Identification of exercise capacity QTL using association mapping in inbred mice

Sean M. Courtney and Michael P. Massett

Department of Health and Kinesiology, Texas A&M University, College Station, Texas

Submitted 20 April 2012; accepted in final form 20 August 2012

Courtney SM, Massett MP. Identification of exercise capacity QTL using association mapping in inbred mice. Physiol Genomics 44: 948–955, 2012. First published August 21, 2012; doi:10.1152/physiolgenomics.00051.2012.—There are large inter-individual differences in exercise capacity. It is well established that there is a genetic basis for these differences. However, the genetic factors underlying this variation are undefined. Therefore, the purpose of this study was to identify novel putative quantitative trait loci (QTL) for exercise capacity by measuring exercise capacity in inbred mice and performing genome-wide association mapping. Exercise capacity, defined as run time and work, was assessed in male mice (n = 6) from 34 strains of classical and wild-derived inbred mice performing a graded treadmill test. Genome-wide association mapping was performed with an efficient mixed-model association (EMMA) algorithm to identify QTL. Exercise capacity was significantly different across strains. Run time varied by 2.7-fold between the highest running strain (CSB/J) and the lowest running strain (A/J). These same strains showed a 16.5-fold difference in work. Significant associations were identified for exercise time on chromosomes 1, 2, 7, 11, and 13. The QTL interval on chromosome 2 (~168 Mb) contains one gene, Nfatc2, and overlaps with a suggestive QTL for training responsiveness in humans. These results provide phenotype data on the widest range of inbred strains tested thus far and indicate that genetic background significantly influences exercise capacity. Furthermore, the novel QTLs identified in the current study provide new targets for investigating the underlying mechanisms for variation in exercise capacity.

Intrinsic endurance exercise capacity, as assessed by exercise time during a graded treadmill test, is commonly used to assess cardiorespiratory fitness and is highly correlated to risk of cardiovascular disease (26). In sedentary humans, there are substantial individual differences in exercise capacity, and it is well established that there is a genetic basis for these individual differences (4, 6, 31). Because of the health benefits associated with high levels of fitness, the physiological factors determining exercise capacity have been widely studied. However, little is known about the underlying genetic determinants of exercise capacity (6, 13). To elucidate these genetic factors, Bouchard and colleagues used genome-wide linkage analysis to identify quantitative trait loci (QTL) for exercise capacity, quantified by maximal oxygen consumption (VO2max), in the sedentary state and in response to training. Several promising and suggestive QTL for pretraining exercise capacity were reported (4, 31), and these regions differed from those for the VO2max response to training. These findings imply that the genes that determine sedentary-state VO2 are different from those that determine the response to training (4, 6, 31). Thus, the underlying genes or DNA variants determining the genetic effect on sedentary-state or intrinsic exercise capacity will need to be determined to understand the link between exercise capacity and susceptibility to disease.

Several rodent models have been utilized to investigate the genetic factors contributing to intrinsic endurance exercise capacity assessed by treadmill running (2, 22, 24, 25, 42). Using traditional genome-wide linkage analysis, investigators have identified several QTL for intrinsic endurance exercise capacity in rats (2, 42) and mice (22, 25). However, traditional QTL analyses are limited by the variation present in the genomes of the two mouse strains (3, 16), which can reduce mapping resolution. A relatively new approach for refining large QTL regions and identifying novel QTL for disease and behavior-related traits, including physical activity is genome-wide association mapping (3, 9, 20). Early attempts at genome-wide association mapping in mice were criticized for low statistical power and differences in population structure across inbred strains leading to a high rate of false positive associations due limited genomic information from a small number of mouse strains (12, 41). Recent advances in single nucleotide polymorphism (SNP) discovery and genomic sequencing capabilities have led to the creation of dense SNP maps available for a large number of inbred strains (10, 44). These large SNP databases have facilitated the development of genome-wide association mapping approaches in mice that account for some of the concerns raised with earlier studies (12, 16, 28). In the current study, we employed efficient mixed model association (EMMA) mapping. EMMA relies on a kinship matrix to account for genetic relatedness and population structure in inbred strains of mice and other model organisms, which can reduce the number of false positive associations (16, 17). EMMA has been used successfully to identify QTL for physiological traits such as body weight, bone mineral density, HDL cholesterol, airway responsiveness, and pulmonary adenomas in inbred mice (3, 9, 16). Thus, given the limited information regarding the genetic determinants of intrinsic endurance exercise capacity and the improved methodology for in silico genetic analyses, the purpose of this study was to characterize intrinsic endurance exercise capacity in 34 strains of classical and wild-derived inbred mice and to apply association mapping to identify novel putative QTL for endurance exercise capacity.

METHODS

All procedures in this study were approved by the Institutional Animal Care and Use Committee at Texas A&M University and adhered to the American Physiological Society’s Guiding Principles in the Care and Use of Vertebrate Animals in Research and Training.
**Association Mapping of Exercise Capacity**

**Animals.** Seven-week-old male mice from 34 inbred strains (129S1/SvImJ, 129X1/SvJ, A/J, AKR/J, BALB/cByJ, BUB/BalJ, C3H/HeJ, C57BL/6J, C57BLKS/J, C57L/J, C58/J, CAST/Eij, CBA/J, CE/J, DBA/2J, FVB/NJ, KK/HJ, I/LnJ, LP/J, MA/MyJ, MRL/Mpj, NOD/ShiLtJ, NON/ShiLtJ, NZB/BINJ, NZ/O/HI/J, NZW/LacJ, PL/J, PWDF/PJ, PWK/PFJ, RIJS/J, SJL/J, SM/J, SWR/J, and WSB/Eij) (n = 6 per strain) were purchased from Jackson Laboratory (Bar Harbor, ME). These strains were chosen for genetic diversity incorporating inbred and wild-derived strains (27, 44). Animals were housed in group cages and kept on a 12 h light-dark schedule (7:00 AM–7:00 PM) in a temperature-controlled environment (21.0–22.0°C) with food and water provided ad libitum. All mice were allowed 1 wk to become accustomed to the housing facility before completing the exercise protocol. All exercise tests were performed between 9:00 AM and 11:00 AM.

**Endurance exercise test.** Intrinsic endurance exercise capacity was defined for each of the 34 strains as time (minutes) and work (kg-m) performed during a graded exercise test. Work performed (kg-m) or vertical work was calculated as a product of body weight (kg) and vertical distance (meters), where vertical distance = (distance run) × sinθ, where θ is equal to the angle of the treadmill from 0° to 15° (24, 25). To identify strain-specific phenotypic differences in endurance exercise capacity, we had all mice perform two graded exercise tests on a six-lane treadmill (Columbus Instruments, Columbus, OH). Prior to completing the exercise testing, mice were familiarized for 2 days running for 10 min up a 10° incline at 9.0 m/min on day 1 and 11.0 m/min on day 2. This protocol introduced the mice to the treadmill running but was limited to avoid inducing any training adaptations. Following familiarization, intrinsic exercise capacity was measured in 8 wk old mice by two graded exercise tests separated by 48 h. The test started at a speed of 9.0 m/min for 9 min at 0°. After 9 min the speed was increased to 10 m/min and the grade was increased 5°. Thereafter, speed was increased 2.5 m/min every 3 min and the grade was increased 5° every 9 min to a maximum grade of 15°. The graded exercise test continued until mice maintained contact with the shock grid for more than 15 s continuously or could no longer be motivated to run (24, 25). For each mouse, the average exercise capacity for the two trials was used to calculate strain means and for association mapping. Individual body weights were measured before each exercise test.

**Association mapping.** Genome-wide association mapping for exercise phenotypes and body mass were performed with EMMA using the UCLA web-based server (http://mouse.cs.ucla.edu/emmaserver). EMMA uses a linear mixed model algorithm to account for relatedness among inbred strains, which reduces the rate of false positive associations (3, 16, 17). Analysis of individual phenotype data was performed with SNP panels consisting of 132,285 SNPs (132 K) and 4 million SNPs (4 M) (3, 17). The initial analysis was conducted using the 132 K panel to identify putative loci for endurance exercise capacity. The SNP density of this panel is comparable to previously published mouse genome-wide association mapping studies conducted on a similar number of mouse strains. A second analysis then was conducted using the 4 M SNP panel to compare with the results from the 132 K panel and to identify novel loci not identified with the smaller panel. The 4 M SNP panel includes the SNPs in the 132 K panel. We conducted association mapping on all 34 strains and after omitting phenotype data from the wild-derived inbred strains to assess the affect these strains have on QTL discovery. Genome-wide significance thresholds were calculated based on a false discovery rate of 5% (or q value < 0.05) using the R package q value (34). We determined confidence intervals by expanding the interval around the peak SNP to include all neighboring SNPs surpassing the significance threshold. For single SNP associations, the QTL confidence interval was set at 400 kb (200 kb on either side of the peak SNP), but only the location of the peak SNP is reported. If two QTL overlapped or were separated by <1 Mb they were considered one QTL. All SNP and gene locations were mapped to Build 37.2 of the National Center for Biotechnology Information’s mouse genome. SNP-associated P values were transformed with −log10 (P value) for graphing association scores.

**Statistical analysis.** All data are represented as means ± SE. Exercise capacity variables and body mass data from the 34 strains was analyzed by JMP 9.0 (SAS Institute, Cary, NC). Analysis of variance (ANOVA) followed by a Tukey’s post hoc test was used to identify significant strain-specific differences across all 34 strains. To assess the allele effect for significant peak SNPs, we grouped strains by genotype at each individual SNP and compared exercise time between groups by Mann-Whitney test. We calculated the proportion of variance explained by significant peak SNPs by fitting exercise time against the genotypes using a linear model. Statistical significance was set at P < 0.05.

**Results**

Body mass in 8-wk-old male mice varied significantly among inbred strains (Fig. 1 and Table 1). There was an approximately threefold difference in body mass across strains. The strain with the lowest body mass was the wild-derived CAST/Eij mice (13.8 ± 0.2 g), whereas the heaviest strain was the MRL/Mpj (39.1 ± 0.8 g). Because there are strain-specific differences in body mass, work performed during the graded exercise test was reported in addition to time.

**Intrinsic endurance exercise capacity.** A strain screen was performed to identify strain-specific differences in intrinsic exercise capacity across 34 strains of male mice in a graded exercise test. All mice completed both graded exercise tests. Both tests were conducted by the same individual and were...
highly reproducible (within-mouse coefficient of variation = 1.56 ± 0.14%). The strain distribution patterns for time and work are shown in Fig. 2, and significant differences are indicated in Table 1. There were significant differences across the strains for each exercise phenotype. For run time, there was a nearly threefold difference (2.73) between the lowest (A/J: 0.2 min) and highest (C58/J: 50.3 min) performing strains (Fig. 2A). The variation among strains was greater for work than for time (Fig. 2B). There is a 15-fold difference between the lowest performing (A/J: 0.29 ± 0.02 kg·m) and highest performing (C58/J: 4.78 ± 0.12 kg·m) strains. The strain distribution pattern also changed when exercise capacity, we conducted genome-wide association mapping using body mass from the 34 inbred strains and the 30 classical inbred strains (Table 4). This analysis was performed to identify any body mass QTL that might overlap with QTL for exercise capacity. A significance threshold corresponding to \( P = 10^{-5} \) was used for this analysis to maximize the number of significant associations identified. None of the significant associations for body mass overlapped with significant associations for time or work, suggesting that none of the exercise-related QTL were primarily due to genetic factors controlling body mass. Seven of the putative QTL identified for body mass overlapped with previously identified QTL for body mass or body mass change after an experimental intervention (e.g., high-fat diet) (Table 4), with the exception of the QTL on chromosome 11 located at 54.08 Mb.

Two peak SNPs (rs27288988 and rs4135796) were identified in multiple analyses. To assess the effect of each allele on exercise capacity, we grouped strains according to their genotype and compared them for exercise time. For rs27288988 on chromosome 2, 25 strains carry the G allele and 9 the T allele. Strains carrying the G allele had significantly greater exercise capacity, whereas the variance accounted for by the peak on chromosome 11 was 36.2% (rs4135796) on chromosome 2, 25 strains carry the G allele and 9 the T allele.

Table 1. Statistical differences among 34 inbred strains for time, work, and body mass

<table>
<thead>
<tr>
<th>Strain</th>
<th>Time (Fig. 2A)</th>
<th>Work (Fig. 2B)</th>
<th>Body Mass (Fig. 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C58/J</td>
<td>A</td>
<td>A</td>
<td>GHI</td>
</tr>
<tr>
<td>AKR/J</td>
<td>B</td>
<td>B</td>
<td>DEF</td>
</tr>
<tr>
<td>CAST/EjJ</td>
<td>BC</td>
<td>IKL</td>
<td>M</td>
</tr>
<tr>
<td>WSB/EjJ</td>
<td>CD</td>
<td>JK</td>
<td>M</td>
</tr>
<tr>
<td>SWR/J</td>
<td>DE</td>
<td>DE</td>
<td>GH</td>
</tr>
<tr>
<td>PWD/PhJ</td>
<td>DEF</td>
<td>JKL</td>
<td>LM</td>
</tr>
<tr>
<td>PWK/PhJ</td>
<td>EFG</td>
<td>HIJK</td>
<td>K</td>
</tr>
<tr>
<td>MA/MjJ</td>
<td>EFG</td>
<td>CDE</td>
<td>EFG</td>
</tr>
<tr>
<td>SJJ/1J</td>
<td>EFG</td>
<td>EFG</td>
<td>GHI</td>
</tr>
<tr>
<td>NON/ShiLtJ</td>
<td>EFG</td>
<td>C</td>
<td>CD</td>
</tr>
<tr>
<td>DBA/2J</td>
<td>EFG</td>
<td>EFGH</td>
<td>GHI</td>
</tr>
<tr>
<td>SIJ/1J</td>
<td>EFG</td>
<td>EFG</td>
<td>GHI</td>
</tr>
<tr>
<td>FVB/NJ</td>
<td>FG</td>
<td>DE</td>
<td>DEF</td>
</tr>
<tr>
<td>C57LIJ</td>
<td>G</td>
<td>EF</td>
<td>EFG</td>
</tr>
<tr>
<td>NOD/ShiLtJ</td>
<td>G</td>
<td>CD</td>
<td>BC</td>
</tr>
<tr>
<td>LPI/1J</td>
<td>G</td>
<td>GHJ</td>
<td>HIJ</td>
</tr>
<tr>
<td>CBA/1J</td>
<td>H</td>
<td>FGH</td>
<td>DE</td>
</tr>
<tr>
<td>BAL/bcByJ</td>
<td>HI</td>
<td>IJK</td>
<td>EFGH</td>
</tr>
<tr>
<td>CE/1J</td>
<td>HI</td>
<td>HIJK</td>
<td>DE</td>
</tr>
<tr>
<td>C57BLKS/1J</td>
<td>HI</td>
<td>JKL</td>
<td>MN</td>
</tr>
<tr>
<td>NZB/BINJ</td>
<td>H</td>
<td>KLMN</td>
<td>GH</td>
</tr>
<tr>
<td>129X1/SvJmj</td>
<td>J</td>
<td>MNO</td>
<td>HIJ</td>
</tr>
<tr>
<td>C3H/HeJc</td>
<td>J</td>
<td>KL</td>
<td>EFG</td>
</tr>
<tr>
<td>PL/J</td>
<td>NO</td>
<td>NO</td>
<td>GH</td>
</tr>
<tr>
<td>C57Bl/dJ</td>
<td>J</td>
<td>NO</td>
<td>GH</td>
</tr>
<tr>
<td>129X1/SvJ</td>
<td>K</td>
<td>O</td>
<td>GH</td>
</tr>
<tr>
<td>KK/HJ</td>
<td>L</td>
<td>LMNO</td>
<td>BC</td>
</tr>
<tr>
<td>MKL/Mpj</td>
<td>K</td>
<td>IJK</td>
<td>A</td>
</tr>
<tr>
<td>ILjJ</td>
<td>L</td>
<td>PQ</td>
<td>JK</td>
</tr>
<tr>
<td>NZW/LacJ</td>
<td>L</td>
<td>PQ</td>
<td>CD</td>
</tr>
<tr>
<td>NZO/HilLj</td>
<td>L</td>
<td>P</td>
<td>B</td>
</tr>
<tr>
<td>RIIIS/J</td>
<td>L</td>
<td>Q</td>
<td>KL</td>
</tr>
<tr>
<td>BUB/BnJ</td>
<td>L</td>
<td>PQ</td>
<td>CD</td>
</tr>
<tr>
<td>AJ</td>
<td>M</td>
<td>Q</td>
<td>GH</td>
</tr>
</tbody>
</table>

Significant differences are based on results from 1-way ANOVA followed by Tukey post hoc analysis. Strains not connected by same letter are significantly different. Strains are organized based on time from the longest time (A) to the shortest time (M). The number of strain groupings varies by phenotype.

with only the classical inbred strains. This was done to investigate the influence of the wild-derived strains on the genetic architecture of the exercise phenotypes. For the 132 K panel, significant associations for exercise time using all 34 strains were identified on chromosomes 2, 7, 11, and 13 (Fig. 3A, Table 2). When analyses were repeated on data from the classical inbred strains only, significant associations for exercise time were found on chromosomes 2 (168.40–168.43 Mb) and 11 (21.66–22.56 Mb) (Fig. 3B). Two associations on chromosome 11 (~25 Mb and ~70 Mb) and those on chromosomes 7 and 13 were no longer significant. Repeating association mapping for exercise time using the 4 M SNP panel identified two significant associations in the 34-strain cohort and four significant associations in the 30-strain cohort (Table 3). A novel significant association was identified on chromosome 1 in both the 34-strain and 30-strain groups and a suggestive (false discovery rate of 10%) association on chromosome 11 in the 30 classical inbred strains.

Genome-wide association mapping for work using the 132 K SNP panel is shown in Fig. 4. Only one significant association was identified for work in the 34-strain cohort (Chr. 2 at 168.4 Mb). No significant associations were identified using the 30 classical inbred strains only. Similar analyses using the 4 M SNP panel yielded no significant associations (data not shown).

To assess the contribution of body mass to QTL for exercise capacity, we conducted genome-wide association mapping using body mass from the 34 inbred strains and the 30 classical inbred strains. The variation among strains was greater for the 132 K panel, two significant associations in the 34-strain cohort and four significant associations in the 30-strain cohort (Table 3). A novel significant association was identified on chromosome 1 in both the 34-strain and 30-strain groups and a suggestive (false discovery rate of 10%) association on chromosome 11 in the 30 classical inbred strains.

Genome-wide association mapping for work using the 132 K SNP panel was performed using 132 K and 4 M SNP panels in EMMA. For each SNP panel, analyses were run two ways, one with all 34 strains including classical and wild-derived inbred mouse strains and a second genome-wide association mapping for exercise phenotypes was performed using 132 K and 4 M SNP panels in EMMA.
DISCUSSION

In general, efforts to determine the genetic basis for variation in exercise capacity have been limited due to small sample size in human studies (13) or limited genetic diversity in animal studies (2, 22, 25, 42). Advances in sequencing technology and SNP discovery have led to the development of large SNP datasets for inbred mice (10, 44). In turn, these datasets have facilitated the advancement of genome-wide association studies in mice (12, 16, 28). Therefore, the aim of this study was to identify strain-specific differences in intrinsic endurance exercise capacity across 34 strains of inbred mice and to identify significant QTL regulating intrinsic exercise capacity using genome-wide association mapping. Using this large strain set, we found significant variation for intrinsic endurance exercise capacity across strains. This variation allowed for the identification of novel QTL for exercise capacity, including one QTL on chromosome 2 that overlaps with a suggestive QTL for the response to exercise training (change in VO2max) in a human linkage study. The majority of the peak SNPs for these QTL are located within genes, suggesting that these genes are prospective candidates for influencing variation in intrinsic endurance exercise capacity. Although these QTL require confirmation by traditional approaches, they provide potential new targets for identifying the underlying genetic basis for variation in endurance exercise capacity.

In the current study, phenotype data from 34 classical and wild-derived inbred mouse strains were used to conduct a genome-wide association study for intrinsic endurance exercise capacity. This large phenotype dataset represents a three- to fourfold expansion of exercise capacity measurements over previous reports. Using male mice from seven inbred strains, Lerman et al. (19) reported a 1.7-fold difference in maximal running speed across strains. Lightfoot et al. (21) showed a similar degree of variation in running distance across 10 inbred strains. In the present study, we used 34 strains and included mice from all seven strain groupings from the mouse family tree (27), indicating a wide range of genetic diversity. Many of these strains were chosen based on their priority for inclusion in the Mouse Phenome Database (http://phenome.jax.org/) or their role as a founder strain for the Collaborative Cross recombinant inbred line panel (129S1/SvImJ, A/J, C57BL/6J, CAST/EiJ, NOD/LtJ, PWK/PhJ, WSB/EiJ) (http://compgen.unc.edu). There was a 2.5- to 3-fold difference between the highest and lowest performing strains with respect to time. This range is comparable to previous reports, despite differences in testing protocols and equipment. Only two strains, C57BL/6J and DBA/2J, were common among all three studies with the C57BL/6J strain ranking at or near the bottom in all three studies. High-performing strains FVB/NJ (19) and SWR/J (21) also performed well in our study, suggesting that these strains might be useful for future studies contrasting high (FVB, SWR) and low (C57BL/6J) performing strains.

Interestingly, despite being a low-performing strain in treadmill-based exercise tests (19, 21, 24, 25), C57BL/6J mice are generally considered a high-performing strain for wheel run-
ASSOCIATION MAPPING OF EXERCISE CAPACITY

Table 2. Genome-wide association mapping for exercise time in male mice from classical and wild-derived inbred strains using a 132 K SNP panel

<table>
<thead>
<tr>
<th>Chr.</th>
<th>QTL Location, Mb</th>
<th>Peak SNP</th>
<th>P Value</th>
<th>Known Genes, n</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Classical and wild-derived inbred strains</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>168.27–168.45</td>
<td>rs27288988</td>
<td>1.61 × 10^{-8}</td>
<td>1 (Nfatc2)</td>
</tr>
<tr>
<td>7</td>
<td>16.97–16.99</td>
<td>rs31142151</td>
<td>1.93 × 10^{-6}</td>
<td>2 (Sae1)</td>
</tr>
<tr>
<td>11</td>
<td>21.66–22.59</td>
<td>rs4135796</td>
<td>1.73 × 10^{-8}</td>
<td>9 (AV249152)</td>
</tr>
<tr>
<td>11</td>
<td>24.51–25.56</td>
<td>rs13480916</td>
<td>3.17 × 10^{-6}</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>70.26–70.81</td>
<td>rs7077722</td>
<td>3.73 × 10^{-6}</td>
<td>30 (Med11)</td>
</tr>
<tr>
<td>13</td>
<td>58.95</td>
<td>rs29804611</td>
<td>2.09 × 10^{-6}</td>
<td>2 (Ntrk2)</td>
</tr>
<tr>
<td></td>
<td>Classical inbred strains only</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>168.40–168.43</td>
<td>rs27288988</td>
<td>2.36 × 10^{-7}</td>
<td>1 (Nfatc2)</td>
</tr>
<tr>
<td>11</td>
<td>21.65–22.56</td>
<td>rs4135796</td>
<td>2.41 × 10^{-7}</td>
<td>9 (AV249152)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs3708339</td>
<td>2.41 × 10^{-7}</td>
<td>(AV249152)</td>
</tr>
</tbody>
</table>

All results are significant based on genome-wide thresholds corresponding to P values of 7.21 × 10^{-6} for 34-strain and 1.87 × 10^{-6} for 30-strain cohorts. Gene symbol in parentheses indicates the gene containing the peak SNP. Nfatc2, kinesin family member 26B; AV249152, WD repeat containing planar cell polarity effector; Med11, mediator of RNA polymerase II transcription, subunit 11 homolog; Auh, AU RNA binding protein/enooyl-coenzyme A hydratase; Ntrk2, neurotrophic tyrosine kinase receptor type 2.

Fig. 3. Genome-wide association mapping for endurance exercise capacity, expressed as time in minutes, in male mice from 34 classical and wild-derived inbred strains (A) and 30 classical inbred strains only (B) using the 132 K SNP panel. Association mapping was conducted with an empirical mixed model algorithm (EMMA) and a 132,285 SNP panel. The x-axis indicates genomic position divided by chromosome. Values on the y-axis are P values transformed using −log_{10}(P value). Horizontal dashed lines in each plot indicate genome-wide significance thresholds. Significance thresholds were calculated using a 5% false discovery rate and correspond to P values of 7.21 × 10^{-6} and 1.87 × 10^{-6} for 34 strain and 30 strain cohorts, respectively.

Fig. 4. Genome-wide association mapping for endurance exercise capacity, expressed as work in kg·m, in male mice from 34 classical and wild-derived inbred strains (A) and 30 classical inbred strains only (B) using the 132 K SNP panel. Association mapping was conducted with EMMA and a 132,285 SNP panel. The x-axis indicates genomic position divided by chromosome. Values on the y-axis are P values transformed using −log_{10}(P value). The horizontal dashed line in A indicates genome-wide significance threshold.

Table 3. Genome-wide association mapping for exercise time in male mice from classical and wild-derived inbred strains using a 4 million SNP panel

<table>
<thead>
<tr>
<th>Chr.</th>
<th>QTL Location, Mb</th>
<th>Peak SNP</th>
<th>P Value</th>
<th>Known Genes, n</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Classical and wild-derived inbred strains</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>180.76</td>
<td>rs39502136</td>
<td>1.01 × 10^{-7}</td>
<td>2 (Kif26b)</td>
</tr>
<tr>
<td>11</td>
<td>21.65–22.58</td>
<td>rs4135796</td>
<td>2.75 × 10^{-8}</td>
<td>(AV249152)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs26852365</td>
<td>2.75 × 10^{-8}</td>
<td>(Ehbp1)</td>
</tr>
<tr>
<td></td>
<td>Classical inbred strains only</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>178.54–180.76</td>
<td>rs39502136</td>
<td>2.43 × 10^{-8}</td>
<td>24 (Kif26b)</td>
</tr>
<tr>
<td>2</td>
<td>168.40–168.43</td>
<td>rs27288988</td>
<td>1.13 × 10^{-6}</td>
<td>1 (Nfatc2)</td>
</tr>
<tr>
<td>11</td>
<td>21.65–22.56</td>
<td>rs4135796</td>
<td>3.29 × 10^{-7}</td>
<td>(AV249152)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs3708339</td>
<td>3.29 × 10^{-7}</td>
<td>(AV249152)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs26852365</td>
<td>3.29 × 10^{-7}</td>
<td>(Ehbp1)</td>
</tr>
<tr>
<td>11</td>
<td>70.92–71.09</td>
<td>rs282333136</td>
<td>3.31 × 10^{-6}</td>
<td>3 (Nlrpc)</td>
</tr>
</tbody>
</table>

*P value corresponding to a genome-wide false discovery rate of ≤10% (q value of ≤0.10). All other results are significant based on genome-wide thresholds corresponding to P values of 1.01 × 10^{-7} for 34-strain and 2.54 × 10^{-8} for 30-strain cohorts. Gene symbol in parentheses indicates the gene containing the peak SNP. Kif26b, kinesin family member 26B; AV249152, WD repeat containing planar cell polarity effector; Ehbp1, EH domain binding protein 1; Nfatc2, nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2; Nlrpc, NLR family, pyrin domain containing 1C.
 ASSOCIATION MAPPING OF EXERCISE CAPACITY

Table 4. Genome-wide association mapping for body mass in male mice from classical and wild-derived inbred strains

<table>
<thead>
<tr>
<th>Chr.</th>
<th>Location, Mb</th>
<th>Peak SNP</th>
<th>Known Genes, n</th>
<th>Previous QTL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Classical and wild-derived inbred strains</td>
</tr>
<tr>
<td>1</td>
<td>54.08</td>
<td>rs6406541</td>
<td>1 (Hevc2)</td>
<td>1 (Hevc2)</td>
</tr>
<tr>
<td>3</td>
<td>109.58</td>
<td>rs29791710</td>
<td>2</td>
<td>Bwq8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Classical inbred strains only</td>
</tr>
<tr>
<td>2</td>
<td>59.19</td>
<td>rs6238344</td>
<td>5</td>
<td>Pbwg1, Bwq8</td>
</tr>
<tr>
<td>4</td>
<td>41.42–41.69</td>
<td>rs6222957</td>
<td>13</td>
<td>Wtr3, Bwq9, Wta1</td>
</tr>
<tr>
<td>9</td>
<td>93.76–98.14</td>
<td>rs6360988</td>
<td>28 (Acpl2)</td>
<td>Bwq6</td>
</tr>
<tr>
<td>9</td>
<td>99.78</td>
<td>rs6387071</td>
<td>3</td>
<td>Bwq6</td>
</tr>
<tr>
<td>15</td>
<td>29.20</td>
<td>rs13482501</td>
<td>0</td>
<td>Dob3</td>
</tr>
<tr>
<td>17</td>
<td>48.77</td>
<td>rs29537880</td>
<td>3</td>
<td>Obwpq4, W10q12, W6q11, Bodwt2, Wuw4</td>
</tr>
</tbody>
</table>

All associations with P value <10^{-8} are reported. Gene symbol in parentheses indicates the gene containing the peak SNP. For single SNP associations, the QTL interval for candidate gene identification was estimated to be 400 kb centered around the peak SNP. Hevc2, HECT, C2 and WW domain containing E3 ubiquitin protein ligase 2; Acpl2, acid phosphatase-like 2; Bwq8, body weight QTL 8 (39); Pbwg1, postnatal body weight growth 1 (55); Bwq8, body weight QTL 8 (36); Wtr3, weight loss response 3 (32); Bwq9, body weight QTL 9 (36); Wta1, weight adult 1 (7); Bwq6, body weight QTL 6 (30); Dob3, dietary obesity 3 (43); Obwpq4, obesity and body weight QTL 4 (36); W10q12, weight 10 wk QTL 12 (33); W6q11, weight 6 wk QTL 11 (33); Bodwt2, body weight 2 (38); Wuw4, weight adult 4 (7).

Physiol Genomics • doi:10.1152/physilgenomics.00051.2012 • www.physiolgenomics.org
types in male and female mice from 38 inbred and wild-derived strains. This number was notably reduced after removing the wild-derived strains. Although the use of a large number of inbred strains and a high-density SNP map is recommended for genome-wide association mapping (41), the effect of wild-derived strains on association mapping observed by Lightfoot and by us suggests that it would be prudent to conduct analyses with and without these strains to identify significant associations and the effect of population structure on the genetic architecture of complex traits.

QTL were identified for exercise capacity by using EMMA to perform association mapping. One criticism of genome-wide association mapping approaches including single SNP marker and haplotype mapping, which relies on a three-SNP window to define haplotype blocks, is that it can be affected by population structure, sparse SNP data, regions of low polymorphisms, or high haplotype diversity, potentially resulting in a large number of false positives (41). Although EMMA also has limitations (45), it was selected for association mapping because it incorporates a kinship matrix that corrects for relatedness among inbred strains (16). This should correct for differences in population structure due to the interrelatedness of inbred strains, which has been shown to reduce the number of false positive associations compared with other mapping approaches (16). Although EMMA has been used successfully to identify significant associations for a variety of physiological and disease-related traits (3, 9, 16), traditional linkage analysis needs to be performed to confirm results from mouse genome-wide association studies (23, 37).

In summary, we performed a large strain survey for exercise capacity using a graded treadmill test. This study demonstrated that a large strain survey and association mapping could be used to identify new QTL for exercise capacity. As with most genome-wide association studies, replication studies are needed to confirm our findings. One option is to use the strain survey data presented here, coupled with large SNP databases to identify the most appropriate strains for traditional linkage analysis to confirm these QTL. In addition, the narrow QTL regions identified here contain a small number of potential candidate genes that can be systematically studied to confirm their role in modulating exercise capacity. Future studies confirming these QTL and identifying the physiological relationship between candidate gene and exercise capacity are required to establish this gene or genetic variant as being the underlying cause of differences in exercise capacity.

GRANTS
This work was supported by National Heart, Lung, and Blood Institute Grant HL-085918 (M. P. Massett).

DISCLOSURES
No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS
Author contributions: S.M.C. and M.P.M. conception and design of research; S.M.C. and M.P.M. performed experiments; S.M.C. and M.P.M. analyzed data; S.M.C. and M.P.M. interpreted results of experiments; S.M.C. and M.P.M. prepared figures; S.M.C. and M.P.M. drafted manuscript; S.M.C. and M.P.M. edited and revised manuscript; S.M.C. and M.P.M. approved final version of manuscript.

REFERENCES


