ELSEVIER

Physiology & Behavior 77 (2002) 27–38

PHYSIOLOGY & BEHAVIOR

Dominance, plasma testosterone levels, and testis size in house mice artificially selected for high activity levels

Kurt F. Klomberg, Theodore Garland, Jr.*, John G. Swallow[1], Patrick A. Carter[2]

Department of Zoology, University of Wisconsin, Madison, WI 53706, USA

Received 3 February 2000; received in revised form 13 August 2001; accepted 22 April 2002

Abstract

Male house mice (*Mus domesticus*) from four replicate lines selectively bred for high voluntary wheel-running behavior were compared with four random-bred control lines with respect to dominance, testis size, and plasma testosterone level. Behavior was measured with a tube apparatus in which focal mice encountered a standard opponent from an inbred strain, and positions of mice were scored over a 10-min period; the test was replicated the following day. Blood samples were taken from undisturbed mice 1 week prior to testing (baseline condition) and immediately after the first tube test; plasma testosterone was measured by enzyme immunoassay with chromatography. As compared with control lines, mice from selected lines tended to be smaller in body mass, to have larger testes, and were significantly less likely to advance towards their opponent during the second tube-test encounter. However, no significant differences in either baseline or post-encounter testosterone levels were detected. Significant differences in body mass, relative testis size, position during the first tube-test encounter, and baseline testosterone were found among the replicate lines within linetype, which indicates founder effects, random genetic drift, unique mutations, and/or multiple responses to selection. At the level of individual variation (residuals from nested analysis of covariance models), an inverse relationship between baseline testosterone and advancing in the tube test was observed, and the relationship was stronger during the second test day. This unexpected result may reflect an alternate coping strategy.

Keywords: Aggression; Artificial selection; Dominance; Genetic correlation; Testosterone; Testis; Tube test; Voluntary activity; Wheel running

1. Introduction

Artificial selection (selective breeding) has proven to be a useful tool for studying the genetics and evolution of many different traits [1–3]. A response to selection can be used simply to demonstrate that the trait under selection has significant narrow-sense heritability (ratio of additive genetic to total phenotypic variance). More interestingly, when a trait responds to selection other traits often show correlated responses. These correlated responses, if consistent across replicate lines [4], indicate the presence of genetic correlations. Genetic correlations (reflecting additive genetic covariances) indicate that the trait selected upon and the traits showing a correlated response share, to some extent, common physiological, biochemical or developmental pathways, caused by pleiotropic gene action [5,6] or by linkage disequilibrium, although the latter tends not to persist for more than two generations in artificial selection experiments.

Here, we study replicate lines of house mice artificially selected for high voluntary wheel-running behavior [7] to test for possible genetic relationships with dominance, testis size, and plasma testosterone levels. The possibility of such relationships is suggested by correlated responses observed in previous artificial selection experiments. For example, in house mice selected for thermoregulatory nest building [2], high-selected lines exhibited shorter attack latencies as compared with both random-bred control and low-selected lines [8], whereas the low-selected lines ran significantly more on wheels than did either control or high-selected lines [9].

Although we are not aware of studies that have selected for social dominance in rodents, correlated responses are

* Corresponding author. Present address: Department of Biology, University of California, Riverside, CA 92521, USA. Tel.: +1-909-787-3524; fax: +1-909-265-6320.

E-mail address: tgarland@citrus.ucr.edu (T. Garland).

[1] Present address: Department of Biology, University of South Dakota, Vermillion, SD 57069, USA.

[2] Present address: Department of Zoology, Washington State University, Pullman, WA 99164, USA.

0031-9384/02/$ – see front matter
PII: S0031-9384(02)00767-9

also common in studies that have selected on aggression in rodents. Mice and other rodents typically use aggression to establish dominance, although the relationship between dominance and aggression is neither simple nor direct. In any case, mice from the first selective breeding program for aggression showed correlated responses in various traits, including locomotor activity in rotating-drum cages and in the open field; water and alcohol consumption; body mass; and testis and forebrain size [10,11] (reviewed in Ref. [12]). Another long-term experiment used a base population of wild house mice and selected for short and long attack latency (review in Ref. [13]). These lines also showed various correlated responses, such as in nest building [8], maternal care effort [14], testis size and testosterone secretory capacity [15], and plasma testosterone level [16]. Unfortunately, the foregoing selection experiments lacked replicate lines, so the changes in other traits cannot reliably be inferred to reflect additive genetic covariances.

Many other researchers have found that testosterone relates positively to aggression and/or dominance in house mice [17–20]. In addition, testosterone may have a variety of other behavioral and physiological effects. Bronson et al. [19], for example, demonstrated that female mice treated with testosterone exhibited reduced voluntary use of a running wheel. In mice selected for increased body mass, levels of circulating testosterone were elevated [21]. Testosterone can also affect organ sizes (e.g. see Ref. [22]) and muscle properties [23], which may in turn affect locomotor abilities, and plasma testosterone levels appear to be heritable in both house mice [24] and humans [25].

The aforementioned studies demonstrate or suggest relationships among locomotor activity, aggression or dominance, body size, testis size, and testosterone. Moreover, as noted by Sandnabba [12], "Since aggressive behavior contains more elements of motor activity than does nonaggressive behavior, a positive correlation between aggression and the overall motor activity was expected" (p. 484). Hence, we hypothesized that our lines of house mice selected for increased voluntary wheel-running [7] would exhibit correlated responses in dominance, testis size, and testosterone levels. We have found previously that, as compared with control lines, mice from the selected lines are smaller in body mass at maturity [26], exhibit higher maximum oxygen consumption during forced treadmill exercise, at least at some ages [27], and build smaller thermoregulatory nests [28], but show few other behavioral differences when housed individually in cages with attached wheels [29]. Here, we tested mice for dominance in a modified tube-test apparatus and also measured body mass, testis size, and both baseline and post-encounter plasma testosterone concentration. In addition to comparing selected and control lines by nested analysis of variance, we also compared the replicate lines within the selected and control groups with respect to all of the same traits. Divergence of replicate lines may occur because of founder effects, genetic drift, unique mutations or multiple responses to the imposed selection, processes that are rarely considered in behavioral endocrinology.

2. Materials and methods

2.1. Animals

Focal mice (*Mus domesticus*) were from second litters from the 10th generation of an artificial selection experiment for high voluntary wheel running [7]. The mice were originally derived from the genetically variable Hsd:ICR strain [30,31]. In generation 10, male mice from the four selected lines ran, on average, about 75% more than those from the four control lines [7]. This increase in wheel-running activity has been caused primarily by selected mice running at greater velocities rather than for a greater number of min each day (see also Refs. [27,29,32]). A total of 80 focal mice, one male from each of 10 different families from each of the eight lines, was studied (see Table 1 for ages). All procedures were consistent with guidelines in the National Institutes of Health *Guide for the Care and Use of Laboratory Animals*.

2.2. Design

We employed a 'standard opponent test' in which individuals of one or another genotype are tested against mice of a third, uniform, genotype [13,33,34]. Specifically, we tested focal mice for aggression by pairing them with approximately same-age male stimulus mice of the inbred C57BL/6Au strain (C57BL/6 was also used by Hyde and Ebert [35]) obtained from Dr. Robert Auerbach's breeding colony in the Department of Zoology at the University of Wisconsin—Madison. Because the stimulus mice came from a single inbred strain and were reared and housed uniformly, they were treated as a standard opponent. For testing, the focal mice were assigned randomly to 1 of 10 different groups (8 in each group). Each of the first 40 focal mice was tested against a different stimulus mouse. Then, to reduce the number of stimulus mice required, each of the remaining 40 focal mice was tested against one of the same stimulus mice used in the first 40 trials, with random assignment.

Mice housed individually have higher levels of aggression than do group-housed mice [36,37]; hence, focal mice were housed individually from weaning to testing, whereas stimulus mice were housed in groups of four males. Extremely clean housing conditions also may increase aggression in mice [36,38]; therefore, cages of the focal mice were cleaned 24 h prior to testing.

2.3. Test procedure

A simplified variation of the tube test was used to measure aggression of focal mice (see Ref. [39]; Fig. 33.2 in Ref. [40]; [41–43]). This method was used to reduce the

Table 1
Nested analysis of covariance comparing body and testis size of mice from selected and control lines

Character	Transform	*n*	Significance level				
			Linetype	Line	Log body mass	Age	Age^2
Body mass at baseline blood sample	none	76	.0843	.0216	–	.0001	.7472
Body mass at baseline blood sample (mean age 57.8 days, range 33–74)	none	74	.0347	.0366[a]	–	.0001	.5035
Body mass at Day 1 tube test (mean age 64.8 days, range 40–81)	none	76	.1242	.0287[b]	–	.0001	.9356
Body mass at sacrifice (mean age 66.8 days, range 42–83)	none	76	.0598	.2307	–	.0001	.9523
Testes mass (g)	log	76	.2171	.0005	.0001	.4169	.0360
Testes mass (g)	log	74	.0581	.0098[c]	.0022	.3157	.2209
Testes mass (g)	log	76	.4362	.0045	–	.1661	.0410
Testes mass (g)	log	72	.0377	.4341	–	.0055	.2128
Testes diameter (mm)	log	76	.2547	.0451	.0050	.7492	.1743
Testes diameter (mm)	log	71	.0929	.0480[d]	.0484	.2028	.0557
Testes diameter (mm)	log	76	.5452	.0484	–	.1874	.1595
Testes diameter (mm)	log	72	.1602	.1127	–	.0102	.0334

[a] Adjusted line means ± standard errors from SAS PROC GLM were 32.7 ± 0.90, 32.5 ± 0.89, 34.4 ± 0.95, and 35.4 ± 0.94 g, for control lines 1, 2, 4, and 5, respectively. Values were 33.1 ± 0.94, 30.8 ± 0.95, 29.5 ± 0.94, and 30.4 ± 0.94, for selected lines 3, 6, 7, and 8, respectively.

[b] Adjusted line means were 33.0 ± 0.94, 32.7 ± 0.93, 34.1 ± 0.99, and 35.0 ± 0.99 g, for control lines 1, 2, 4, and 5, respectively. Values were 34.4 ± 0.93, 31.2 ± 0.99, 30.1 ± 0.93, and 31.3 ± 0.98, for selected lines 3, 6, 7, and 8, respectively.

[c] Adjusted line means were −0.648 ± 0.0161, −0.654 ± 0.0159, −0.649 ± 0.0171, and −0.726 ± 0.0172 for control lines 1, 2, 4, and 5, respectively. Values were −0.626 ± 0.0160, −0.595 ± 0.0172, −0.608 ± 0.0167, and −0.642 ± 0.0192, for selected lines 3, 6, 7, and 8, respectively.

[d] Adjusted line means were 0.889 ± 0.0076, 0.891 ± 0.0067, 0.884 ± 0.0074, and 0.863 ± 0.0073, for control lines 1, 2, 4, and 5, respectively. Values were 0.900 ± 0.0068, 0.908 ± 0.0077, 0.889 ± 0.0071, and 0.894 ± 0.0081, for selected lines 3, 6, 7, and 8, respectively.

possibility of injury during any agonistic interaction. The apparatus was constructed of four clear, 0.5-m long acrylic tubes with an internal diameter of 2.5 cm. These were mounted on a 2.5-m board, end to end, with a removable opaque divider between the two center tubes. Along the length of the board, a metric ruler was placed with the 0.0-cm position to the left end of the tube set-up and the 200-cm position at the right end. Each tube had air holes of 2.0-mm diameter at 2.5-cm intervals along the top of the tube.

One week prior to a group's (see above) testing, blood was drawn (see below) and body mass was recorded. Within a group, testing order was random, and the observer was blind with respect to the focal mouse's status as belonging to a selected or control line.

During testing, one inbred stimulus mouse was weighed and placed inside one end of the apparatus, and the focal mouse was weighed and placed inside the end opposite the stimulus mouse. The side of the apparatus that the stimulus mouse entered was determined by a coin flip. Two plastic plungers were used to move the mice into the tubes and then briefly hold them near the center divider. The divider was removed immediately, the plungers were pulled rapidly out of the tube, and two rubber stoppers were placed in the opposite ends of the tube apparatus. The observer used a hand-held stopwatch to time encounters, and recorded the position of the focal mouse's ears every 30 s for 10 min. The same person observed all trials.

At the end of the test, the focal mouse was removed by disassembling the apparatus and, if necessary, using the plunger to force the individual from the tube. A blood sample was then taken immediately (see below). Before each test, the tube apparatus was cleaned with soapy water. The next day the procedure was repeated with the same eight mice. A focal mouse faced the same stimulus mouse during both tests.

2.4. Gonad extraction

The day following the second tube test, focal mice were killed by cervical dislocation and their testes were removed and weighed to the nearest 0.001 g. The greatest diameter of each testis was measured to the nearest 0.01 mm with digital calipers.

2.5. Blood sampling

Blood samples were collected from the suborbital sinus [44] of each focal mouse on two separate occasions. Although we are aware that mice exhibit pulsatile secretion of testosterone [45,46], we chose not to attempt to take multiple blood samples (for pooling) prior to behavioral testing because of ethical concerns and because it might have had adverse effects on behavior and/or health (e.g. see Ref. [47]).

At each blood sample, two heparinized micro-hematocrit tubes were filled, which yielded a total blood volume of approximately 160 μl per sample. Blood samples were collected as quickly as possible, and times between initial restraint and the end of blood drawing ranged from 22 to 229 s with a mean of 69 s (median = 60 s) for baseline samples (n = 77) and 20 to 170 s with a mean of 56 s (median = 50 s) for post-tube-test samples (n = 77) (both distributions were positively skewed). The first sample was taken 1 week prior to tube testing (1030–1345 h), from the right eye, the second immediately after the first tube test (0900–1600 h), from the left eye. Blood was centrifuged in a micro-hematocrit centrifuge for 8 min and

the plasma was frozen and stored at −80 °C until testosterone assays.

2.6. Testosterone assays

Assays were performed in facilities of the Wisconsin Regional Primate Research Center. Because of competitive binding of different steroids with the antibody, chromatographic separation of steroids was performed [48] after extraction. Chromatography columns consisted of 5-ml glass pipettes packed with 0.7 g of Celite mixed with 0.175 ml of ethylene glycol and 0.175 ml propylene glycol (Fisher Scientific). The columns were rinsed with two 3.5-ml washes of isooctane (Fisher Scientific). This was eluted under nitrogen pressure and discarded. After chromatography the samples were vortexed and refrigerated until assay. To determine procedural losses in the extraction and chromatography procedures, two recoveries for each assay were collected. These recoveries utilized [^{3}H]testosterone (2000 cpm; New England Nuclear, SA 100 Ci/mmol) in phosphate-buffered saline (PBS; 0.1 M, pH 7.0).

Plasma testosterone concentrations were analyzed by enzyme immunoassay adapted from Munro and Lasley [49] and Saltzman et al. [50]. Microtiter plates (Nunc Maxisorb-96F) were coated with testosterone antibody (R156/7, raised in rabbits against testosterone-6-carboxymethyl oxime:BSA and provided by C. Munro, University of California, Davis). Testosterone:horseradish peroxidase (T:HRP) solution (1:50,000 in PBS–BSA) was also provided by C. Munro.

Serial dilution of a testosterone-spiked female mouse-plasma pool gave a displacement curve parallel to that obtained with testosterone standards. The recovery of testosterone standards added to 50 μl of sample was 99.75 ± 1.13% (mean ± S.E.M.). The sensitivity of the assay at 90% binding was 0.6 pg. The inter- and intra-assay coefficients of variation [51] of a mouse plasma pool (22% binding) assayed in duplicate were 10.0% and 2.8%, respectively (eight assays). [^{3}H]Testosterone sample recoveries for chromatography were 84.98 ± 2.05% (mean ± S.E.M.).

2.7. Statistical analysis

We used nested analysis of covariance (ANCOVA) to compare selected and control lines, using Type III sums of squares in the SAS GLM procedure. Replicate line was nested within linetype (selected vs. control); covariates (e.g. body mass, age, the square of the *z*-transform of age (to allow for nonlinear effects of age, as shown for plasma testosterone levels by Bartke and Dalterio [45])) and interactions (linetype and line by all covariates) were included in the initial models. Models were then analyzed iteratively to remove nonsignificant ($P > .05$) interactions. Degrees of freedom for testing the linetype effect were always 1 and 6. Procedure GLM in SAS calculated regression coefficients for covariates and least-squares adjusted means for selected and control lines. Adjusted means were calculated using all covariates, regardless of significance levels. Only the significant ($P < .05$) Linetype × Covariate interactions were retained in the final models.

Traits were transformed as necessary to improve normality and linearity of relationships with various covariates (based on inspection of bivariate scatterplots). Body mass required no transformation when analyzed as a dependent variable. Both testis mass and diameter were log-transformed for analysis, as was body mass as the covariate, because internal organs, including the testes (e.g. see Ref. [52]), are generally expected to be linearly related to body mass on a log–log scale. Although we have no information on body composition of mice from the present study, at generation 13 mice from selected lines had less body fat than did mice from control lines [53]. If this difference also existed for the present mice, then an ANCOVA with total body mass as a covariate might indicate that selected lines have greater mass-adjusted testis size simply because they were leaner. Therefore, testis size was also analyzed without body mass as a covariate.

For consistency, body mass was also log-transformed when used as a covariate in all other analyses. Plasma

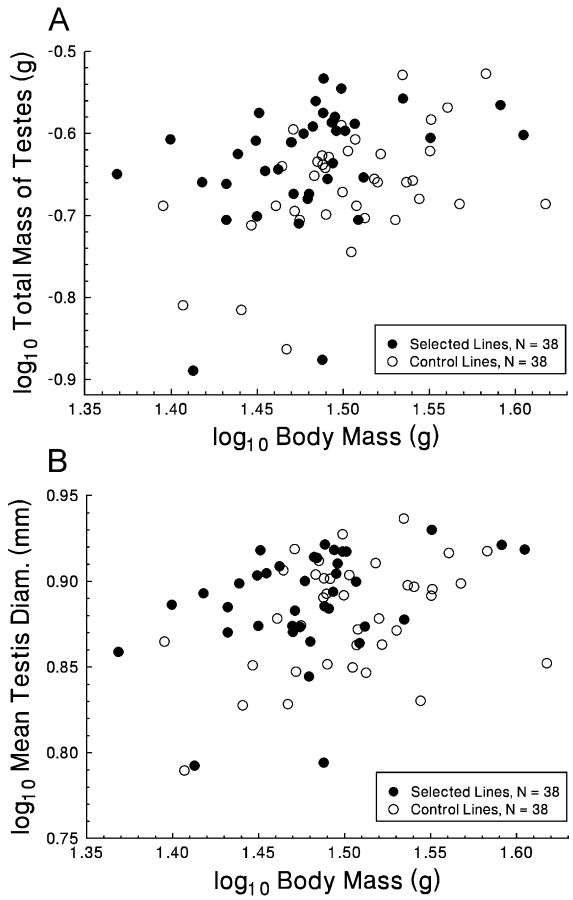

Fig. 1. (A) Log total mass of testes versus log body mass, (B) log mean testis diameter versus log body mass. Mice from selected lines tend to have larger testes both absolutely and relative to their body mass (see Table 1).

Table 2
Nested analysis of covariance comparing plasma testosterone levels of mice from selected and control lines

Character	Transform	n	Significance level								
			Linetype	Line	Log body mass	Age	Age^2	Time	Time^2	Sampling delay	Interaction
Baseline testosterone	log	76	.8519	.0172	.0765	.9331	.1360	.2122	.6041	.9474	[a]
Baseline testosterone	log	75	.9332	.0132[b]	.0602	.9929	.1839	.1243	.4248	.8811	[c]
Post-tube Day 1 testosterone	log	73	.2310	.0778[d]	.5075	.1364	.7927	.1086	.5339	.0948	none

[a] Time × Linetype P=.8670, Time × Line P=.0180.
[b] Adjusted line means ± standard errors from SAS PROC GLM were 0.558 ± 0.1556, 0.156 ± 0.1597, − 0.088 ± 0.1704, and 0.537 ± 0.1713, for control lines 1, 2, 4, and 5, respectively. Values were 0.514 ± 0.1548, 0.472 ± 0.1705, 0.489 ± 0.1686, and 0.089 ± 0.1659, for selected lines 3, 6, 7, and 8, respectively.
[c] Time × Linetype P=.9254, Time × Line P=.0138.
[d] Adjusted line means were 0.579 ± 0.2056, 1.004 ± 0.2137, 1.066 ± 0.2238, and 0.238 ± 0.2370, for control lines 1, 2, 4, and 5, respectively. Values were 0.671 ± 0.2123, 0.128 ± 0.2407, 0.406 ± 0.2177, and 0.452 ± 0.2192, for selected lines 3, 6, 7, and 8, respectively.

testosterone levels were right-skewed and so were log-transformed for analyses. Position in the tube tests was often skewed and/or kurtotic, so rank-transformed values were analyzed [54,55].

Because parametric statistical tests such as ANCOVA can be highly sensitive to departures from normality of residuals, we paid close attention to the distribution of residuals. For several traits, some individuals were notable outliers (e.g. standardized residuals greater than 2.5) and/or skewness of the overall distribution was relatively high. For these traits, we report results for both (1) the full data set and (2) after removing individuals one at a time while rechecking the distribution of residuals as well as the stability of P values for main effects (linetype and line-within-linetype). We accepted as final models those with skewness between 0.2 and − 0.2 and no standardized residuals greater than approximately 2.5 in magnitude. We emphasize interpretation of analyses of the reduced data sets.

In the figures, we show all individuals, with one exception: one individual mouse had extremely small testes (0.058 g total mass, mean diameter of 4.2 mm, body mass = 31.50 g, age = 79 days) as compared with all other individuals (compare Fig. 1). It was an extreme outlier in all analyses of testis size, though not in analyses of other traits. We excluded it from statistical analyses of all traits and from all figures because testicular hypoplasia can indicate other malformations including, but not limited to, neural development [56].

Adjusted line means are reported for some of the traits in the tables. These are from the SAS GLM procedure, and include adjustments for all factors and covariates in the models, regardless of statistical significance. Note that lines designated 1, 2, 4, and 5 are control, whereas 3, 6, 7, and 8 are selected.

3. Results

3.1. Body mass

Mice from selected lines tended to be smaller than mice from control lines (Table 1; see also Ref. [32] concerning siblings used in another experiment). When baseline blood samples were collected (1 week before tube tests), the difference in mass was statistically significant (P=.0347); adjusted means were 33.7 g for control mice and 31.0 g for selected mice. Lines within linetype also showed statistically significant differences for two of the three measures of body mass (Table 1). The stimulus mice were smaller than either selected or control-line focal mice

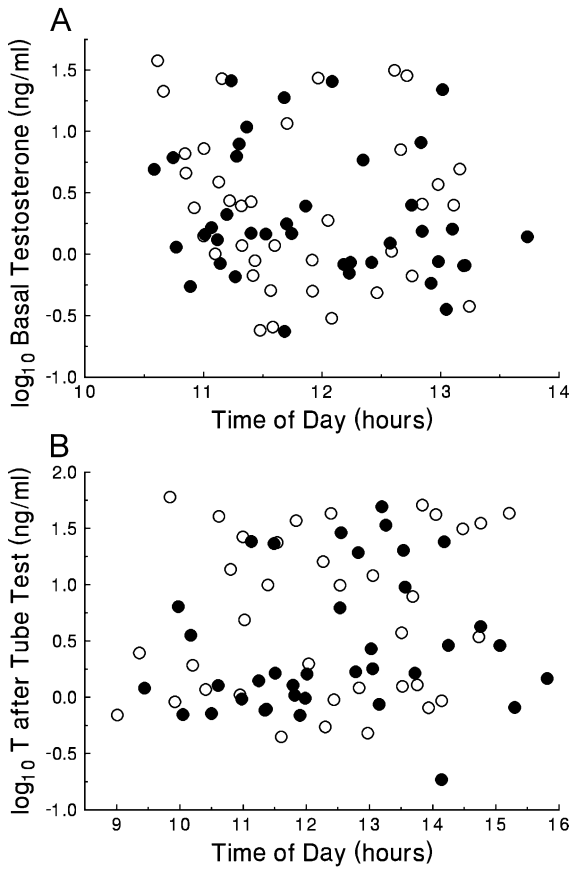

Fig. 2. Lack of relationship between testosterone levels and time of day (closed circles are mice from selected lines, open circles are control lines). (A) Log baseline (n = 38 selected, 38 control); (B) log post-Day 1 tube test (n = 37 selected, 36 control). Also note that testosterone levels did not differ significantly between selected and control lines (see Table 2).

(at time of behavioral tests, mean = 23.7 g, S.D. = 2.31, range = 16.2–27.6 g, $N = 41$).

3.2. Testis size

Sizes of right and left testes were highly correlated ($n = 76$: mass $r = .945$; diameter $r = .814$). For both measures, right testes were, on average, larger than left (mass: paired $t = 6.31$, $P < .0005$; diameter: paired $t = 5.59$, $P < .0005$), a result consistent with previous studies of house mice and other mammals [52]. Mean testis mass and mean testis diameter were also highly correlated ($n = 76$, $r = .820$).

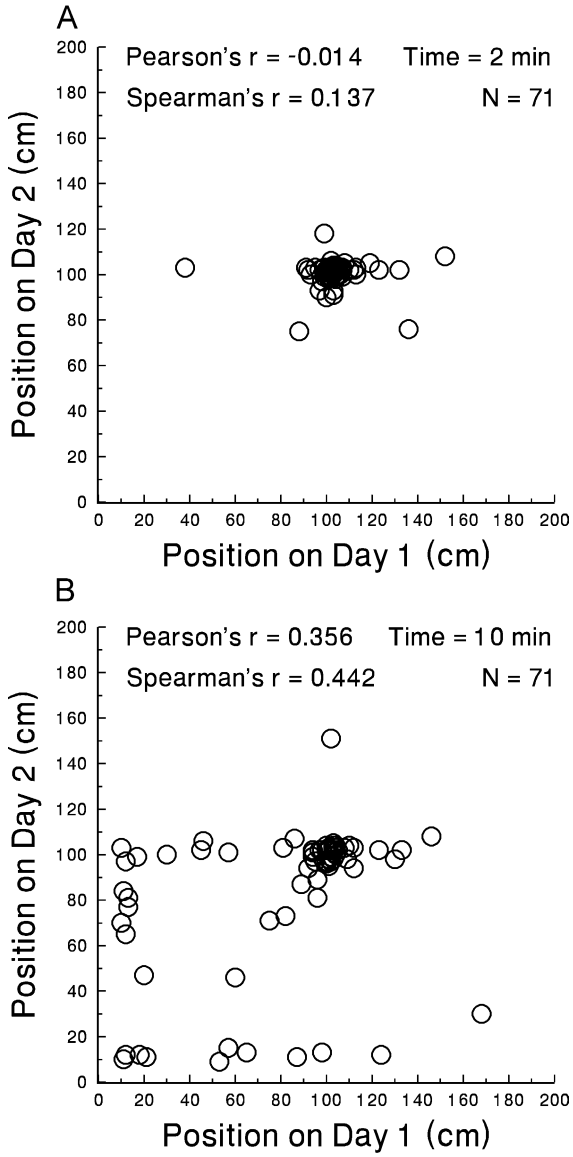

Fig. 3. Relationship between position of focal mouse in the tube-test apparatus on trial days 1 and 2. Middle of tube was at 100 cm; positions between 0 and 100 cm indicate a focal mouse that advanced (was more aggressive than its opponent). (A) Positions at 2 min; (B) positions at 10 min (end of test).

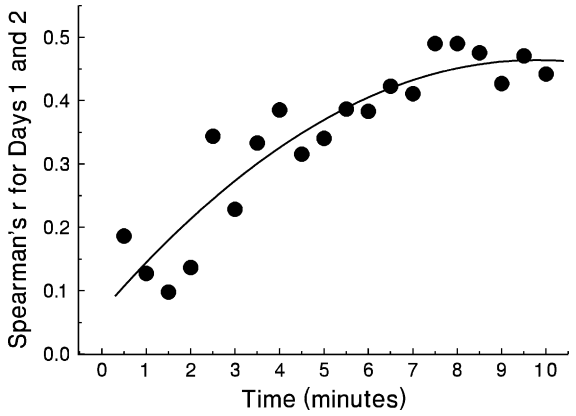

Fig. 4. Correlation between position of focal mouse on the 2 trial days (rank-transformed data) in relation to time of trial. The strength of the relationship increases and then stabilizes by about 7 min; for heuristic purposes, a second-degree polynomial is fitted to the data.

As is apparent from inspection of Fig. 1, several of the 76 measured individuals had unusually small testes for their body mass (see also Methods concerning one omitted individual), and were outliers in statistical analyses. Mice from selected lines tended to have larger testes than did mice from control lines (Fig. 1). The difference was significant for total testis mass when log body mass was not included as a covariate ($P = .0377$), but marginally nonsignificant when body mass was a covariate ($P = .0581$). Mean testis diameter showed a similar trend (Fig. 1), but the differences were less statistically significant (Table 1); this difference in P values probably reflects the lower repeatability of diameter measurements (see below), and consequent reduction in statistical power. When log body was included as a covariate, both testes mass and diameter showed significant differences among replicate lines within linetype (Table 1).

3.3. Testosterone

Baseline plasma testosterone levels (1 week before tube tests) ranged from 0.24 to 37.7 ng/ml, with a mean of 5.62 and a median of 1.57 ng/ml ($n = 76$), and post first tube-test levels ranged from 0.19 to 60.1 ng/ml with a mean of 11.2 and a median of 2.46 ng/ml ($n = 73$). Testosterone levels after the first tube test were significantly higher than baseline levels (for raw values: paired $t = 2.63$, $df = 72$, 2-tailed $P = .011$; for log-transformed values: paired $t = 2.11$, $df = 72$, 2-tailed $P = .038$).

Neither baseline plasma testosterone levels nor levels after the first tube test differed significantly between selected and control lines (Table 2). Baseline levels did, however, show significant differences among replicate lines within linetype. This difference among lines remained significant when the interaction term (see Table 2) was removed from the model ($P = .0293$ for log-transformed data). Post-encounter testosterone levels showed marginally nonsignificant differences among lines.

None of the potential covariates (age, age^2, time of day, time of day^2, delay time for blood sampling) was a significant predictor of testosterone levels, although body mass showed a trend to be positively related to baseline levels (P=.06: Table 2). Fig. 2 demonstrates the lack of relationship with time of day, and also shows the lack of difference between mice from selected and control lines.

We also analyzed individual animals' differences between post-tube test and baseline testosterone levels, as well as the log of the ratio of post-tube/baseline levels. As was the case for the individual measures, selected and control lines did not differ significantly, but replicate lines did (results not shown).

Testosterone levels after the tube test were uncorrelated with baseline testosterone levels (r=.028); results were similar when residuals from the nested ANCOVA models were analyzed (r=.024). Testosterone levels were also uncorrelated with residual testis mass ($r = -.035$ and r=.001 for tube and baseline values, respectively). Finally, adjusted line means for the two measures of testosterone (see notes to Table 2) were uncorrelated with line means for either the mass or diameter of testes (see notes to Table 1).

3.4. Behavior in tube test

Some focal animals advanced quickly, pushing the opponent mouse backwards. Others retreated rapidly, often without apparent provocation from the opponent. Many focal animals exhibited little forward or backward movement and became less active as time elapsed. Physical attacks were rare because the animals could not achieve an attack position inside the tube. Vocalizations audible to the observer were common, but we were unable to determine which animal was vocalizing at a given time.

On both trial days, position of the focal mouse always showed a highly kurtotic and/or skewed distribution: a few individuals either advanced or retreated substantially more than the average mouse. Fig. 3 shows the bivariate distributions at times of 2 and 10 min (end of trial); as can be seen from the 10-min data, the focal mice more often advanced beyond the starting point as compared with the stimulus mice.

For a given time into the trial, positions on the 2 days were generally only weakly related. Correlations between daily positions ranged from $-.014$ to .461 (Pearson's r) or

Table 3
Nested analysis of covariance comparing selected and control lines with respect to position of focal mouse (rank transformed) at various times during a 10-min tube test of aggression administered on 2 consecutive days

Time (min)	N	Significance level								
		Line type	Line	Log body mass of		Age	Age^2	Time of day	Time of day^2	Interaction
				Focal mouse	Stimulus mouse					
Day 1										
0.5	75	.4934	.0122	.7874	.0377	.3846	.1109	.2872	.2973	none
0.5	73	.4509	.0020	.3077	.0327	.0708	.3956	.1617	.1148	none
1.0	75	.3472	.0360	.9485	.1925	.2900	.0935	.2691	.5610	none
1.5	75	.5960	.0249	.5147	.0760	.3320	.2725	.3511	.6153	[a]
2.0	75	.6838	.0410	.5687	.2690	.2127	.4439	.2739	.5847	[b]
5.0	75	.8191	.0177	.2439	.3815	.6383	.9806	.5494	.4794	[c]
10.0	75	.7737	.0171	.8773	.2023	.0440	.4606	.2359	.4862	[d]
10.0	67	.4422	.0002	.3168	.0773	.0334	.1863	.0636	.2391	[e]
Day 2										
0.5	71	.0013	.9405	.9278	.1504	.2793	.1431	.4524	.1983	none
0.5	68	.0105	.2320	.9765	.0455	.0122	.0119	.5466	.0006	none
1.0	71	.0243	.5978	.6448	.1925	.4911	.0801	.3565	.5026	none
1.0	70	.0315	.2970	.7475	.1186	.1848	.0563	.7970	.0981	none
1.5	71	.0393	.7377	.7577	.6975	.9265	.0661	.7049	.8687	[f]
1.5	69	.0013	.9947	.7637	.4030	.3424	.0121	.8103	.1289	[g]
2.0	71	.0089	.9722	.4153	.2302	.2215	.3333	.4287	.2391	none
2.0	69	.0187	.7244	.3424	.1530	.1314	.1447	.8182	.0599	none
5.0	71	.4810	.4332	.2805	.6786	.0906	.4264	.4622	.3961	none
10.0	69	.1141	.7535	.8828	.7162	.0819	.9402	.2962	.2907	none
10.0	69	.1659	.5221	.5775	.4209	.0236	.8872	.3416	.1567	none

[a] Linetype × Log Stimulus Mass P=.5779, Line × Log Stimulus Mass P=.0239.
[b] Linetype × Log Stimulus Mass P=.6649, Line × Log Stimulus Mass P=.0397.
[c] Linetype × Log Focal Mass P=.8316, Line × Log Focal Mass P=.0192.
[d] Linetype × Log Focal Mass P=.5291, Line × Log Focal Mass P=.0086; Linetype × Log Stimulus Mass P=.2511, Line × Log Stimulus Mass P=.0230; Linetype × Age P=.7572, Line × Age P=.0087.
[e] Linetype × Log Focal Mass P=.8316, Line × Log Focal Mass P=.0192; Linetype × Log Stimulus Mass P=.2110, Line × Log Stimulus Mass P=.0003; Linetype × Age P=.7819, Line × Age P=.0001.
[f] Linetype × Log Focal Mass P=.0423, Line × Log Focal Mass P=.7465.
[g] Linetype × Log Focal Mass P=.0016, Line × Log Focal Mass P=.9946.

Table 4
Pairwise Pearson correlations of plasma T levels and of testes mass with position in tube test at different times (rank transformed)

	0.5 min	2 min	5 min	10 min	
Day 1					
Log baseline testosterone	.08	.28 *	.15	−.04	*r*
	72	74	74	67	*n*
	.51	.01	.20	.78	*P*
Log post-tube-test testosterone	−.14	−.07	−.06	−.07	*r*
	71	72	72	65	*n*
	.26	.58	.64	.55	*P*
Log testes mass	.06	−.16	−.10	.08	*r*
	68	70	70	62	*n*
	.64	.20	.41	.53	*P*
Day 2					
Log baseline testosterone	.23	.33 *	.23	.13	*r*
	67	68	70	68	*n*
	.07	.01	.06	.30	*P*
Log post-tube-test testosterone	.05	.04	.12	−.04	*r*
	66	67	69	67	*n*
	.71	.76	.31	.76	*P*
Log testes mass	.16	.22	.04	−.14	*r*
	63	64	66	64	*n*
	.22	.08	.72	.29	*P*

All variables are residuals from nested analyses of covariance (with reduced sample sizes, as presented in Tables 1–3).
Values are correlation, *n*, 2-tailed significance.
*$P < 0.05$.

.098 to .490 (Spearman's *r*), with stronger correlations occurring for later times (Fig. 4). Because behavior of the focal mouse on the 2 trial days was only weakly consistent it is perhaps best considered as two different traits (for a similar argument concerning forced sprint running speed in Hsd:ICR mice, see Ref. [30]). Data for the two trials were, therefore, analyzed separately.

Position of the focal mouse on Day 1 did not differ significantly between selected and control lines but did differ among replicate lines (Table 3). On Day 2, however, mice from selected lines moved towards their opponent less than did mice from control lines, at least for the first few min of the test. By 5 min into the test, differences between selected and control mice were no longer statistically significant (Table 3).

Position of the focal mouse was positively correlated with baseline plasma testosterone levels, especially for the 2-min time during the second trial: individuals with higher testosterone were less likely to advance towards their opponent (Table 4). Position of the focal mouse was uncorrelated with plasma testosterone levels after the first-day's tube test (Table 4).

4. Discussion

House mice from four replicate lines selected for high levels of voluntary wheel-running behavior were less likely to advance towards their opponent during the second of two daily tube tests of dominance, tended to be smaller in body mass (see also Refs. [26,32]), but tended to have larger testes, as compared with mice from four randomly bred control lines. Selected and control lines did not differ significantly in plasma testosterone levels measured either under baseline conditions or immediately following the first day's tube test.

Replicate lines within linetype showed several significant differences: body mass, relative testis size, position during the first tube-test encounter, and baseline testosterone. Such differences indicate genetic divergence resulting from one or more of four sources (e.g. see Refs. [1,4,6]): founder effects when our lines were initially established, random genetic drift, unique mutations or multiple genetic responses to selection (which are contingent on the genetic variation that exists in each generation). Our data do not allow us to discriminate among these processes, which are rarely studied with respect to behavioral (but see Refs. [57,58]) or endocrinological phenotypes. In any case, the differences among lines may make them useful for comparative studies of other phenotypes.

At the level of differences among individual mice (residuals from nested ANCOVA models), baseline testosterone correlated with position of the focal mouse during the tube test (higher T levels were found in individuals that were less likely to advance), but post-encounter testosterone level did not exhibit such a relationship. The unexpected direction of the association between the measure of dominance and baseline testosterone (e.g. see Ref. [59]) may reflect an alternate coping strategy (see below).

4.1. Plasma testosterone levels and testis size

Two limitations of our data should be noted. First, whereas the level of plasma testosterone is important in assessing the hormone's behavioral influences, it may be of equal importance to assess the sensitivity and density of testosterone receptors in various tissues. Individuals with similar levels of plasma testosterone may still differ in behavior because of different sensitivities and/or densities of testosterone receptors. Consistent with such possibilities, mouse strains have been shown to differ in the sensitivity of seminal vesicles to testosterone [60]. Second, house mice exhibit an episodic testosterone secretion, causing rapid changes in plasma testosterone level [45,46]. As we only took a single sample of blood each time we drew from a given mouse, our data may not reflect 'typical' testosterone level for a given individual.

In principle, the foregoing limitations may have reduced our ability to detect differences among groups and/or correlations at the level of individual variation. Nevertheless, our values for plasma testosterone levels are consistent with those reported by other authors (e.g. see Refs. [45,61,16 (wild mice)]). The fact that we were able to detect statistically significant differences among replicate lines (Table 2), and correlations at the level of individual variation (Table 4), indicates that our testosterone measurements

were sensitive enough to detect effects when they existed. Moreover, we were able to detect significant linetype effects for other traits. Thus, the lack of testosterone differences (Table 2) between selected and control lines is probably a reliable result. In future studies, we plan to examine testosterone levels as indexed by the pooling of multiple blood samples (e.g. see Ref. [62]).

Although selected and control lines did not differ in plasma testosterone level, they tended to differ in testis size, at least when outlier individuals were excluded from the analysis (Fig. 1, Table 1). Moreover, mass of testes did not correlate with plasma testosterone concentration at the level of individual variation, nor at the level of line means (as reported in Tables 1 and 2). Future studies should examine relative proportions of Leydig and Sertoli cells in relation to testis size and testosterone secretory capacity. Leydig cells produce testosterone, and the density of these cells may determine secretory capacity (e.g. see Ref. [15]). Sertoli cell number, however, may be the primary mode through which genes regulate testis size in house mice [63]. If so, then size of testes would not necessarily correlate with testosterone output. Alternatively, a better approach may be to measure seminal vesicle mass, as this tissue changes in response to, and may reflect variations in, testosterone level [60].

4.2. Behavior in tube test

Behavior in the tube test on Day 1 correlated weakly with behavior on Day 2. This suggests that tests on successive days may be measuring different aspects of behavior, which is unsurprising (see also [30]). Day 1 provided the animal with a novel environment and unfamiliar opponent, but on Day 2 the animal was experienced with both the situation and the opponent. On Day 2, mice from the selected lines were significantly more likely to retreat as compared with mice from control lines. One possible interpretation of these results was suggested by an anonymous reviewer of this manuscript. On Day 1, focal mice may have learned the size of their opponent. On Day 2, they may have behaved in accordance with this knowledge. Given that mice from selected lines were smaller than those from control lines, they might hence have been more likely to retreat on Day 2. However, as noted in the Results, stimulus mice were several grams lighter than focal mice from either the selected or control lines, which differed by about only 2 g in body mass.

The relationship between behavior in the tube test, typically considered to indicate dominance [39–43], and aggression per se is unclear. However, if retreat in the tube test reflects lower aggression, then our selected lines would appear to be less aggressive than the random-bred control lines. How does this result compare with correlated responses observed in other selection experiments? Lagerspetz's [10] high-aggression line showed higher locomotor activity in rotating-drum cages than did the low-selected line (reviewed in Ref. [12]), which seems inconsistent with our results; however, rotating-drum cages are not directly comparable to those we used, which constitute a home cage attached to a separate wheel [7,27,29,32]. In wild house mice selected for short or long attack latency (review in Ref. [13]), the former (more aggressive) build larger nests, but wheel-running behavior has not been reported. In Lynch's [2] bidirectionally selected nesting lines, low-nesting lines run more on wheels than do either control or high-selected lines (i.e. her control and high lines are similar in wheel running: 9), whereas high-nesting lines are more aggressive than either control or low lines (i.e. her control and low lines are similar in aggression: 8). Thus, our high wheel-running lines (see also Ref. [28]) might appear to be behaviorally convergent with Lynch's low-nesting lines: both exhibit high wheel running, low nesting, and low aggression. However, we used a neutral testing situation involving a tube, whereas Sluyter et al. [8] employed a resident–intruder test in a cage, and these paradigms may yield different results. In addition, it should be noted that a consideration of all six of the Lynch lines does not strongly support an association between wheel running and aggression.

4.3. Individual variation and alternate coping strategies

Many researchers have found that testosterone is positively related to aggression, and sometimes dominance, in mice (see also Introduction). Compaan et al. [17], for example, found that testosterone administered to intact mice increased aggression in both sexes. Matochik et al. [18] observed that testosterone administered to castrated mice resulted in a normal frequency of aggressive behavior. Similarly, in castrated male mice from lines selected for short attack latency and from control lines, testosterone treatment led to reductions in attack latency [16].

Given the foregoing studies, and presuming a relationship between aggression and dominance, then an inverse relationship between baseline testosterone and tube dominance (Table 4) may seem surprising. Note, however, that none of the above-cited studies tested for correlations between natural (unmanipulated) levels of testosterone and measures of either dominance or aggression. Moreover, other studies call into question the generality of a positive relationship between testosterone and aggression or dominance. Leshner and Politch [64], for example, demonstrated that castrated and sham-castrated mice displayed no difference in time to submission or in the number of aggressive acts needed to elicit their submission, whereas manipulation of the pituitary–adrenocortical hormones did influence these behaviors. Similarly, Bhasin et al. [23] found no behavioral differences between groups of men administered testosterone and those administered vehicle only. On the other hand, Salvador et al. [65] have recently shown positive associations between testosterone and offensive behaviors in human males engaging in judo contests (see also review and discussion in Ref. [59], and following papers).

But what could account for a *negative* relationship between a measure of aggression and plasma testosterone levels at the level of individual variation? We speculate that our finding may reflect an 'alternative' or 'individual coping strategy' [66] employed by relatively aggressive individuals. Our tube test constituted a confining situation in which attack was difficult. As noted by Boissy [67], behavioral patterns indicative of fear can be contradictory, including both active defense (attack, threat) and active avoidance (flight, escape). Aggressive individuals, faced with a situation in which actual attack was difficult (because of the confinement of the tube), may have instead attempted to remove themselves from the source of stress, which is considered a form of 'active manipulation' [66]. More generally, as reviewed by Henry [68] and Boissy [67], an animal's perception of its situation can have a large influence on its response to 'stressful' situations. According to this model, an aggressive individual in our tube test may have perceived itself not to be in control, experienced activation of the hypothalamic–pituitary–adrenal (HPA) axis, a rise in glucocorticoids, and an increased probability of flight (retreat). Benus et al. [69] found that mice from a high-aggression selected line were more likely than low-aggression mice to attempt escape from an encounter that occurred on another mouse's territory, and that once the aggressive animal had experience with the route of escape (as in Day 2 of our tube test), it used it more rapidly than in the initial encounter.

Arguing against the foregoing speculation is the fact that mice from our selected lines were less likely to advance towards their opponent than were control mice on Day 2 of the tube test, yet did not differ significantly in testosterone levels. However, a lack of correspondence between trait correlations from comparisons of selected and control lines and from individual variation within lines (residuals from nested ANCOVA models), or from mean values for each replicate line, is not uncommon (e.g. see Ref. [6]). Indeed, we have identified such differences in these mice with respect to the relationship between amount of wheel running and litter size [70] (see also Ref. [71] on food consumption). These differences may have several causes. First, trait variation and covariation among individuals within lines are caused by both genetic and environmental variation, whereas trait variation between sets of selected and control lines are mainly the result of differences in allele frequencies at loci that affect the trait that has been under selection (i.e. wheel running in the present experiment). Allelic variation at these loci may or may not have pleiotropic effects on other phenotypic traits (e.g. behavior in a tube test, testosterone levels); thus, the pattern of trait 'correlation' apparent from linetype versus individual comparisons need not be congruent. Second, correlations based on mean values of each line ($N = 8$ in our experiment) may differ in still other ways because of random genetic effects (e.g. see Ref. [71]). Finally, the statistical power for detecting correlations at the individual level (N approximately 70) may be higher than for testing differences between selected and control lines ($N = 8$). In future studies, we plan to compare selected and control lines with respect to other measures of aggression and plasma corticosterone levels (see also Ref. [72]), in various testing situations.

Acknowledgements

We would like to thank W. Saltzman for comments on the manuscript. Research was supported by NSF grants IBN-9111185 and IBN-9728434 to T.G. and an Edward A. and Anna G. Birge Scholarship to K.F.K. We would also like to thank S.J. Davis and S.S. Newton for aid in animal care and data collection, and G. Scheffler and D.J. Wittwer for aid in performing testosterone assays.

References

[1] Garland T. Selection experiments: an underutilized tool in biomechanics and organismal biology. In: Gasc J-P, Casinos A, Bels VL, editors. Vertebrate biomechanics and evolution, vol. I. Oxford, UK: Bios Scientific Publishers, 2002. p. 23–56.
[2] Lynch CB. Evolutionary inferences from genetic analyses of cold adaptation in laboratory and wild populations of the house mouse. In: Boake CRB, editor. Quantitative genetic studies of behavioral evolution. Chicago: University of Chicago Press, 1994. p. 279–301.
[3] Gibbs AG. Laboratory selection for the comparative physiologist. J Exp Biol 1999;202:2709–18.
[4] Henderson ND. Spurious associations in unreplicated selected lines. Behav Genet 1997;27:145–54.
[5] Finch CE, Rose MR. Hormones and the physiological architecture of life history evolution. Q Rev Biol 1995;70:1–52.
[6] Falconer DS, Mackay TFC. Introduction to quantitative genetics. 4th ed. UK: Longman, Essex, 1996.
[7] Swallow JG, Carter PA, Garland T. Artificial selection for increased wheel-running behavior in house mice. Behav Genet 1998;28:227–37.
[8] Sluyter F, Bult A, Lynch CB, van Oortmerssen GA, Koolhaas JM. A comparison between house mouse lines selected for attack latency or nest-building: evidence for a genetic basis of alternative behavioral strategies. Behav Genet 1995;25:247–52.
[9] Bult A, Hiestand L, van der Zee EA, Lynch CB. Circadian rhythms differ between selected mouse lines: a model to study the role of vasopressin neurons in the suprachiasmatic nuclei. Brain Res Bull 1993;32:623–7.
[10] Lagerspetz KMJ. Genetic and social causes of aggressive behaviour in mice. Scand J Psychol 1961;2:167–73.
[11] Selander RK, Kvist BM. Open-field parameters and maze learning in aggressive and nonaggressive male mice. Percept Mot Skills 1991;73:811–24.
[12] Sandnabba NK. Selective breeding for isolation-induced intermale aggression in mice: associated responses and environmental influences. Behav Genet 1996;26:477–88.
[13] Sluyter F, van Oortmerssen GA, de Ruiter AJH, Koolhaas JM. Aggression in wild house mice: current state of affairs. Behav Genet 1996;25:489–96.
[14] Benus RF, Rondigs M. Patterns of maternal effort in mouse lines bidirectionally selected for aggression. Anim Behav 1996;51:67–75.
[15] de Ruiter AJH, Koolhaas JM, Keijser JN, van Oortmerssen GA, Bohus B. Differential testosterone secretary capacity of the testes of

aggressive and nonaggressive house mouse during ontogeny. Aggress Behav 1992;18:149–57.
[16] van Oortmerssen GA, Dijk DJ, Schuurman T. Studies in wild house mice: II. Testosterone and aggression. Horm Behav 1987;21:139–52.
[17] Compaan JC, Van Wattum G, de Ruiter AJH, van Oortmerssen GA, Koolhaas JM, Bohus B. Genetic differences in female house mice in aggressive response to sex steroid hormone treatment. Physiol Behav 1993;54:899–902.
[18] Matochik JA, Sipos ML, Nyby JG, Barfield RJ. Intercranial androgenic activation of male-typical behaviors in house mice: motivation verses performance. Behav Brain Res 1994;60:141–9.
[19] Bronson FH, Nguyen KQ, De La Rosa J. Effect of anabolic steroids on behavior and physiological characteristics of female mice. Physiol Behav 1996;59:49–55.
[20] Zielinski WJ, Vandenbergh JG. Testosterone and competitive ability in male house mice, *Mus musculus*: laboratory and field studies. Anim Behav 1993;45:873–91.
[21] O'Keane JC, Brien TG, Hooper ACB, Graham A. Testicular activity in mice selected for increased body weight. Andrologia 1986;18:190–5.
[22] Hooper ACB, Brien TG, Lawlor PG. The effects of orchidectomy and the role of testosterone in determining the growth of male mice selected for increased body weight. Andrologia 1986;18:509–15.
[23] Bhasin S, Storer TW, Berman N, Callegari C, Clevenger B, Phillips J, Bunnell TJ, Tricker R, Shirazi A, Casaburi R. The effects of supraphysiologic doses of testosterone on muscle size and strength in normal men. N Engl J Med 1996;335:1–7.
[24] van Oortmerssen GA, Benus RF, Sluyter F. Studies on wild house mice: IV. On the heredity of testosterone and readiness to attack. Aggress Behav 1992;18:143–8.
[25] Harris JA, Vernon PA, Boomsma DI. The heritability of testosterone: a study of Dutch adolescent twins and their parents. Behav Genet 1988;28:165–71.
[26] Swallow JG, Koteja P, Carter PA, Garland T. Artificial selection for increased wheel-running activity in house mice results in decreased body mass at maturity. J Exp Biol 1999;202:2513–20.
[27] Swallow JG, Garland T, Carter PA, Zhan W-Z, Sieck GC. Effects of voluntary activity and genetic selection on aerobic capacity in house mice (*Mus domesticus*). J Appl Physiol 1998;84:69–76.
[28] Carter PA, Swallow JA, Davis SJ, Garland T. Nesting behavior of house mice (*Mus domesticus*) selected for increased wheel-running activity. Behav Genet 2000;30:85–94.
[29] Koteja P, Garland T, Sax JK, Swallow JG, Carter PA. Behaviour of house mice artificially selected for high levels of voluntary wheel running. Anim Behav 1999;58:1307–18.
[30] Dohm MR, Hayes JP, Garland T. Quantitative genetics of sprint running speed and swimming endurance in laboratory house mice (*Mus domesticus*). Evolution 1996;50:1688–701.
[31] Carter PA, Garland T, Dohm MR, Hayes JP. Genetic variation and correlations between genotype and locomotor physiology in outbred laboratory house mice (*Mus domesticus*). Comp Biochem Physiol A 1999;123:155–62.
[32] Koteja P, Swallow JG, Carter PA, Garland T. Energy cost of wheel running in house mice: implications for coadaptation of locomotion and energy budgets. Physiol Zool 1999;72:238–49.
[33] Maxson SC. Methodological issues in genetic analysis of an agonistic behavior (offense) in male mice. In: Goldwitz D, Wahlsten D, Wimer RE, editors. Techniques for the genetic analysis of brain and behavior. Amsterdam: Elsevier, 1992. p. 349–73.
[34] Roubertoux PL, Le Roy I, Mortaud S, Perez-Diaz F, Tordjman S. Measuring aggression in the mouse. In: Crusio WE, Gerlai RT, editors. Handbook of molecular-genetic techniques for brain and behavior research. Amsterdam: Elsevier, 1999. p. 696–709.
[35] Hyde JS, Ebert PD. Correlated response in selection for aggressiveness in female mice: I. Male aggressiveness. Behav Genet 1976;6:421–7.
[36] Burright RG, Freeman MJ, Donovick PJ. Repeated tests of intermale aggression in mice (*Mus musculus*) are influenced by test conditions. J Comp Psychol 1988;102:303–5.
[37] Rodgers RJ, Cole JC. Influence of social isolation, gender, strain and prior novelty on plus-maze behavior in mice. Physiol Behav 1993;54:729–36.
[38] Gray S, Hurst JL. The effects of cage cleaning on aggression within groups of male laboratory mice. Anim Behav 1995;49:821–6.
[39] Lindzey G, Winston H, Manosevitz M. Social dominance in inbred mouse strains. Nature 1961;191:474–6.
[40] Wimer RE, Fuller JL. Patterns of behavior. In: Green EL, editor. Biology of the laboratory mouse. 2nd ed. New York: McGraw-Hill, 1966. p. 629–53.
[41] Messeri P, Eleftheriou BE, Oliverio A. Dominance behavior: a phylogenetic analysis in the mouse. Physiol Behav 1975;14:53–8.
[42] Benton D, Dalrymple-Alford JC, Brain PF. Comparisons of measures of dominance in the laboratory mouse. Anim Behav 1980;28:1274–9.
[43] van de Weerd HA, van den Broek FAR, Beynen AC. Removal of vibrissae in male mice does not influence social dominance. Behav Process 1992;27:205–8.
[44] Cunliffe-Beamer TL. Biomethodology and surgical techniques. In: Fosler HL, Small JD, Fox JG, editors. The mouse in biomedical research, vol. 3. New York: Academic Press, 1983. p. 401–30.
[45] Bartke A, Dalterio S. Evidence for episodic secretion of testosterone in laboratory mice. Steroids 1975;26:749–56.
[46] Coquelin A, Desjardins C. Luteinizing hormone and testosterone secretion in young and old mice. Am J Physiol 1982;243:E257–63.
[47] McGuill MW, Rowan AN. Biological effects of blood loss: implications for sampling volumes and techniques. ILAR News 1989;31:5–18.
[48] Abraham GE, Buster JE, Lucal LA, Corrales PC, Teller RC. Chromatographic separation of steroid hormones for use in radioimmunoassay. Anal Lett 1972;5:509–17.
[49] Munro CJ, Lasley BL. Non-radiometric methods for immunoassay of steroid hormones. In: Albertson BD, Haseltine FP, editors. Non-radiometric assays: technology and application in polypeptide and steroid hormone detection. New York: Alan R. Liss, 1988. p. 289–329.
[50] Saltzman W, Schultz-Darken NJ, Scheffler G, Wegner FH, Abbott DH. Social and reproductive influences on plasma cortisol in female marmoset monkeys. Physiol Behav 1994;56:801–10.
[51] Rodbard D. Statistical quality control and routine data processing for radioimmunoassays and immunoradiometric assays. Clin Chem 1974;20:1255–70.
[52] Kenagy GJ, Trombulak SC. Size and function of mammalian testes in relation to body size. J Mammal 1986;67:1–22.
[53] Swallow JG, Koteja P, Carter PA, Garland T. Food consumption and body composition in mice selected for high wheel-running activity. J Comp Physiol B 2001;171:651–9.
[54] Conover WJ, Iman RL. Rank transformations as a bridge between parametric and nonparametric statistics. Am Stat 1981;35:124–9.
[55] Potvin C, Roff DA. Distribution-free and robust statistical methods: viable alternatives to parametric statistics? Ecology 1993;74:1617–28.
[56] Benirschke K, Garner FM, Jones TC. Pathology of laboratory animals, vol. I. New York: Springer-Verlag, 1978.
[57] Smith RH. Behavioral measures of drift in *Mus musculus*. Behav Genet 1985;15:483–97.
[58] Bult A, Lynch CB. Multiple selection responses in house mice bidirectionally selected for thermoregulatory nest-building behavior: crosses of replicate lines. Behav Genet 1996;26:439–46.
[59] Mazur A, Booth A. Testosterone and dominance in men. Behav Brain Sci 1998;21:353–97 (including Open Peer Commentaries and Author's Response).
[60] Bartke A. Increased sensitivity of seminal vesicles to testosterone in a mouse strain with low plasma testosterone levels. J Endocrinol 1974;60:145–8.
[61] Eisen EJ, Johnson BH. Correlated responses in male reproductive

traits in mice selected for litter size and body weight. Genetics 1981;99:513–24.
[62] Goldzieher JW, Dozier TS, Smith KD, Steinberger E. Improving the diagnostic reliability of rapidly fluctuating plasma hormone levels by optimized multiple-sampling techniques. J Clin Endocrinol Metab 1976;43:824–30.
[63] Chubb C. Genes regulating testis size. Biol Reprod 1992;47:29–36.
[64] Leshner AI, Politch JA. Hormonal control of submissiveness in mice: irrelevance of the androgens and relevance of the pituitary–adrenal hormones. Physiol Behav 1979;22:531–4.
[65] Salvador A, Suay F, Martinez-Sanchis S, Simon VM, Brain PF. Correlating testosterone and fighting in male participants in judo contests. Physiol Behav 1999;68:205–9.
[66] Benus RF, Bohus B, Koolhaas JM, van Oortmerssen GA. Heritable variation for aggression as a reflection of individual coping strategies. Experientia 1991;47:1008–19.
[67] Boissy A. Fear and fearfulness in animals. Q Rev Biol 1995;70: 165–91.
[68] Henry JP. Biological basis of the stress response. Integr Physiol Behav Sci 1992;27:66–83.
[69] Benus RF, Koolhaas JM, van Oortmerssen GA. Individual differences in behavioral reaction to a changing environment in mice and rats. Behaviour 1987;100:105–23.
[70] Girard I, Swallow JG, Carter PA, Koteja P, Rhodes JS, Garland T. Maternal-care behavior and life-history traits in house mice (*Mus domesticus*) artificially selected for high voluntary wheel-running activity. Behav Process 2002;57:37–50.
[71] Koteja P, Swallow JG, Carter PA, Garland T. Maximum cold-induced food consumption in mice selected for high locomotor activity: implications for the evolution of endotherm energy budgets. J Exp Biol 2001;204:1177–90.
[72] Girard I, Garland T. Plasma corticosterone response to acute and chronic voluntary exercise in female house mice. J Appl Physiol 2002;92:1553–61.