiol.* **285**, H1537-H1545.

Wagner, P. D. (1996). Determinant of maximum oxygen transport and utilization. *Annu. Rev. Physiol.* **58**, 21-50.

Wahlsten, D. (1990). Insensitivity of the analysis of variance to heredity-environment interaction. *Behav. Brain Sci.* **13**, 109-161.

Zar, J. H. (1999). *Biostatistical Analysis*, 4th edn. London: Prentice-Hall International, Inc.

Zhan, W.-Z, Swallow, J. G., Garland, T., Proctor, D. N., Carter, P. A. and Sieck, G. C. (1999). Effects of genetic selection and voluntary activity on the medial gastrocnemius muscle in house mice. *J. Appl. Physiol.* **87**, 2326-2333.