

How to run far: multiple solutions and sex-specific responses to selective breeding for high voluntary activity levels

Theodore Garland, Jr, Scott A. Kelly, Jessica L. Malisch, Erik M. Kolb, Robert M. Hannon, Brooke K. Keeney, Shana L. Van Cleave and Kevin M. Middleton

Proc. R. Soc. B 2011 **278**, 574-581 first published online 1 September 2010
doi: 10.1098/rspb.2010.1584

Supplementary data

["Data Supplement"](#)

<http://rspb.royalsocietypublishing.org/content/suppl/2010/08/27/rspb.2010.1584.DC1.html>

References

[This article cites 48 articles, 17 of which can be accessed free](#)

<http://rspb.royalsocietypublishing.org/content/278/1705/574.full.html#ref-list-1>

Subject collections

Articles on similar topics can be found in the following collections

[behaviour](#) (1782 articles)
[biomechanics](#) (190 articles)
[evolution](#) (2404 articles)

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)

How to run far: multiple solutions and sex-specific responses to selective breeding for high voluntary activity levels

Theodore Garland, Jr^{*}, Scott A. Kelly[†], Jessica L. Malisch,
Erik M. Kolb, Robert M. Hannon, Brooke K. Keeney,
Shana L. Van Cleave and Kevin M. Middleton[‡]

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

The response to uniform selection may occur in alternate ways that result in similar performance. We tested for multiple adaptive solutions during artificial selection for high voluntary wheel running in laboratory mice. At generation 43, the four replicate high runner (HR) lines averaged 2.85-fold more revolutions per day as compared with four non-selected control (C) lines, and females ran 1.11-fold more than males, with no sex-by-linetype interaction. Analysis of variance indicated significant differences among C lines but not among HR for revolutions per day. By contrast, average speed varied significantly among HR lines, but not among C, and showed a sex-by-linetype interaction, with the HR/C ratio being 2.02 for males and 2.45 for females. Time spent running varied among both HR and C lines, and showed a sex-by-linetype interaction, with the HR/C ratio being 1.52 for males but only 1.17 for females. Thus, females (speed) and males (speed, but also time) evolved differently, as did the replicate selected lines. Speed and time showed a trade-off among HR but not among C lines. These results demonstrate that uniform selection on a complex trait can cause consistent responses in the trait under direct selection while promoting divergence in the lower-level components of that trait.

Keywords: adaptation; experimental evolution; genetic drift; locomotion; parallel evolution; trade-off

1. INTRODUCTION

A long-standing principle in evolutionary biology holds that a given type of natural selection, often influenced by other evolutionary processes (e.g. random genetic drift, sexual selection), may lead to different outcomes (e.g. [1–3]). When a given type of selection (e.g. natural selection favouring increased water efficiency in deserts) acts on different kinds of organisms (e.g. birds versus mammals), different adaptive responses or different degrees of response are not so surprising because the starting point (e.g. avian versus mammalian kidney function and their respective genomes) is likely to influence (possibly ‘constrain’) the subsequent adaptive trajectory (e.g. [4–6]). For example, whereas small mammals can adapt to extreme temperatures in part by burrowing, this is less possible for large mammals. Birds can move long distances on a seasonal or even daily basis to avoid low temperatures, but tortoises cannot.

More surprising is the observation that uniform selection from a given starting point often leads to multiple ‘solutions’ to the adaptive ‘problem’. For example,

Futuyma ([7], pp. 256–257) notes that plumage patterns of newly hatched ptarmigans and grouse ‘doubtless provide cryptic protection, but it is quite possible that the differences among the species have little adaptive significance’. If the ancestral form also possessed a cryptic pattern, then the current differences among species may have arisen by random genetic drift, within some confines set by ongoing natural selection. Alternatively, a cryptic pattern may have evolved multiple times. Well-designed phylogenetically based comparative analyses can sometimes distinguish between such alternatives [8–10], but they cannot provide direct observations of the adaptive process.

A literature review indicated that conspecific populations have diverged under putatively similar selection for a variety of traits (table 1 in reference [11]). However, in many of these cases the ‘replicate’ starting populations may already have differed genetically to some extent prior to experiencing the uniform selection, either inherently or because of founder effects (see also [12,13]).

Replicated selection experiments of various types provide a powerful way to explore multiple solutions that may occur in response to relatively well-defined types of selection [13–18]. As noted by Mayr ([2], p. 1505), ‘Breeders and students of natural selection have discovered again and again that independent parallel lines exposed to the same selection pressures will respond at different rates and with different effects, none of them predictable’ (but see [19]). Most commonly, experimental evolution

^{*} Author for correspondence (tgarland@ucr.edu).

[†] Present address: Departments of Nutrition and Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.

[‡] Present address: Department of Biology, California State University, San Bernardino, CA 92407, USA.

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2010.1584> or via <http://rspb.royalsocietypublishing.org>.

approaches begin with replicate populations derived from the same genetic stock, i.e. lines whose 'gene pools' are initially identical except for sampling (founder) effects.

As one example, Weber *et al.* [20] compared five replicate pairs of *Drosophila melanogaster* lines that had been divergently selected with respect to wing shape. They found that 29 loci showed consistent expression differences in all five paired comparisons. However, for a pair of lines that derived from a different base population the significant loci were almost entirely different. Thus, the gene pool of the starting population influenced the evolutionary outcome at the level of gene expression. Weber *et al.* [20] did not indicate whether the replicate selected lines showed consistent divergences with respect to the trait under selection, i.e. wing shape.

Many other examples of multiple genetic paths underlying parallel evolution are found in the microbial experimental evolution literature (e.g. reviews in [21–23]). For example, in *Escherichia coli*, on either lactate or glycerol minimal media, growth phenotypes at the evolutionary endpoint were convergent and repeatable, but had different underlying gene expression states [24]. In *E. coli* evolved in a glucose environment, a candidate gene approach demonstrated widespread parallelism in the direct response to selection, but considerable heterogeneity in (pleiotropic) mutant effects in five novel environments [25].

The purpose of the present paper was to compare replicate lines of laboratory house mice that have been selectively bred for high voluntary wheel running [26]. Since reaching an apparent selection limit around generation 16, the four high runner (HR) lines have been running 2.5–3.0-fold more revolutions per day as compared with four non-selected control (C) lines ([27,28] and references therein). These lines have been the subject of numerous studies aimed at identifying the motivational and physiological foundations of elevated activity levels (e.g. [27–33]).

Here, we first compare the HR and C lines with respect to divergence in mean values for wheel revolutions per day—the trait under direct selection—and its components, time spent running and average running speed. We also analyse body mass, which has evolved as a correlated response to selection [31,34]. We then test for among-replicate line differences in the HR and C mice, analyse the correlation of sex-specific line means, and finally analyse correlations at the level of individual variation within the lines. Most analyses are done separately by sex because of various differences in wheel running as well as body size, and because the potential sex-specificity of (alternate) responses to selection has rarely been considered.

2. MATERIAL AND METHODS

(a) *Animals*

We studied mice from generation 43 of the ongoing selection experiment [15,26–28]. Briefly, this experiment involves eight lines descended from 224 outbred Hsd:ICR house mice (*Mus domesticus*) purchased in 1993. Further details regarding this strain can be found in [35,36]. Each line is perpetuated by at least 10 families each generation, and average litter size is about 10.5 [36]. Offspring are weaned at 21 days of age, housed four per cage (same sex), and then housed

with access to wheels (1.12 m circumference) for 6 days at 6–8 weeks of age. Computerized photocell counters record revolutions at 1 min intervals. Mice are weighed immediately before and after wheel access.

In the four HR lines, breeders are chosen based on the average number of revolutions run on days 5 + 6. The highest-running male and female are chosen from each family, with some second-highest runners used to form backup pairings. In the four C lines, breeders are chosen randomly from within families. This within-family selection causes effective population size to be approximately constant across all eight lines (typical N_e approx. 35 [26]). Sibling matings are never allowed.

In a typical generation, more than 400 HR mice are tested with wheels (4 lines \times 10 families \times 10.5 pups per family). In the four control (C) lines, only two males and two females are chosen (randomly) at weaning, so about 160 control-line mice are tested in a typical generation (4 lines \times 10 families \times 4 pups per family). For the present study, we analysed a generation in which a substantial number of mice had been set aside at weaning for other experiments (e.g. [37,38]), thus leaving 230 HR mice and 157 C mice for wheel testing. These mice were tested in one room with 100 wheels over a period of four weeks.

At all times, mice were maintained on a 12L:12D cycle and provided with water and food (Harlan Teklad Rodent Diet 8604) ad libitum (breeding females were given Teklad S-2335 Mouse Breeder diet: see also [39]).

(b) *Traits*

We analysed body mass prior to wheel testing and the means of days 5 + 6 for total revolutions, the number of 1 min intervals with at least one revolution, and average speed during the active intervals (revolutions per interval; [26]).

(c) *Statistical analyses*

We first compared mean values for the two linetypes (HR versus C) and sexes using SAS procedure mixed (SAS Institute, Cary, NC, USA) with Type-III tests of fixed effects. Random effects were family nested within line, line nested within linetype, and the sex \times line interaction. Each of these covariance components was estimated separately for the HR and C lines (cf. [40]) using the /GROUP = LINE-TYPE option in the RANDOM statement. Because entire families were tested within one of four weekly batches, the family variable also accounted for any among-batch variation. Age was used as a covariate, as was a measure of wheel freeness to rotation. Residuals from these models were retained for the analysis of individual variation within line and within sex. We also repeated the comparisons of means with body mass prior to wheel testing as an additional covariate.

To compute sex-specific least-squares line means and standard errors, we also used SAS procedure mixed, but excluded linetype from the model and treated line as a fixed effect (and retained the covariates other than body mass). Line means were analysed by partial correlation and analysis of covariance (ANCOVA).

We used two approaches to compare the among-line variability of the two linetypes. First, we performed separate analyses of the HR and C lines, treating line as a fixed effect. In these analyses, different results for HR and C lines can be a consequence of the amount of among-line variance (the main quantity of interest) and/or the amount

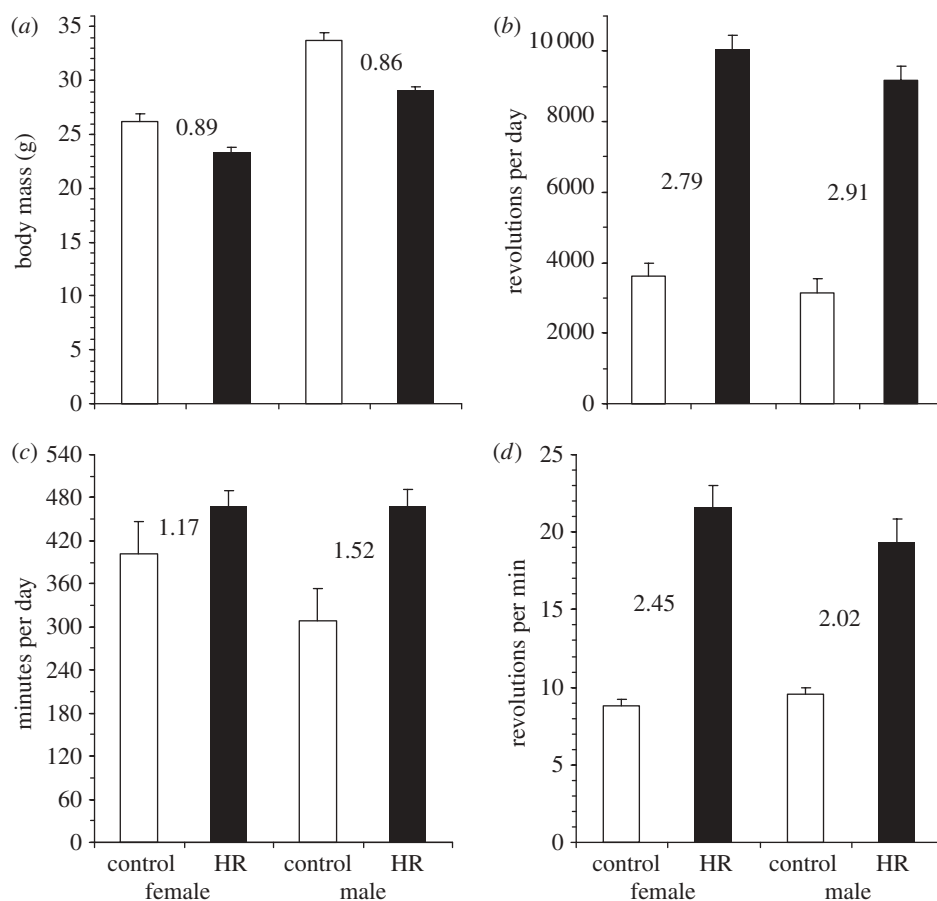


Figure 1. Least-squares means and associated standard errors (corresponding to analyses presented in the electronic supplementary material, table S1 without body mass as a covariate). Numbers indicate ratios of those means for HR/C lines (interaction: (a), $p = 0.0355$; (b), $p = 0.4732$; (c), $p = 0.0142$; (d), $p = 0.0195$).

of among-family variance, which is not directly relevant to the main topic of this paper. Therefore, separately for each sex, we also compared the differences in ln restricted maximum likelihoods for models that estimated separate among-line variance components for the HR and C lines (i.e. six random effects in total) with those that estimated pooled among-line variance components for the HR and C lines. Twice the difference in ln likelihoods of these models should asymptotically follow a χ^2 distribution with one degree of freedom. Linetype was a fixed effect in these models. Separate variance components for HR and C lines were estimated for both among-family and among-individual within family components.

Statistical significance was judged at $p < 0.05$, and all tests were two-tailed.

3. RESULTS

(a) Differences in means between HR and C lines

Statistical comparisons of the mean values for the four HR lines with their four C counterparts can be found in the electronic supplementary material, table S1. As expected from numerous previous studies, HR ran nearly threefold farther on a daily basis than C (figure 1), and females ran significantly farther than males in both HR and C lines, with no sex \times linetype interaction ($p = 0.4732$). The increased running distance of HR lines (figures 1 and 2) was accomplished almost entirely by increased average speeds in females, but for males the HR lines also showed a statistically significant

increase in time spent running (figure 1). Separate analyses by sex indicated no difference in time spent running between HR and C lines for females ($p = 0.2575$), but a significant difference for males ($p = 0.0146$). For average speed, separate analyses indicated highly significant HR versus C differences for both females ($p = 0.0004$) and males ($p = 0.0002$).

Males were significantly larger than females in both HR and C lines, and HR mice were smaller than C for both sexes (figure 1), with the size reduction greater for males (sex \times linetype interaction $p = 0.0355$). Analyses of wheel running were, therefore, repeated with body mass as an additional covariate. Body mass was not a significant predictor of any wheel trait, and the p -values did not change dramatically (see electronic supplementary material, table S1).

(b) Differences in among-line variability between HR and C lines

Separate ANCOVAs (table 1) showed that revolutions per day varied among the four C lines ($p = 0.0022$) but not among the four HR lines ($p = 0.3707$). By contrast (see figure 3), average speed varied significantly among the HR lines ($p = 0.0013$), but not among the C lines ($p = 0.6940$). Minutes per day spent running varied significantly among lines for both HR and C mice (table 1, figure 3). Importantly, time spent running differed between the sexes for C lines ($p < 0.0001$), but not for HR lines ($p = 0.9501$). As an additional covariate, body

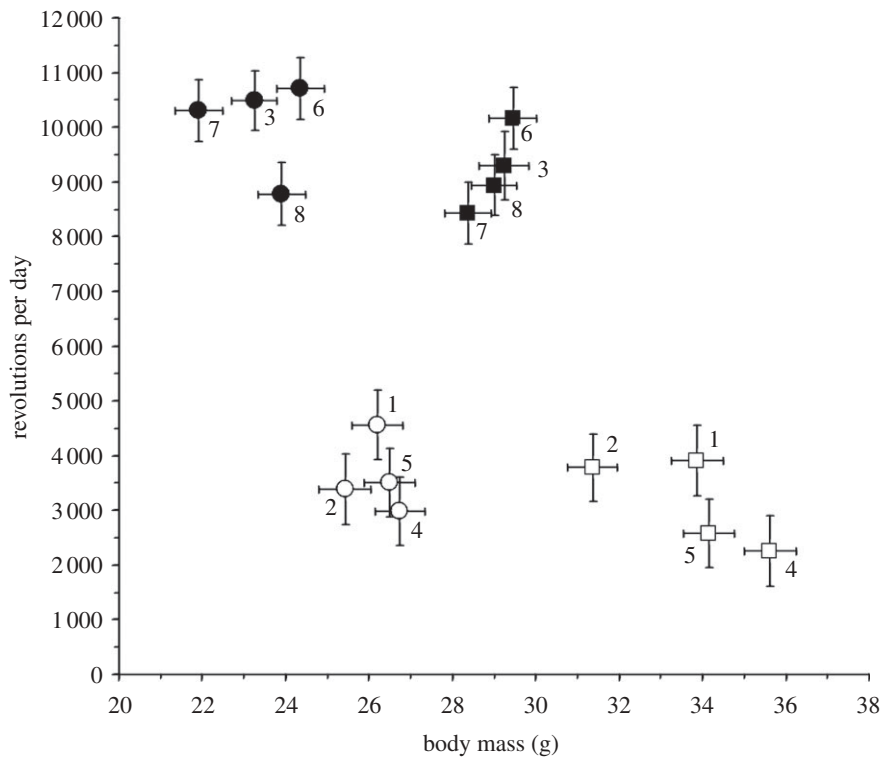


Figure 2. Sex-specific least-squares means and standard errors (from SAS procedure mixed analysis of all eight lines, excluding linetype as a factor) for each of the eight lines for mean daily wheel running and body mass prior to the start of the 6-day wheel test. This graph illustrates the 2.85-fold higher daily running by HR mice (lines numbered 3, 6, 7 and 8), as well as the smaller body sizes of HR mice (filled circle, high runner female; filled square, high runner male; open circle, control female; open square, control male).

Table 1. Separate ANCOVAs comparing the four control and four high runner lines of mice. (Models without and with body mass as a covariate are shown. Additional covariates were age and wheel freeness (results not shown).)

trait	effect	control			high runner		
		d.f.	<i>F</i>	<i>p</i>	d.f.	<i>F</i>	<i>p</i>
revolutions per day	sex	1,36	5.84	0.0209	1,37	5.95	0.0197
	line	3,35	5.92	0.0022	3,37	1.08	0.3707
	sex × line	3,36	2.00	0.1316	3,37	1.60	0.2057
revolutions per day	sex	1,36	5.11	0.0300	1,37	2.02	0.1639
	line	3,35	6.67	0.0011	3,37	1.04	0.3854
	sex × line	3,36	2.39	0.0845	3,37	1.58	0.2108
	body mass	1,74	1.67	0.2008	1,143	0.01	0.9076
minutes per day	sex	1,36	42.99	<.0001	1,37	0	0.9501
	line	3,35	20.65	<.0001	3,37	6.28	0.0015
	sex × line	3,36	2.11	0.1164	3,37	1.77	0.1705
minutes per day	sex	1,36	13.69	0.0007	1,37	3.14	0.0844
	line	3,35	20.85	<.0001	3,37	6.10	0.0018
	sex × line	3,36	2.23	0.1018	3,37	2.05	0.1235
	body mass	1,74	0.54	0.4647	1,143	5.49	0.0205
average speed (revolutions per min)	sex	1,36	3.64	0.0645	1,37	12.89	0.0010
	line	3,35	0.49	0.6940	3,37	6.46	0.0013
	sex × line	3,36	1.70	0.1843	3,37	1.25	0.3044
average speed (revolutions per min)	sex	1,36	0.03	0.8717	1,37	0.83	0.3694
	line	3,35	0.23	0.8768	3,37	6.65	0.0010
	sex × line	3,36	2.01	0.1302	3,37	1.14	0.3475
	body mass	1,74	1.47	0.2292	1,143	1.50	0.2226

mass was a significant predictor of wheel-running traits in only one of six instances, and the *p*-values for line and for the sex × line interactions were virtually unchanged in either the control lines or the HR lines. On the other

hand, some of the *p*-values for the main effect of sex were altered dramatically. For example, the higher average running speeds of females were statistically eliminated by inclusion of body mass in the model (see table 1).

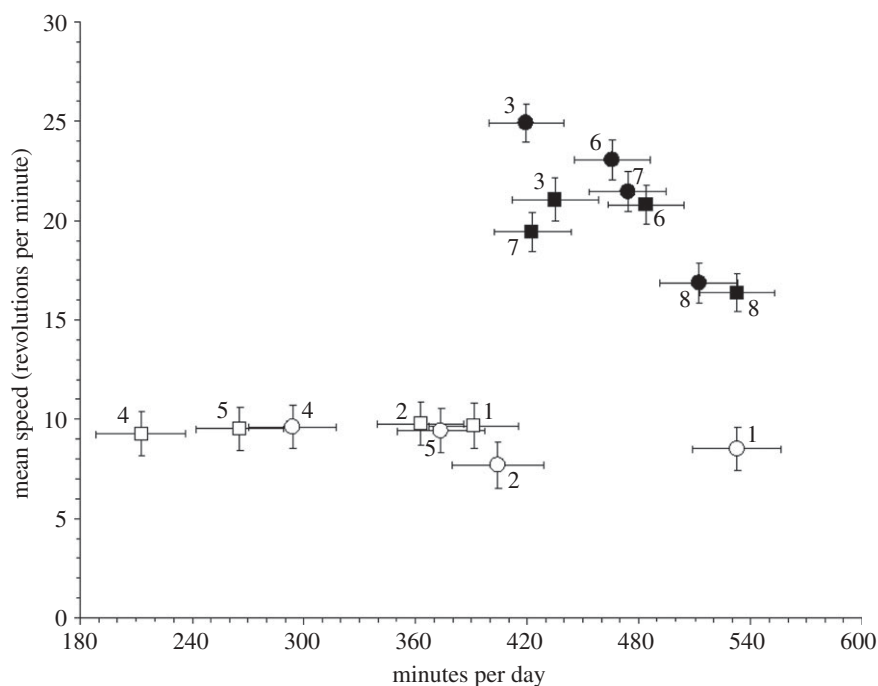


Figure 3. Sex-specific least-squares means and standard errors (from SAS procedure mixed analysis of all eight lines, excluding linetype as a factor) for each of the eight lines for the components of mean daily wheel running, number of 1 min intervals with at least one revolution and mean speed (revolutions per day divided by number of active intervals). This graph illustrates that the among-line variation in speed (*y*-axis) is significantly greater for high runner lines (numbered 3, 6, 7 and 8) than for control lines (numbered 1, 2, 4 and 5), whereas the among-line variation in time (*x*-axis) is greater for control lines (see electronic supplementary material, tables S1 and S2). The HR lines show a significant negative relation ('trade-off'), but the C lines do not (see Results; filled circle, high runner female; filled square, high runner male; open circle, control female; open square, control male).

Models with separate estimates of the among-line variance components for the HR and C lines fit better than models with a pooled among-line variance for average running speed (see electronic supplementary material, table S2: $p = 0.0583$ and 0.0169 for females and males, respectively). The among-line variance component for average speed was much larger for HR than for C lines.

(c) Correlations at the level of line means

The partial correlation between speed and duration of running (figure 3), controlling for sex, was non-significant for the four control lines (partial $r = -0.263$, two-tailed $p = 0.569$), but significant for the four HR lines (partial $r = -0.788$, $p = 0.035$). Using all 16 sex- and line-specific means, consideration of adjusted r^2 and significance levels for independent variables indicated that the best-fitting ANCOVA model predicted speed with duration ($p = 0.590$), linetype ($p = 0.001$), and the linetype \times duration interaction ($p = 0.011$). Thus, the relation between speed and duration of running (figure 3) differs significantly for the HR and C lines.

(d) Correlations at the level of individual (residual) variation

The correlation between speed and duration of wheel-running behaviour are shown separately by sex for each of the eight lines in the electronic supplementary material, figure S1. Considering the 16 separate correlation coefficients, for males, the average Pearson correlation was significantly lower in the HR lines ($r = 0.168$) than in the control lines ($r = 0.574$; ANOVA, two-tailed

$p = 0.0267$; Mann–Whitney, $p = 0.0433$). For females, the trend was the same (HR, $r = -0.034$; control, $r = 0.201$), but not statistically significant (ANOVA, two-tailed $p = 0.3907$; Mann–Whitney, $p = 0.5637$).

4. DISCUSSION

Previous research has provided numerous examples in which conspecific populations have responded to uniform selection in different ways at the genetic level [11], and these empirical results are consistent with some theoretical genetic models [41]. However, most such examples involve populations that were probably somewhat different in genetic composition prior to imposition of the uniform selection, and whether different responses are a common occurrence when (finite) replicate starting populations are virtually identical for their initial gene pools is not well documented, especially in multicellular organisms. Moreover, few studies of multicellular organisms have examined 'multiple solutions' with respect to lower-level phenotypes that compose a higher-level trait subject to selection [14,42–44], although here, too, theoretical models have suggested that uniform selection does not necessarily produce convergence at the suborganismal level [45]. Finally, possible sex-specific responses to selection have been largely ignored (but see [42,43]).

We found that uniform selection for high wheel running led to a consistent response in each of four replicate HR lines of mice. After 43 generations of selective breeding, we found an average increase of 2.85-fold in revolutions run per day (figure 1) and no statistical difference among the four HR lines, while

four non-selected control lines showed highly significant differences in revolutions per day (table 1). Thus, the trait under intentional selection was ‘clamped’ at equivalent levels in selected lines, whereas the control populations diverged via random genetic processes.

In contrast to revolutions per day, mean speed showed greater variation among the HR lines than among the C lines (table 1, figure 3, see electronic supplementary material, table S2). Thus, the selection regimen promoted phenotypic divergence at a level immediately below that of the trait under direct selection. Moreover, we find evidence of a trade-off between speed and duration of wheel running among the HR lines (figure 3), supported by the observation that at the level of individual variation the speed-duration correlation is, on average, lower (less positive) in the HR lines as compared with the C lines (see electronic supplementary material, figure S1). These results are consistent with the idea that trade-offs may only occur in organisms that are near some sort of limit, whether genetic or functional in nature (e.g. pp. 268–270 in reference [46]).

Also of note is the lack of among-line variation in average running speed for the C lines (table 1, figure 3). One possible explanation would be a low heritability of speed in the original base population. However, Swallow *et al.* [26] reported the heritabilities (\pm s.e.) based on midparent–offspring regressions of 0.18 (\pm 0.064) for revolutions per day, 0.14 (\pm 0.088; not statistically different from zero) for minutes per day, and 0.28 (\pm 0.074) for average speed. The relatively high heritability of speed suggests an alternative hypothesis that stabilizing selection may be acting on running speed through unidentified fitness components.

Identification of the genetic loci and regulatory networks [47] responsible for divergence in wheel running between the HR and C lines, or among the replicate lines within linetype, has begun [29,48], and previous studies have provided some evidence that the level of genetic divergence is also greater for the HR lines than for the C lines. A Mendelian recessive allele present at a frequency of approximately 7 per cent in the base population causes hindlimb muscle mass to be reduced by approximately 50 per cent, in addition to many other pleiotropic effects [28,49–51]. The so-called ‘mini-muscle’ phenotype was only ever observed in three of the lines; apparently, the underlying allele was lost by drift early in the experiment in the other five lines. One C line showed the mini-muscle phenotype at low frequency (less than or equal to 10%) for 22+ generations, then subsequently lost it. Two of the HR lines exhibited a dramatic increase in mini-muscle frequency, and one eventually went to fixation by generation 36 [49,52]. Thus, the four HR lines are heterogeneous with respect to mini-muscle frequency, whereas the four C lines are uniform in lacking it.

In the present study, male (2.91) and female (2.79) HR mice showed a similar fold increase in revolutions per day relative to the C lines (figure 1), as has been documented previously (e.g. [15,26,34,53]). However, as also documented previously ([54] and references therein), the increased daily running distance is accomplished mainly by speed in female HR mice, but by both speed and duration of running in HR males (figure 1). The increased time running in HR males leads to relatively higher

energy costs of running, as compared with HR females, and may result in sex-specific limitations to the evolution of further increases in daily running distances [54]. Male HR mice also run with shorter, more frequent bouts, as compared with HR females ([54]; see also [55]). Another interesting sex-related aspect of the response to selection is the observation that male and female HR mice have converged on the same amount of time spent running per day (figure 3, table 1), whereas in C, and in the starting population [26], females run more minutes per day. This result suggests that the limit to further increases in running may in part be related to time constraints, possibly having to do with the need for sleep [54].

The morphological, physiological and, ultimately, genetic bases of the sex differences in response to selection are starting to be elucidated [56,57]. For example, the increase in maximal oxygen consumption (measured during forced treadmill exercise) exhibited for HR lines is greater for males than for females [54,58]. Moreover, the neurochemical underpinnings of elevated wheel running have been shown to differ for HR males and females, specifically with respect to the endocannabinoid system [53]. At the genetic level, sex-specific effects of quantitative trait loci (QTL) have been identified [48].

The evolution of increased voluntary wheel running is analogous to the evolution of other composite traits, such as certain measures of organismal performance [46] or even brain size, as discussed by Atchley *et al.* [59]. Increased wheel-running distance is the product of increases in mean speed and minutes of running. Similarly, increased adult brain size is the product of hyperplasia (increased cell numbers) and hypertrophy (increased cell size). In both cases, different combinations of the two component phenotypes can yield the same increase in the higher-level phenotype. As another example, the number of ova shed and the rate of loss of ova and conceptuses both differ in two replicate lines mice that were subject to index selection for high fertility [44].

As the number of component phenotypes within a higher-level, complex trait increases, the number of different, functionally equivalent solutions is predicted to increase and, as a consequence, the correlation between variation in any one (or even a few) lower-level traits and variation in the putative selective regime should become weaker [45]. This perspective emphasizes the importance of taking a holistic, integrative, and multilevel approach to the study of adaptation. Moreover, one needs to consider the possibility of sex-specific solutions to a given adaptive problem.

Animals were maintained in accordance with NIH guidelines, and all procedures were approved by the IACUC of UCR, which is accredited by AAALAC.

T.G. thanks Jon Seger for encouragement. We thank D. A. Roff for helpful discussions about analyses. We thank M. Blows and two anonymous referees for extensive and insightful comments on previous versions of this manuscript. This work was supported by US NSF grant IOB-0543429 to T.G.

REFERENCES

- 1 Bock, W. J. 1959 Preadaptation and multiple evolutionary pathways. *Evolution* **13**, 194–211. (doi:10.2307/2405873)

- 2 Mayr, E. 1961 Cause and effect in biology. *Science* **134**, 1501–1506. (doi:10.1126/science.134.3489.1501)
- 3 Irschick, D. J., Herrel, A. & Vanhooydonck, B. 2006 Whole-organism studies of adhesion in pad-bearing lizards: creative evolutionary solutions to functional problems. *J. Comp. Physiol. A* **192**, 1169–1177. (doi:10.1007/s00359-006-0145-2)
- 4 Harvey, P. H. & Pagel, M. D. 1991 *The comparative method in evolutionary biology*. Oxford, UK: Oxford University Press.
- 5 Middleton, K. M. & Gatesy, S. M. 2000 Theropod forelimb design and evolution. *Zool. J. Linn. Soc.* **128**, 149–187. (doi:10.1111/j.1096-3642.2000.tb00160.x)
- 6 Langerhans, R. B. & DeWitt, T. J. 2004 Shared and unique features of evolutionary diversification. *Am. Nat.* **164**, 335–349. (doi:10.1086/422857)
- 7 Futuyma, D. J. 1986 *Evolutionary biology*, 2nd edn. Sunderland, MA: Sinauer Associates.
- 8 Maddison, D. R. & Maddison, W. P. 2000 *MacClade 4: analysis of phylogeny and character evolution*. Sunderland, MA: Sinauer Associates.
- 9 Garland Jr, T., Bennett, A. F. & Rezende, E. L. 2005 Phylogenetic approaches in comparative physiology. *J. Exp. Biol.* **208**, 3015–3035. (doi:10.1242/jeb.01745)
- 10 Revell, L. J., Johnson, M. A., Schulte II, J. A., Kolbe, J. J. & Losos, J. B. 2007 A phylogenetic test for adaptive convergence in rock-dwelling lizards. *Evolution* **61**, 2898–2912. (doi:10.1111/j.1558-5646.2007.00225.x)
- 11 Cohan, F. M. 1984 Can uniform selection retard random genetic divergence between isolated conspecific populations? *Evolution* **38**, 495–504. (doi:10.2307/2408699)
- 12 Beall, C. M. 2007 Two routes to functional adaptation: Tibetan and Andean high-altitude natives. *Proc. Natl Acad. Sci. USA* **104**, 8655–8660. (doi:10.1073/pnas.0701985104)
- 13 Simões, P., Santos, J., Fragata, I., Mueller, L. D., Rose, M. R. & Matos, M. 2008 How repeatable is adaptive evolution? The role of geographical origin and founder effects in laboratory adaptation. *Evolution* **62**, 1817–1829. (doi:10.1111/j.1558-5646.2008.00423.x)
- 14 Endler, J. A., Basolo, A., Glowacki, S. & Zerr, J. 2001 Variation in response to artificial selection for light sensitivity in guppies (*Poecilia reticulata*). *Am. Nat.* **158**, 36–48. (doi:10.1086/320862)
- 15 Garland Jr, T. 2003 Selection experiments: an under-utilized tool in biomechanics and organismal biology. In *Vertebrate biomechanics and evolution* (eds V. L. Bels, J.-P. Gasc & A. Casinos), pp. 23–56. Oxford, UK: BIOS Scientific Publishers.
- 16 Bradley, T. J. & Folk, D. G. 2004 Analyses of physiological evolutionary response. *Physiol. Biochem. Zool.* **77**, 1–9. (doi:10.1086/381466)
- 17 Swallow, J. G. & Garland Jr, T. 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. (doi:10.1093/icb/45.3.387)
- 18 Garland Jr, T. & Rose, M. R. 2009 *Experimental evolution: concepts, methods, and applications of selection experiments*. Berkeley, CA: University of California Press.
- 19 Bell, G. 2008 *Selection: the mechanism of evolution*, 2nd edn. Oxford, UK: Oxford University Press.
- 20 Weber, K. E., Greenspan, R. J., Chicoine, D. R., Fiorentino, K., Thomas, M. H. & Knight, T. L. 2008 Microarray analysis of replicate populations selected against a wing-shape correlation in *Drosophila melanogaster*. *Genetics* **178**, 1093–1108. (doi:10.1534/genetics.107.078014)
- 21 Forde, S. E. & Jessup, C. M. 2009 Understanding evolution through the phages. In *Experimental evolution: concepts, methods, and applications of selection experiments* (eds T. Garland, Jr & M. R. Rose), pp. 391–418. Berkeley, CA: University of California Press.
- 22 Rosenzweig, F. & Sherlock, G. 2009 Through a glass clearly: experimental evolution as a window on adaptive genome evolution. In *Experimental evolution: concepts, methods, and applications of selection experiments* (eds T. Garland, Jr & M. R. Rose), pp. 353–388. Berkeley, CA: University of California Press.
- 23 Travisano, M. 2009 Long-term experimental evolution and adaptive radiation. In *Experimental evolution: concepts, methods, and applications of selection experiments* (eds T. Garland, Jr & M. R. Rose), pp. 111–133. Berkeley, CA: University of California Press.
- 24 Fong, S. S., Joyce, A. R. & Palsson, B. O. 2005 Parallel adaptive evolution cultures of *Escherichia coli* lead to convergent growth phenotypes with different expression states. *Genome Res.* **15**, 1365–1372. (doi:10.1101/gr.3832305)
- 25 Ostrowski, E., Woods, R. & Lenski, R. 2008 The genetic basis of parallel and divergent phenotypic responses in evolving populations of *Escherichia coli*. *Proc. R. Soc. B* **275**, 277–284. (doi:10.1098/rspb.2007.1244)
- 26 Swallow, J. G., Carter, P. A. & Garland Jr, T. 1998 Artificial selection for increased wheel-running behavior in house mice. *Behav. Genet.* **28**, 227–237. (doi:10.1023/A:1021479331779)
- 27 Rhodes, J. S., Gammie, S. C. & Garland Jr, T. 2005 Neurobiology of mice selected for high voluntary wheel-running activity. *Integr. Comp. Biol.* **45**, 438–455. (doi:10.1093/icb/45.3.438)
- 28 Middleton, K. M., Kelly, S. A. & Garland Jr, T. 2008 Selective breeding as a tool to probe skeletal response to high voluntary locomotor activity in mice. *Integr. Comp. Biol.* **48**, 394–410. (doi:10.1093/icb/icn057)
- 29 Bronikowski, A. M., Rhodes, J. S., Garland Jr, T., Prolla, T. A., Awad, T. & Gammie, S. C. 2004 The evolution of gene expression in mouse hippocampus in response to selective breeding for increased locomotor activity. *Evolution* **58**, 2079–2086. (doi:10.1111/j.0014-3820.2004.tb00491.x)
- 30 Belke, T. W. & Garland Jr, T. 2007 A brief opportunity to run does not function as a reinforcer for mice selected for high daily wheel-running rates. *J. Exp. Anal. Behav.* **88**, 199–213. (doi:10.1901/jeab.2007.62-06)
- 31 Malisch, J. L., Saltzman, W., Gomes, F. R., Rezende, E. L., Jeske, D. R. & Garland Jr, T. 2007 Baseline and stress-induced plasma corticosterone concentrations of mice selectively bred for high voluntary wheel running. *Physiol. Biochem. Zool.* **80**, 146–156. (doi:10.1086/508828)
- 32 Gomes, F. R., Rezende, E. L., Malisch, J. L., Lee, S. K., Rivas, D. A., Kelly, S. A., Lytle, C., Yaspelkis III, B. B. & Garland Jr, T. 2009 Glycogen storage and muscle glucose transporters (GLUT-4) of mice selectively bred for high voluntary wheel running. *J. Exp. Biol.* **212**, 238–248. (doi:10.1242/jeb.025296)
- 33 Nehrenberg, D. L., Hua, K., Estrada-Smith, D., Garland Jr, T. & Pomp, D. 2009 Voluntary exercise and its effects on body composition depend on genetic selection history. *Obesity* **17**, 1402–1409. (doi:10.1038/oby.2009.51)
- 34 Swallow, J. G., Koteja, P., Carter, P. A. & Garland Jr, T. 1999 Artificial selection for increased wheel-running activity in house mice results in decreased body mass at maturity. *J. Exp. Biol.* **202**, 2513–2520.
- 35 Hauschka, T. S. & Mirand, E. A. 1973 The ‘breeder: HA(ICR)’ Swiss mouse, a multi-purpose stock selected for fecundity. In *Perspectives in cancer research and treatment* (eds G. P. Murphy, D. Pressman & E. A. Mirand), pp. 319–331. New York, NY: Alan R. Riss, Inc.
- 36 Girard, I., Swallow, J. G., Carter, P. A., Koteja, P., Rhodes, J. S. & Garland Jr, T. 2002 Maternal-care

- behavior and life-history traits in house mice (*Mus domesticus*) artificially selected for high voluntary wheel-running activity. *Behav. Processes* **57**, 37–50. (doi:10.1016/S0376-6357(01)00206-6)
- 37 Malisch, J. L., Breuner, C. W., Kolb, E. M., Wada, H., Hannon, R. M., Chappell, M. A., Middleton, K. M. & Garland Jr, T. 2009 Behavioral despair and home-cage activity in mice with chronically elevated baseline corticosterone concentrations. *Behav. Genet.* **39**, 192–201. (doi:10.1007/s10519-008-9246-8)
- 38 Eisenmann, J. C., Wickel, E. E., Kelly, S. A., Middleton, K. M. & Garland Jr, T. 2009 Day-to-day variability in voluntary wheel running among young mice selectively bred for high activity. *Eur. J. Appl. Physiol.* **106**, 613–619. (doi:10.1007/s00421-009-1056-z)
- 39 Krugner-Higby, L., Girard, I., Welter, J., Gendron, A., Rhodes, J. S. & Garland Jr, T. 2006 Clostridial enteropathy in lactating outbred swiss-derived (ICR) mice. *J. Am. Assoc. Lab. Anim. Sci.* **45**, 80–87.
- 40 Cohan, F. M. & Hoffmann, A. A. 1989 Uniform selection as a diversifying force in evolution: evidence from *Drosophila*. *Am. Nat.* **134**, 613–637. (doi:10.1086/285000)
- 41 Cohan, F. M. 1984 Genetic divergence under uniform selection. I. Similarity among populations of *Drosophila melanogaster* in their responses to artificial selection for modifiers of *ci^D*. *Evolution* **38**, 55–71. (doi:10.2307/2408547)
- 42 Huey, R. B., Gilchrist, G. W., Carlson, M. L., Berrigan, D. & Serra, L. 2000 Rapid evolution of a geographic cline in size in an introduced fly. *Science* **287**, 308–309. (doi:10.1126/science.287.5451.308)
- 43 Calboli, F. C. F., Gilchrist, G. W. & Partridge, L. 2003 Different cell size and cell number contribution in two newly established and one ancient body size cline of *Drosophila subobscura*. *Evolution* **57**, 566–573. (doi:10.1111/j.0014-3820.2003.tb01548.x)
- 44 Spitschak, M., Langhammer, M., Schneider, F., Renne, U. & Vanselow, J. 2007 Two high-fertility mouse lines show differences in component fertility traits after long-term selection. *Reprod. Fertil. Dev.* **19**, 815–821. (doi:10.1071/RD07009)
- 45 Alfaro, M., Bolnick, D. I. & Wainwright, P. C. 2004 Evolutionary dynamics of complex biomechanical systems. *Evolution* **58**, 495–503. (doi:10.1554/03-404)
- 46 Garland Jr, T. 1994 Quantitative genetics of locomotor behavior and physiology in a garter snake. In *Quantitative genetic studies of behavioral evolution* (ed. C. R. B. Boake), pp. 251–277. Chicago, IL: University of Chicago Press.
- 47 Chouard, T. 2008 Beneath the surface. *Nature* **456**, 300–303. (doi:10.1038/456300a)
- 48 Nehrenberg, D. L., Wang, S., Hannon, R. M., Garland Jr, T. & Pomp, D. 2010 QTL underlying voluntary exercise in mice: interactions with the ‘mini muscle’ locus and sex. *J. Hered.* **101**, 42–53. (doi:10.1093/jhered/esp066)
- 49 Garland Jr, T., Morgan, M. T., Swallow, J. G., Rhodes, J. S., Girard, I., Belter, J. G. & Carter, P. A. 2002 Evolution of a small-muscle polymorphism in lines of house mice selected for high activity levels. *Evolution* **56**, 1267–1275. (doi:10.1111/j.0014-3820.2002.tb01437.x)
- 50 Hannon, R. M., Kelly, S. A., Middleton, K. M., Kolb, E. M., Pomp, D. & Garland Jr, T. 2008 Phenotypic effects of the ‘mini-muscle’ allele in a large HR × C57BL/6J mouse backcross. *J. Hered.* **99**, 349–354. (doi:10.1093/jhered/esn011)
- 51 Bilodeau, G. M., Guderley, H., Joanisse, D. R. & Garland Jr, T. 2009 Reduction of type IIb myosin and IIB fibers in tibialis anterior muscle of mini-muscle mice from high-activity lines. *J. Exp. Zool. A: Ecol. Genet. Physiol.* **311A**, 189–198. (doi:10.1002/jez.518)
- 52 Syme, D. A., Evashuk, K., Grintuch, B., Rezende, E. L. & Garland Jr, T. 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. (doi:10.1152/jappphysiol.00369.2005)
- 53 Keeney, B. K., Raichlen, D. A., Meek, T. H., Wijeratne, R. S., Middleton, K. M., Gerdeman, G. L. & Garland Jr, T. 2008 Differential response to a selective cannabinoid receptor antagonist (SR141716: rimonabant) in female mice from lines selectively bred for high voluntary wheel-running behavior. *Behav. Pharmacol.* **19**, 812–820. (doi:10.1097/FBP.0b013e32831c3b6b)
- 54 Rezende, E. L., Gomes, F. R., Chappell, M. A. & Garland Jr, T. 2009 Running behavior and its energy cost in mice selectively bred for high voluntary locomotor activity. *Physiol. Biochem. Zool.* **82**, 662–679. (doi:10.1086/605917)
- 55 Girard, I., McAleer, M. W., Rhodes, J. S. & Garland Jr, T. 2001 Selection for high voluntary wheel running increases intermittency in house mice (*Mus domesticus*). *J. Exp. Biol.* **204**, 4311–4320.
- 56 Houle-Leroy, P., Garland Jr, T., Swallow, J. G. & Guderley, H. 2000 Effects of voluntary activity and genetic selection on muscle metabolic capacities in house mice *Mus domesticus*. *J. Appl. Physiol.* **89**, 1608–1616.
- 57 Swallow, J. G., Rhodes, J. S. & Garland Jr, T. 2005 Phenotypic and evolutionary plasticity of organ masses in response to voluntary exercise in house mice. *Integr. Comp. Biol.* **45**, 426–437. (doi:10.1093/icb/45.3.426)
- 58 Rezende, E. L., Kelly, S. A., Gomes, F. R., Chappell, M. A. & Garland Jr, T. 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**, 83–99. (doi:10.1086/498187)
- 59 Atchley, W. R., Xu, S. & Vogl, X. 1994 Developmental quantitative genetic models of evolutionary change. *Dev. Genet.* **15**, 92–103. (doi:10.1002/dvg.1020150110)

Electronic Supplemental Table 1. Mixed-model ANCOVAs comparing trait mean values for the four replicate Control (C) and four replicate High Runner (HR) lines of mice, without and with body mass as a covariate. (d.f. for Sex, Linetype and Sex * Linetype interaction are 1 and 6. Additional covariates were age for all traits and wheel freeness for wheel-running traits [results not shown]. Sign after *p*-value for Body Mass indicates direction of effect. Associated least-squares means and standard errors are presented in Figure 1 of the publication.)

Trait	Effect	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>
Body Mass (g)	Sex	355.30	<.0001		
	Linetype	30.08	0.0015		
	Sex*Linetype	7.29	0.0355		
Revolutions/Day	Sex	5.86	0.0518	2.91	0.1392
	Linetype	161.56	<.0001	136.08	<.0001
	Sex*Linetype	0.59	0.4732	0.48	0.5155
	Body Mass			0.10	0.7539 +
Minutes/Day	Sex	11.19	0.0155	11.37	0.0150
	Linetype	5.35	0.0600	6.37	0.0451
	Sex*Linetype	11.68	0.0142	11.71	0.0141
	Body Mass			2.81	0.0945 +
Average Speed (revs/min)	Sex	2.38	0.1736	0.55	0.4844
	Linetype	61.38	0.0002	57.03	0.0003
	Sex*Linetype	10.00	0.0195	9.98	0.0196
	Body Mass			0.06	0.8120 -

Electronic Supplemental Table 2. Likelihood ratio tests comparing models that estimate separate among-line variances for Control and High Runner lines of mice (**in bold**) with those that estimate a pooled among-line variance component. (The fit of models is compared with ln maximum likelihood ratio tests that employ a χ^2 statistic [twice the difference in ln likelihoods] with 1 d.f. A significant *p* value [2-tailed] indicates that the HR and Control lines differ in the amount of among-line variance.)

Trait	ln maximum likelihood for simpler model	ln maximum likelihood for model with separate variances	χ^2	<i>p</i>	estimated among-line variance for Control	estimated among-line variance for High Runner	estimated among-family variance for Control	estimated among-family variance for High Runner	estimated among-individual variance for Control	estimated among-individual variance for High Runner
Revolutions/Day										
females	-1748.44	-1748.40	0.07	0.7845	328,905	119,789	364,569	3,191,768	1,445,882	7,666,532
males	Did	not	con-	verge						
Minutes/Day										
females	-1154.25	-1153.11	2.27	0.1319	9,259	1,062	2,905	1,305	6,585	8,288
males	-1136.71	-1136.33	0.77	0.3814	6,398	1,949	0	2,559	9,937	7,014
Average Speed (revs/min)										
females	-558.84	-557.05	3.58	0.0583	0.36	9.93	1.02	11.33	6.05	24.72
males	-515.72	-512.87	5.71	0.0169	0.00	5.38	1.68	3.84	5.75	14.82

Electronic Supplemental Figure 1. Individual variation in average speed (residual revolutions • min⁻¹) versus duration of wheel running (residual 1-minute intervals with >0 revolutions). Lines 1, 2, 4, and 5 are non-selected Control, whereas 3, 6, 7, and 8 are High Runner. For males, the average Pearson correlation was significantly lower in the HR lines ($r = 0.168$) than in the Control lines ($r = 0.574$) (ANOVA, 2-tailed $p = 0.0267$; Mann-Whitney, $p = 0.0433$). For females, the trend was the same (HR $r = -0.034$, Control $r = 0.201$), but not statistically significant (ANOVA, 2-tailed $p = 0.3907$; Mann-Whitney, $p = 0.5637$).

