Title

A novel intronic single nucleotide polymorphism in the myosin heavy polypeptide 4 gene is responsible for the mini-muscle phenotype characterized by major reduction in hind-limb muscle mass in mice.

Journal

Genetics, 195(4)

Authors

Kelly, Scott

Bell, Timothy

Selitsky, Sara

et al.

Publication Date

2013-12-01

DOI

10.1534/genetics.113.154476

Peer reviewed

A Novel Intronic Single Nucleotide Polymorphism in the *Myosin heavy polypeptide 4* Gene Is Responsible for the Mini-Muscle Phenotype Characterized by Major Reduction in Hind-Limb Muscle Mass in Mice

Scott A. Kelly,*[,1] Timothy A. Bell,[†] Sara R. Selitsky,[†] Ryan J. Buus,[†] Kunjie Hua,[†] George M. Weinstock,[‡] Theodore Garland Jr.,[§] Fernando Pardo-Manuel de Villena,[†] and Daniel Pomp[†]

*Department of Zoology, Ohio Wesleyan University, Delaware, Ohio 43015, [†]Department of Genetics and Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, North Carolina 27599, [‡]The Genome Institute, Washington University School of Medicine, St. Louis, Missouri 63108, and [§]Department of Biology, University of California, Riverside, California 92521

ABSTRACT Replicated artificial selection for high levels of voluntary wheel running in an outbred strain of mice favored an autosomal recessive allele whose primary phenotypic effect is a 50% reduction in hind-limb muscle mass. Within the High Runner (HR) lines of mice, the numerous pleiotropic effects (*e.g.*, larger hearts, reduced total body mass and fat mass, longer hind-limb bones) of this hypothesized adaptive allele include functional characteristics that facilitate high levels of voluntary wheel running (*e.g.*, doubling of mass-specific muscle aerobic capacity, increased fatigue resistance of isolated muscles, longer hind-limb bones). Previously, we created a backcross population suitable for mapping the responsible locus. We phenotypically characterized the population and mapped the *Minimsc* locus to a 2.6-Mb interval on MMU11, a region containing ~100 known or predicted genes. Here, we present a novel strategy to identify the genetic variant causing the mini-muscle phenotype. Using high-density genotyping and whole-genome sequencing of key backcross individuals and HR mice with and without the mini-muscle mutation, from both recent and historical generations of the HR lines, we show that a SNP representing a C-to-T transition located in a 709-bp intron between exons 11 and 12 of the *Myosin heavy polypeptide 4* (*Myh4*) skeletal muscle gene (position 67,244,850 on MMU11; assembly, December 2011, GRCm38/mm10; ENSMUSG00000057003) is responsible for the mini-muscle phenotype, $Myh4^{Minimsc}$. Using next-generation sequencing, our approach can be extended to identify causative mutations arising in mouse inbred lines and thus offers a great avenue to overcome one of the most challenging steps in quantitative genetics.

IN 1993, replicated artificial selection for high voluntary wheel-running behavior was implemented, beginning from a base population of 224 outbred Hsd:ICR mice (Swallow *et al.* 1998). The goal was to elucidate the underpinnings of voluntary exercise at the levels of physiology, neurobiology, and motivation and eventually to identify genetic factors underlying these intermediate phenotypes. The selection criterion was the total revolutions run during days 5 and 6 of a 6-day period of wheel access when mice were young adults (~6–8 weeks of age). After 10 generations of selection, mice from the four replicate High Runner (HR) lines were running >70% more revolutions per day (on days 5 and 6) as compared with four nonselected control (C) lines (Swallow *et al.* 1998). Wheel running continued to increase in the HR lines until generation 16–28 (depending on line and sex) and remained approximately at those levels through a subsequent 50 generations of selective breeding (Middleton *et al.* 2008; Swallow *et al.* 2009; Kolb *et al.* 2010; Careau *et al.* 2012, 2013).

Dissections to monitor potential correlated selection responses in organ masses led to the discovery of a mini-muscle

doi: 10.1534/genetics.113.154476
Manuscript received July 10, 2013; accepted for publication September 10, 2013; published Early Online September 20, 2013.
Supporting Information is available online at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.113.154476/-/DC1.
[1]Corresponding author: Ohio Wesleyan University, Schimmel/Conrades Science Center #346, 61 S. Sandusky St., Delaware, OH 43015. E-mail: sakelly@owu.edu

phenotype that is caused by an autosomal recessive mutation (Garland *et al.* 2002; Hannon *et al.* 2008). The most notable feature of this phenotype is an ~50% reduction in mass of the triceps surae muscle complex (gastrocnemius, plantaris, soleus) and in the mass of the entire hind-limb musculature (Houle-Leroy *et al.* 2003; Swallow *et al.* 2005). Hence, the allele qualifies as a "gene of major effect" (*e.g.*, Agrawal *et al.* 2001; Hannon *et al.* 2008).

The mini-muscle phenotype was observed in two of the four HR lines and in one of the four C lines (Table 1). Initially, the phenotype remained at low frequency in the C line and eventually disappeared, but increased dramatically in frequency in both HR lines, eventually going to fixation in one of the two HR lines by approximately generation 36 (Syme *et al.* 2005). Evidence for the adaptive significance of the allele was obtained by fitting multiple-generation data to hierarchical models that included effects of selection and/or random genetic drift (Garland *et al.* 2002). The mini-muscle allele was estimated to have been present at a frequency of ~7% in the base population, and analyses indicated that strong selection favored the allele in the HR but not in the C lines. It was hypothesized that the mini-muscle phenotype has functional characteristics that facilitate high levels of wheel running (Garland *et al.* 2002), such as a doubled mass-specific aerobic capacity (Houle-Leroy *et al.* 2003) and increased fatigue resistance in at least some hind-limb muscles (Syme *et al.* 2005). Alternatively, the underlying allele might have pleiotropic effects that act on nonmuscle tissues and organs but still facilitate endurance running (*e.g.*, reduced total body mass and fat mass; increased relative heart, liver, and kidney mass; longer hind-limb bones) (Garland *et al.* 2002; Kelly *et al.* 2006; Hannon *et al.* 2008; Kolb *et al.* 2010). However, at generation 22, wheel running of affected individuals did not differ statistically from those with normal-sized muscles (Garland *et al.* 2002), and the magnitude of response to selection was not systematically higher in the two lines that had the mini-muscle phenotype as compared with the two that did not, indicating that multiple genetic "solutions" are possible in response to selection for high levels of voluntary exercise (Garland *et al.* 2011a; Careau *et al.* 2013). To date, obvious deleterious effects of the mini-muscle phenotype have not been identified (*e.g.*, see Girard *et al.* 2002), which is unusual for a gene of major effect (*e.g.*, Carriére *et al.* 1994; Rivero *et al.* 2011).

Previously, we used a backcross population between HR (line 3) and the C57BL/6J inbred strain (Hannon *et al.* 2008) to map the causative locus known as *Minimsc* (MGI:44204080) to a 2.6-Mb interval on MMU11, a region containing ~100 known or predicted genes, many of which have known roles in muscle development and/or function (Hartmann *et al.* 2008). Here, we present a novel strategy to identify the genetic variant(s) causing the mini-muscle phenotype. Using high-density genotyping in a small but carefully selected sample of 10 mice, we first identified two maximally informative backcross samples for whole-genome sequencing, and subsequently we used the sequence data to identify two putative causal SNPs. Using targeted genotyping of these two SNPs in large cohorts of HR mice with and without the mini-muscle mutation, from both recent and historical generations of the HR lines, we show a SNP representing a C-to-T transition located in a 708-bp intron between exons 11 and 12 of the *Myosin heavy polypeptide 4* (*Myh4*) skeletal muscle gene (position 67,244,850 on MMU11; assembly, December 2011, GRCm38/mm10; ENSMUSG00000057003) is responsible for the mini-muscle phenotype. The symbol of this allele is $Myh4^{Minimsc}$.

Table 1 Frequency of the mini-muscle phenotype in four replicate lines selectively bred for high voluntary wheel running (High Runner, HR) and four nonselected lines that serve as their controls (Control, C) in the breeding program

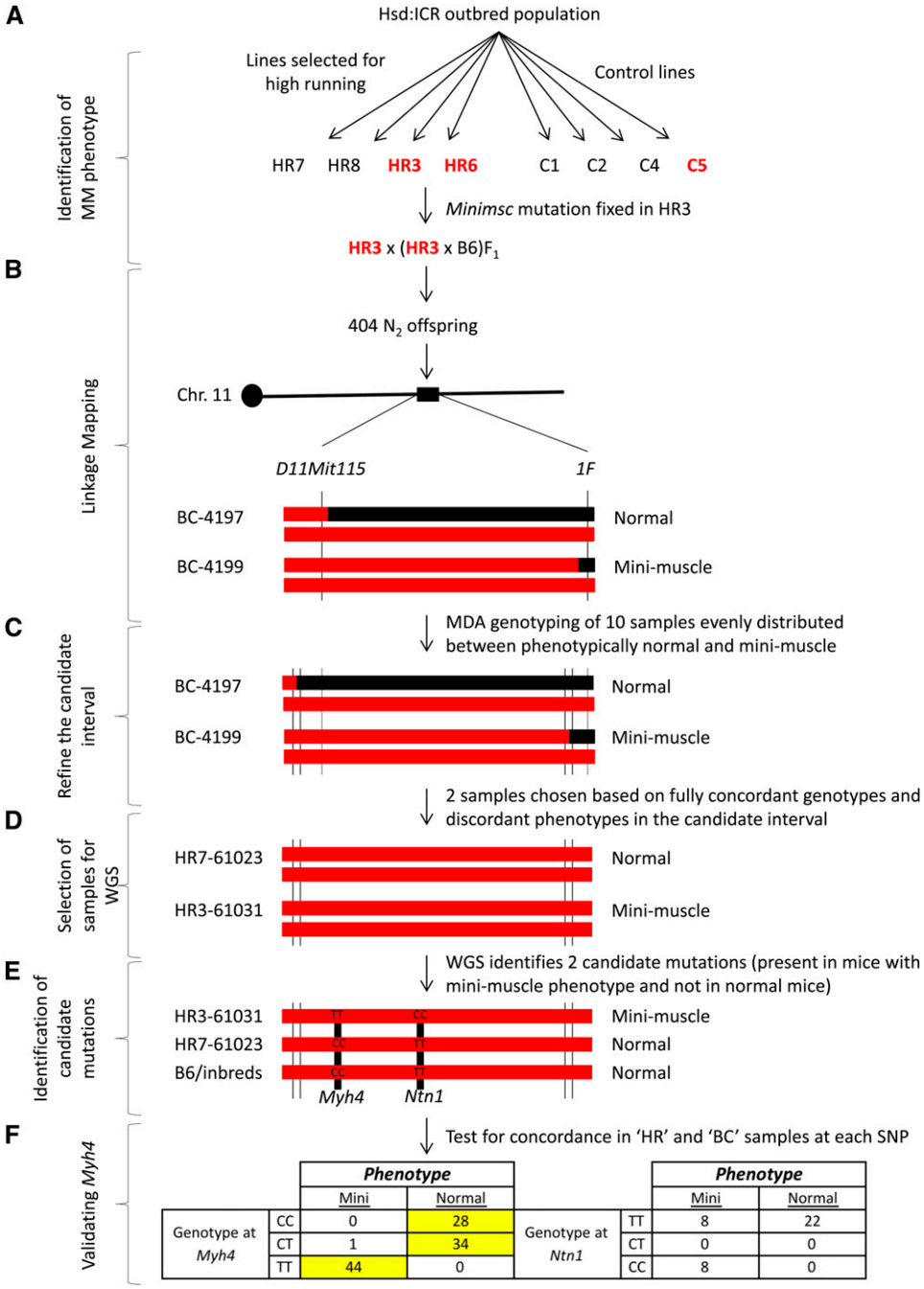
Line	Linetype	Mini-muscle status
1	Control	Never observed
2	Control	Never observed
4	Control	Never observed
5	Control	Observed at least through generation 22
3	High runner	Eventually became fixed for mini phenotype
6	High runner	Still polymorphic for mini phenotype at generation 66
7	High runner	Never observed
8	High runner	Never observed

Materials and Methods

High-density genotyping using the mouse diversity array

To confirm and refine the *Minimsc* candidate interval reported previously (Hartmann *et al.* 2008), we genotyped 10 mice with the high-density mouse diversity array (MDA) (Yang *et al.* 2009). Four of these mice were selected from the backcross (BC) (Hannon *et al.* 2008) and used to map *Minimsc*, while the other six mice were selected from the four replicate HR lines (Figure 1). Samples were evenly distributed between phenotypically normal animals (BC-4197, HR7-61023, HR7-61148, HR8-61095, and HR8-61102) and animals exhibiting the mini-muscle phenotype (BC-4086, BC-4032, BC-4199, HR3-61031, and HR6-527). Three of these BC samples provided the critical recombination events used to define the proximal (BC-4197 and BC-4086) and distal (BC-4199) boundaries for the *Minimsc* candidate interval reported previously (Hartmann *et al.* 2008).

High-quality, high-molecular weight DNA was extracted using phenol–chloroform. DNA concentration was normalized to 50 ng/μl, processed according to the Affymetrix Genome-Wide Human SNP Nsp/Sty Assay Kit 5.0/6.0 protocol, and hybridized to the Affymetrix MDA genotyping array (Yang *et al.* 2009) at the Functional Genomics Core Facility at the University of North Carolina at Chapel Hill (UNC) as previously described (Aylor *et al.* 2011). SNP genotype calling was conducted as described previously (Yang *et al.* 2009). The *Minimsc* candidate interval on MMU11 was refined using the



Genotype at *Myh4*	Phenotype: Mini	Phenotype: Normal
CC	0	28
CT	1	34
TT	44	0

Genotype at *Ntn1*	Phenotype: Mini	Phenotype: Normal
TT	8	22
CT	0	0
CC	8	0

Figure 1 Flow diagram detailing each successive step, from phenotype to genotype, in the identification of *Myh4*^{Minimsc}. (A) The mini-muscle was first observed in some of the HR and C lines (shown in red) but not in others (shown in black). The phenotype became fixed in the HR3 line. (B) A large backcross between HR3 and C57BL/6J was used to map the *Minimsc* locus. Two mice, BC-4197 and BC-4199, provided the initial boundaries for the *Minimsc* candidate interval. (C) High-density genotyping of these two samples provided more accurate boundaries for the candidate interval, including the proximal extension of the candidate interval. (D) Based on high-density genotyping, two HR samples with identical genotypes in the candidate interval but contrasting muscle phenotype were selected for whole-genome sequencing. (E) Identification of candidate genetic variants for the mini-muscle phenotype. (F) Discrimination between candidate variants and validation of *Myh4* as the candidate gene.

heterozygous regions of mice BC-4197 and BC-4199 (BC mice delimiting the candidate region; see Hartmann *et al.* 2008) to fine-map boundaries of the candidate region (Figure 1).

Hybridization intensities at every probe within the *Minimsc* candidate interval were compared between the mini-muscle and normal mice in an effort to identify potential deletions and variable intensity oligonucleotides (Yang *et al.* 2011; Didion *et al.* 2012). Intensity plots spanning from 66.924 to 70.027 Mb on MMU11 among all pairwise comparisons of a mini-muscle and a normal sample were analyzed. Probes with differential hybridization were noted for each pair (mini-muscle and normal) and then compared across the whole population. Mini-muscle, normal, and standard inbred strains with whole-genome sequence publicly available [referred as the Sanger strain set (Keane *et al.* 2011)] were compared using probe intensity plots and genotype calls to select the pair of samples with contrasting phenotypes (*i.e.*, mini-muscle and normal) and most similar probe intensity plots for whole-genome sequencing and analysis.

Whole-genome sequencing and alignment

Five micrograms of high-molecular-weight DNA from samples HR3-61031 and HR7-61023 were shipped from UNC to The Genome Institute at Washington University for

whole-genome sequencing. Genomic DNA was used to construct two Illumina HiSeq libraries, using 0.5 μg starting material, with 300- to 400- and 400- to 500-bp fragment sizes. Each library was sequenced on four lanes of an Illumina HiSeq flowcell, as paired-end reads, with 100-bp read lengths. This produced 155 Gb of high-quality sequence with 41.308× haploid coverage for HR3-61031 and 120 Gb of high-quality sequence with 32.188× haploid coverage for HR7-61023. We aligned the sequences to the University of California at Santa Cruz Mouse Build mm9 using ELAND to assess base accuracy; BAM files were then produced.

We used Bowtie version 0.12.7 and aligned the newly generated sequence data to the mouse reference sequence [July 2007 (NCBI37/mm9)] under the following conditions: (1) Alignments may have no more than two mismatches in the first 28 bases on the high-quality (left) end of the read (the seed). (2) The sum of the Phred quality values at all mismatched positions may not exceed 70. Although we aligned sequence data to assembly [July 2007 (NCBI37/mm9)], all positions have been converted and are presented here to reflect the most recent assembly [December 2011 (GRCm38/mm10)]. Only a portion of the whole-genome sequence was utilized to identify candidate mutations for *Minimsc*, and these sequences are provided for HR3-61031 and HR7-61023 in the Supporting Information (File S1 and File S2, respectively). File S1 and File S2 are provided in pileup format and are accompanied by a word document with an explanation of the SAMtools pileup format (File S3).

Identification of candidate mutations for Minimsc

We used consistency and coverage as quality-control metrics to determine the usefulness of the sequence data. The two samples used in this analysis, HR3-61031 and HR7-61023, are homozygous in the *Minimsc* candidate interval and thus are expected to be homozygous at every position. We classified a given position as consistent when >65% of reads at that position had the same genotype call. Both samples were consistent at >99.5% of positions, and consistency was highly correlated between samples across the candidate interval. Similarly, coverage in each sample varied substantially across the candidate interval in both samples, but coverage was highly correlated between the two samples at a given position. To identify candidate SNPs, we used a custom script to find any position with more than a 10-read coverage in any of the two samples that had different genotype calls between the two samples.

Genotyping at Ntn1 and Myh4 SNPs

To validate and extend the whole-genome sequence results, we genotyped a diverse set of samples using dideoxy Sanger sequencing of PCR amplicons spanning each one of the two candidate SNPs. PCR primers were designed to amplify 289- and 323-bp-long PCR products in the *Myh4* and *Netrin 1* (*Ntn1*) genes, respectively. Primers were designed using Primer Quest software (http://www.idtdna.com). The primers for the *Myh4* SNP are 5′-AGGTGGCCATTTACAAGCTCACAG-3′ and 5′-GGGAATGGACAGAAACAGTAAAGGTCC-3′ (MMU11: 67,244,634–67,244,922). The primers for the *Ntn1* SNP are 5′-AGGATTTGCTGTCTGGCCCTCATT-3′ and 5′-ATGAAGTGCTCCTGGGTGCTTTCT-3′ (MMU11: 68,282,072–68,282,395). The PCR fragment was amplified using standard conditions (initial denaturation at 95° for 5 min, followed by 35 cycles of 95°, 55°, and 72° for 30 sec each, and a final extension at 72° for 10 min) in an Applied Biosystems Veriti 96 Well Thermal Cycler. After the PCR amplification, the product was purified using a Promega SV gel and Wizard Clean Up System. Sequencing of the purified PCR product was performed at the UNC Genome Analysis Core Facility using the Applied Biosystems 3730xl Genetic Analyzer. Sequences were then analyzed using Sequencher.

Samples genotyped from various current and historical generations of the HR lines are summarized in Table 2. In addition, the following inbred strains were genotyped using the *Myh4* and *Ntn1* primers: BULS, BUSNA, CALB/RkJ, CZECHI/EiJ, CZECHII/EiJ, DDY/JclSidSeyFrkJ, JFI/Ms, KK/HlJ, LEWES/EiJ, LG/J, MOLD/DnJ, MOLF/EiJ, NON/ShiJ, NZW/LacJ, PERC/EiJ, POHN, PWD/PhJ, RBF/DnJ, RIIIS/J, SEB, SEG, SFM, SKIVE/EiJ, SMON, STLT, STUF, WMP/PasDnJ, XBS, 129S1/SvImJ, AKR/J, A/J, BALB/cJ, C3H/HeJ, C57BL/6J, CAST/EiJ, CBA/J, DBA/2J, FVB/NJ, LP/J, NOD/ShiLtJ, NZO/HILtJ, PWK/PhJ, SPRET/EiJ, and WSB/EiJ. Finally, genotypes from the following lines were obtained from reported whole-genome sequencing (Keane *et al.* 2011): 129P2/OlaHsd, 129S1/SvImJ, 129S5SvEbBrd, AKR/J, A/J, BALB/cJ, C3H/HeJ, C57BL/6NJ, CAST/EiJ, CBA/J, DBA/2J, LP/J, NOD/ShiLtJ, NZO/HILtJ, PWK/PhJ, SPRET/EiJ and WSB/EiJ. The phylogenetic origin of the haplotype in the candidate region in these strains has been determined using MDA (Yang *et al.* 2011)

Results

High-density genotyping refines the candidate interval for Minimsc

The MDA genotypes of samples BC-4197 and BC-4199 provide the new proximal and distal boundaries, respectively, for the *Minimsc* candidate interval (Figure 2). Specifically, the interval starts between positions 67,110,081 and 67,125,771 [defined by MDA markers JAX00314160 and JAX00314161 (rs26906842), respectively] and ends between 70,209,794 and 70,210,630 [MDA markers JAX00028760 (rs29423020) and JAX00028761 (rs26921141), respectively]. The MDA genotype analysis refines the boundaries, but it also expands the previous candidate interval from the ~2.6 Mb reported in Hartmann *et al.* (2008) to ~3.1 Mb (Figure 2). This discrepancy is explained by a strong likelihood that the one microsatellite marker providing the original proximal boundary (*D11Mit115*) was yielding incorrect genotypes in BC-4197 and BC-4086 (compare the respective genotypes in Figure 2 with those reported in table 5 in Hartmann *et al.* 2008). Note also that this extended and better-annotated interval was critical for the identification of the causal variant (*i.e.*, the causative SNP lies outside the initial candidate interval) (Figure 2 and see below).

Table 2 Genotyped samples from current and historical generations of replicated lines selectively bred for high voluntary wheel running

Sample	Line	Generation	*Ntn1* SNP[a]	*Myh4* SNP	Expected muscle phenotype	Observed muscle phenotype	Concordance for *Myh4*
64358	1	61	T	C	Normal	Normal	Yes
64426	1	61	T	C	Normal	Normal	Yes
64485	1	61	T	Het	Normal	Normal	Yes[b]
64382	2	61	T	C	Normal	Normal	Yes
64728	2	61	T	C	Normal	Normal	Yes
25087	3	22	ND	T	Mini	Mini	Yes
25088	3	22	ND	T	Mini	Mini	Yes
25091	3	22	ND	T	Mini	Mini	Yes
25101	3	22	ND	T	Mini	Mini	Yes
25126	3	22	ND	Het	Normal	Normal	Yes
25127	3	22	ND	C	Normal	Normal	Yes
25243	3	22	ND	Het	Normal	Normal	Yes
25285	3	22	ND	T	Mini	Mini	Yes
25289	3	22	ND	T	Mini	Mini	Yes
25324	3	22	ND	Het	Normal	Normal	Yes
25330	3	22	ND	T	Mini	Mini	Yes
25359	3	22	ND	T	Mini	Mini	Yes
25360	3	22	ND	T	Mini	Mini	Yes
25435	3	22	ND	T	Mini	Mini	Yes
25440	3	22	ND	C	Normal	Normal	Yes
25460	3	22	ND	T	Mini	Mini	Yes
25465	3	22	ND	T	Mini	Mini	Yes
25474	3	22	ND	T	Mini	Mini	Yes
45295	3	42	C	T	Mini	Mini	Yes
45580	3	42	C	T	Mini	Mini	Yes
45581	3	42	C	T	Mini	Mini	Yes
45588	3	42	C	T	Mini	Mini	Yes
61031	3	58	C	T	Mini	Mini	Yes
64262	3	61	C	T	Mini	Mini	Yes
64301	3	61	C	T	Mini	Mini	Yes
64596	3	61	C	T	Mini	Mini	Yes
64170	4	61	T	C	Normal	Normal	Yes
64352	4	61	T	C	Normal	Normal	Yes
25001	5	22	ND	C	Normal	Normal	Yes[c]
25004	5	22	ND	C	Normal	Normal	Yes[c]
25120	5	22	ND	Het	Normal	Normal	Yes[c]
25122	5	22	ND	Het	Normal	Normal	Yes[c]
25182	5	22	ND	Het	Normal	Mini	NO[d]
25185	5	22	ND	C	Normal	Normal	Yes[c]
25279	5	22	ND	Het	Normal	Normal	Yes[c]
25281	5	22	ND	T	Mini	Mini	Yes[c]
25293	5	22	ND	C	Normal	Normal	Yes[c]
25294	5	22	ND	C	Normal	Normal	Yes[c]
25319	5	22	ND	C	Normal	Normal	Yes[c]
25444	5	22	ND	Het	Normal	Normal	Yes[c]
25447	5	22	ND	Het	Normal	Normal	Yes[c]
25449	5	22	ND	Het	Normal	Normal	Yes[c]
25537	5	22	ND	C	Normal	Normal	Yes[c]
25539	5	22	ND	C	Normal	Normal	Yes[c]
25574	5	22	ND	Het	Normal	Normal	Yes[c]
25575	5	22	ND	Het	Normal	Normal	Yes[c]
64406	5	61	T	C	Normal	Normal	Yes[c]
64413	5	61	T	C	Normal	Normal	Yes[c]
257427	5	22	ND	C	Normal	Normal	Yes[c]
25031	6	22	ND	Het	Normal	Normal	*Yes*[e]
25038	6	22	ND	C	Normal	Normal	Yes[e]
25039	6	22	ND	Het	Normal	Normal	Yes[e]
25040	6	22	ND	Het	Normal	Normal	Yes[e]
25072	6	22	ND	T	Mini	Mini	Yes[e]
25075	6	22	ND	Het	Normal	Normal	Yes[e]
25193	6	22	ND	T	Mini	Mini	Yes[e]
25196	6	22	ND	T	Mini	Mini	Yes[e]

(continued)

Table 2, *continued*

Sample	Line	Generation	*Ntn1* SNP[a]	*Myh4* SNP	Expected muscle phenotype	Observed muscle phenotype	Concordance for *Myh4*
25478	6	22	ND	Het	Normal	Normal	Yes[e]
25483	6	22	ND	Het	Normal	Normal	Yes[e]
25584	6	22	ND	Het	Normal	Normal	Yes[e]
25586	6	22	ND	Het	Normal	Normal	Yes[e]
25598	6	22	ND	C	Normal	Normal	Yes[e]
25601	6	22	ND	Het	Normal	Normal	Yes[e]
25606	6	22	ND	T	Mini	Mini	Yes[e]
25615	6	22	ND	Het	Normal	Normal	Yes[e]
25633	6	22	ND	T	Mini	Mini	Yes[e]
25636	6	22	ND	T	Mini	Mini	Yes[e]
25652	6	22	ND	T	Mini	Mini	Yes[e]
25658	6	22	ND	T	Mini	Mini	Yes[e]
40035	6	37	T	C	Normal	Normal	Yes[e]
40036	6	37	T	Het	Normal	Normal	Yes[e]
40038	6	37	T	C	Normal	Normal	Yes[e]
40040	6	37	T	T	Mini	Mini	Yes[e]
40042	6	37	T	T	Mini	Mini	Yes[e]
40044	6	37	T	T	Mini	Mini	Yes[e]
41890	6	38	T	T	Mini	Mini	Yes[e]
41892	6	38	T	T	Mini	Mini	Yes[e]
41895	6	38	T	Het	Normal	Normal	Yes[e]
41896	6	38	T	Het	Normal	Normal	Yes[e]
41897	6	38	T	Het	Normal	Normal	Yes[e]
41899	6	38	T	T	Mini	Mini	Yes[e]
64200	6	61	T	Het	Normal	Normal	Yes[e]
64201	6	61	T	T	Mini	Mini	Yes[e]
64214	6	61	T	Het	Normal	Normal	Yes[e]
64222	6	61	T	T	Mini	Mini	Yes[e]
64839	6	61	T	Het	Normal	Normal	Yes[e]
66015	6	63	ND	T	Mini	Mini	Yes[e]
66320	6	63	ND	Het	Normal	Normal	Yes[e]
66326	6	63	ND	C	Normal	Normal	Yes[e]
66333	6	63	ND	Het	Normal	Normal	Yes[e]
66346	6	63	ND	Het	Normal	Normal	Yes[e]
66347	6	63	ND	Het	Normal	Normal	Yes[e]
66348	6	63	ND	Het	Normal	Normal	Yes[e]
66355	6	63	ND	T	Mini	Mini	Yes[e]
66357	6	63	ND	T	Mini	Mini	Yes[e]
66360	6	63	ND	T	Mini	Mini	Yes[e]
66367	6	63	ND	T	Mini	Mini	Yes[e]
66590	6	63	ND	T	Mini	Mini	Yes[e]
61023	7	58	T	C	Normal	Normal	Yes
64103	7	61	T	C	Normal	Normal	Yes
64135	7	61	T	C	Normal	Normal	Yes
64328	8	61	T	C	Normal	Normal	Yes

Samples were genotyped (C; T; Het, heterozygous) for *Myh4* and *Ntn1*. The expected hind-limb muscle phenotype (based on genotype) was compared with the actual phenotype revealed post dissection. Concordance between the expected muscle phenotype based on the *Myh4* genotype and the observed phenotype is presented in the far right column.

[a] ND, sample not genotyped for respective SNP.

[b] There is concordance between the *Myh4* genotype and the observed phenotype. However, despite the presence of the variant in this sample, the mini-muscle phenotype has never been observed in line 1.

[c] There is concordance between the *Myh4* genotype and the observed phenotypes in all cases except one. Note that the mini-muscle phenotype was observed at least through generation 22 in line 5.

[d] We hypothesize that there is an error in identification, recording, or labeling involving the animal, tissue sample, or extracted DNA from this discordant individual. Note that this sample originated from the historical samples dating back to generation 22 of selection in the year 1999.

[e] Note that line 6 is still polymorphic for the mini-muscle phenotype as of generation 66.

High-density genotyping identifies the pair of optimal samples for identification of the candidate mutation Minimsc through whole-genome sequencing

The six samples genotyped with MDA from four HR selection line replicates are all homozygous within the *Minimsc* candidate interval (Figure 2). All samples from the HR7, HR3, and HR6 lines have identical haplotypes throughout the entire region (Figure 2, HR red haplotype), despite the fact that HR7 has the normal phenotype while HR3 and HR6 have the mini-muscle phenotype. These results confirm that *Minimsc* arose as a spontaneous mutation within the same genetic background shared by different HR lines. The HR8

line has a different haplotype (Figure 2, HR blue haplotype) starting between 69,818,782 and 69,818,849 (JAX00314705 and JAX00314706, respectively) and continuing through the end of the *Minimsc* candidate interval. This change in haplotype covering the distal 392 kb of the *Minimsc* candidate region is unrelated to the mini-muscle phenotype. Analysis of SNP and exon probe intensity plots shows no evidence of large deletions or insertions in the region (results not shown).

Genotype calls and intensity plots were also used to determine which pair of HR samples showed the strongest evidence for identity by decent (IBD) within the *Minimsc* candidate interval between samples with the mini-muscle and normal phenotypes. This analysis indicated that the closest match was between sample HR3-61031 with the mini-muscle phenotype and sample HR7-61023 with the normal phenotype (Figure 2). Only three probes had significantly different probe intensities between these two samples. Other pairwise combinations involving combinations of HR samples with discordant mini-muscle phenotypes have additional probes with significantly different intensities. Therefore, we concluded that HR3-61031 and HR7-61023 represented the best pair of samples that were genotypically matched but phenotypically discordant. Thus, given the available choices, these two samples are optimal for directly identifying candidate mutations for *Minimsc* as any genetic variant that distinguishes them within the new boundaries of the candidate interval using whole-genome sequencing. We also conclude that among the set of Sanger inbred strains (Keane *et al.* 2011) the CBA/J strain shows the strongest evidence for IBD in the *Minimsc* candidate region with HR3-61031 based on MDA genotypes (data not shown).

Whole-genome sequence identifies two candidate mutations for Minimsc

We used the aligned reads from the two sequenced samples (HR3-61031 and HR7-61023) to identify candidate genetic variants for *Minimsc*. We first determined the coverage over the candidate and flanking regions (at least 100 kb proximal and distal flanking regions; Figure S1). We used these data to determine the threshold used in variant calling (see below). We also determined whether there was differential coverage between these two samples in any 500-bp windows. We found strong evidence of a structural variant at the *Nlrp1* gene family locus (71.09–71.29 Mb). This result demonstrates our ability to detect copy-number variations. However, this deletion can be ruled out as underlying *Minimsc* because it is shared by the two samples with discordant phenotypes and it is located outside of the candidate interval.

In the 3.1-Mb *Minimsc* refined candidate interval, we identified only two SNPs with different alleles in our two samples. These two SNPs are located at positions 67,244, 850 and 68,282,254. The first SNP is a C-to-T transition located in a 709-bp intron between exons 11 and 12 of the *Myh4* skeletal muscle gene. The second SNP is a T-to-C transition in a very long intron (>100 kb) separating exons 3 and 4 of the *Ntn1* gene. The variant allele at both SNPs is found in the HR3-61031 sample that has the mini-muscle phenotype [*i.e.*, this allele is not found in the mouse reference genome nor in any inbred strain genotyped or sequenced at that position (Table S1)], and it is well supported by forward and reverse reads with a good read coverage (minimum coverage is 24 reads).

Because it was not immediately obvious whether either of these SNPs would have functional consequences, we explored the possibility of additional undiscovered SNPs in regions with low coverage or at positions with inconsistent genotype calls within a sample. We could reject both scenarios with a high degree of confidence by comparing coverage and the ratio of inconsistent genotypes between the two sequenced samples. Coverage and inconsistent SNPs were almost identical at both samples. Additionally, there was no evidence of any insertion in the HR3-61031 sample based on disruptions of forward and reverse read density (data not shown). We concluded that one of the two SNPs in either the *Myh4* or the *Ntn1* genes was the cause of the mini-muscle phenotype and thus represents *Minimsc*.

Extensive genotyping in additional samples excludes the Ntn1 SNP and provides independent evidence supporting the Myh4 SNP as Minimsc

The variant allele at both the *Ntn1* and *Myh4* SNPs is not present among 46 inbred strains genotyped by us or for which genome sequence had been reported recently (Table S1; Keane *et al.* 2011). These strains represent a wide variety of classical and wild-derived stocks from multiple *Mus* taxa including *M. musculus musculus*, *M. m. domesticus*, *M. m. castaneus*, *M. spretus*, and *M. macedonicus*. These results indicate that the alleles identified in the HR stocks at both SNPs are derived alleles of recent origin.

We genotyped both SNPs in additional HR samples that had been phenotyped for presence or absence of the mini-muscle. Table 2 summarizes the results including the concordance between genotype and phenotype. For the *Ntn1* SNP there are eight discordant samples; all of them have the mini-muscle phenotype and are homozygous for the reference allele. This level of discordance and the fact that only HR3 samples have the alternative allele, rules out this SNP as a candidate for *Minimsc*. In contrast, all but one sample of 107 genotyped at the *Myh4* SNP are concordant with the mini-muscle phenotype (Table 2), including samples from both HR lines that express the mini-muscle phenotype (HR3 and HR6) and historical samples dating from August 1999 (samples beginning with the number 25). The single discordant sample (25182) is heterozygous at the *Myh4* SNP and reportedly has the mini-muscle phenotype. We hypothesize that there is an error in identification, recording, or labeling involving the animal, tissue sample, or extracted DNA from this discordant individual. This is quite plausible given that this sample originated from the historical samples dating back to generation 22 of selection in the year 1999. We therefore conclude that the *Myh4* SNP is *Minimsc* (allele symbol, $Myh4^{Minimsc}$) and that the *Ntn1* variant represents a more recent SNP that arose and was fixed in the HR3 selection line.

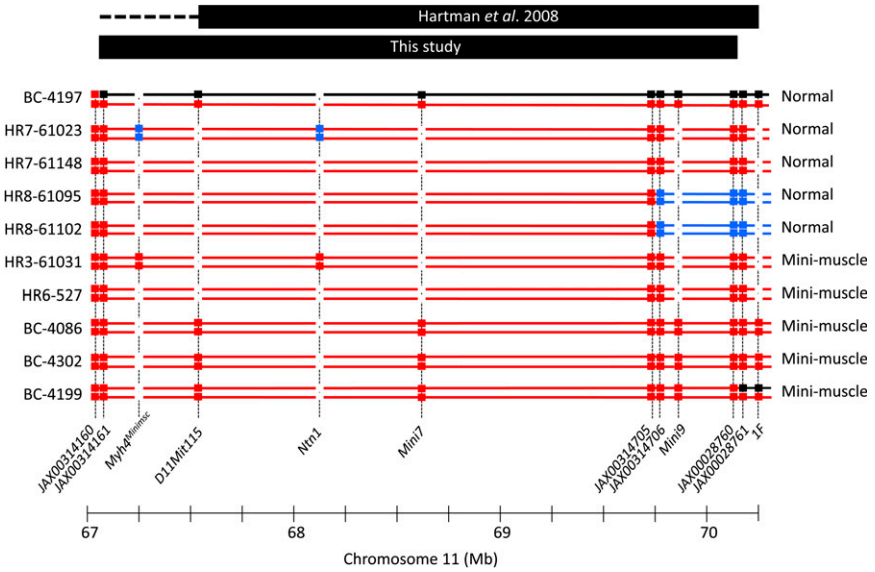


Figure 2 High-density genotyping refines the candidate interval for *Minimsc*. The different haplotypes present in each chromosome of each sample genotyped are shown as lines of different colors. Squares represent actual genotypes (MDA, microsatellites, and PCR sequencing). Red represents genotypes and haplotypes present in HR lines. Blue denotes alternative genotypes and haplotypes present in a subset of HR lines. Black represents genotypes and haplotypes present in C57BL/6J inbred line.

Discussion

Transcriptional profiling of muscle shows a significant reduction in expression of Myh4 in Myh4^{Minimsc} relative to wild type

Previously, genome-wide expression profiling described transcriptome differences in the medial gastrocnemius between HR mice (generation 37) with the mini-muscle phenotype (line 3) and HR mice with wild-type muscles (lines 7 and 8) (Burniston *et al.* 2013). Differential expression analysis revealed that, consistent with the known reduction of type IIB fibers (Guderley *et al.* 2008), *Myh4* expression was expressed at significantly lower levels (−8.82-fold reduction) in muscles taken from adult mice with the mini-muscle phenotype (Burniston *et al.* 2013). This change was accompanied by additional expression differences in structural genes [*e.g.*, myocyte enhancer factor 2C (*Mef2c*), 2-fold greater in mini-muscle samples] and myogenic factors [*e.g.*, myogenin (*Myog*), 3.8-fold greater in mini-muscle] associated with slow-type myofibrils. These findings are an important functional validation of the results presented here and further support that *Myh4^{Minimsc}* is responsible for the mini-muscle phenotype.

Phenotypic consequences of the mini-muscle mutation are strikingly similar in many ways to those observed in a strain of mice lacking IIB myosin heavy chain due to targeted disruption of the *Myh4* gene, which was first reported by Acakpo-Satchivi *et al.* (1997) and later characterized by Allen *et al.* (2001). In general, this includes changes in muscle mass, reduced body mass, larger hearts, no overt differences in locomotor behavior, reduced grip strength (K. M. Middleton, R. M. Hannon, and T. Garland Jr., unpublished results), reduced muscle force (Syme *et al.* 2005; McGillivray *et al.* 2009), and ultrastructural and histological differences (Guderley *et al.* 2006, 2008). In contrast, several findings in the *Myh4* knockout strain do not closely match data from mini-muscle mice. For example, mini-muscle mice have longer tibiafibulae (Kelly *et al.* 2006), eat more food per gram of body mass (Meek *et al.* 2010), have disorganized muscle fibers in certain muscle regions (Guderley *et al.* 2006, 2008), and have reduced contractile speed (Syme *et al.* 2005; McGillivray *et al.* 2009), whereas none of those alterations have been reported in the *Myh4* knockout mice. Some of the differences between the mice may be attributable to context (specific muscle studied, age, sex, rearing environment), while others may be due to the different types of mutations (null in the knockout and likely hypomorphic in the mini-muscle mice. In addition, they could also be the consequence of genetic background, which can exert strong influences on expressions of genes of major effect. As examples, the mini-muscle mutation shows statistically different effects on body mass, muscle mass, relative composition of myosin heavy and light chains, and mitochondrial volume densities *per* plantaris muscle fiber in the two HR lines that exhibit the phenotype (Guderley *et al.* 2006).

With regard to voluntary wheel running, Harrison *et al.* (2002) reported that myosin heavy chain null mice (IIb and IId/x) had reduced running performance [both distance (km/night) and average speed (m/min)] compared to non-transgenic C57BL/6. Although the myosin heavy chain null mice ran slower and for shorter distances compared to non-transgenics, muscles of both null groups showed an ability to train (skeletal muscle plasticity) or proximately adapt to exercise (Harrison *et al.* 2002). In response to voluntary wheel running, both null groups showed increases in muscle fiber size and muscle oxidative capacity (Harrison *et al.* 2002). In comparison, mini-muscle individuals also show an ability to train in response to voluntary wheel running, as demonstrated by Houle-Leroy *et al.* (2000, 2003). However, it is important to note that, unlike the myosin heavy chain null mice discussed above, mini-muscle individuals do not consistently run less on wheels (revolutions/day) than other HR mice (Garland *et al.* 2002).

Potential biomedical significance of Myh4^{Minimsc}

Variable muscle function and physiological properties underlying muscular performance are in part attributable to different isoforms of the molecular myosin heavy chain. Heritable mutations in the various isoforms have been associated with muscular syndromes and diseases. The pathology and mechanisms of these myosin myopathies have been detailed elsewhere (Tajsharghi and Oldfors 2013). Phenotypically, myosin myopathies broadly display disintegration of the myofibrillar network and accumulation of cellular protein aggregates. Collectively, when grouped with myofibrillar myopathies (also displaying protein aggregates), these myopathies are termed "protein aggregate myopathies" (Goebel and Muller 2006). The underlying genetic architecture and the resulting molecular pathway leading to phenotypic consequences is still undetermined for many of these myopathies, although protein conformational changes as a result of mutations seem a likely molecular cause (Ferrer and Olive 2008; Schroder and Schoser 2009).

Myofibrillar myopathy has been observed in a mouse mutant model (*ariel*) as a result of an ENU-induced recessive mutation resulting in a change (L342Q) within the motor domain of *Myh4* (MyHC IIb) (Kurapati *et al.* 2012). By day 13 (concomitant with the switch from developmental to adult myosin isoforms), homozygotes (*Myh4^{arl}/Myh4^{arl}*) develop limb paralysis caused by the aggregate-prone protein (Kurapati *et al.* 2012). No other tissue displayed pathological changes (Kurapati *et al.* 2012). The phenotypes (*e.g.*, muscle atrophy, rigid gait, muscle paralysis) in this *ariel* mutant are in sharp contrast to the *Myh4* null mice discussed above and the *Myh4^{Minimsc}* mice discussed here. The differences between *Myh4* null and *Myh4^{arl}/Myh4^{arl}* mice, as hypothesized by Kurapati *et al.* (2012), may be in part attributable to compensatory mechanisms (muscle fiber hypertrophy, upregulation of the adult myosin heavy chain IIa and IId) chronicled in *Myh4* null mice (Allen *et al.* 2001). Muscle fiber hypertrophy and upregulation of the adult myosin heavy chain IIa have previously been chronicled in *Myh4^{Minimsc}* individuals (Guderley *et al.* 2008). We also hypothesize that phenotypic differences between *Myh4^{Minimsc}* and *Myh4^{arl}/Myh4^{arl}* mice are caused by similar compensatory mechanisms as observed in *Myh4* null mice. Regardless, all three models may serve to elucidate the downstream molecular mechanisms leading to drastically different phenotypic outcomes as a result of variable changes in *Myh4*. In an attempt to determine if the *Myh4^{Minimsc}* variant is unique to the current population, or potentially appears in other species, we utilized a standard nucleotide BLAST (BLASTn) to query 115 bp before and after *Myh4^{Minimsc}*. BLASTn returned zero matching sequences inclusive of the *Myh4^{Minimsc}* variant from nonmurine species.

We agree with Kurapati *et al.* (2012) that establishing genotype–phenotype correlations for myosin mutations is a challenging exercise, but it has now been accomplished in at least three different mouse models for mutated *Myh4*. Although *Myh4* (MyHC IIb) is an essential motor protein in rats and mice, in humans the messenger RNA has been found only in masseter muscles (*e.g.*, Horton *et al.* 2001). However, functional equivalents have been identified in humans (*Myh1*), and a patient with a mutation in *Myh7* had a phenotypically similar myopathy to the *Myh4^{arl}/Myh4^{arl}* mouse (Tasca *et al.* 2012). Moreover, Harrison *et al.* (2011) observed that MyHC IIb is expressed in fetal human cells *in vitro* and may be "reactivated" in degenerative/regenerative adult muscles, thus strengthening the potential for translational relevance of the work presented here.

Evolutionary significance of Myh4^{Minimsc}

As noted previously (Hannon *et al.* 2008), the *Myh4^{Minimsc}* allele exhibits classical properties of a gene of major effect (GOME), including its Mendelian recessive nature, dramatic effect on phenotype (muscle mass), and numerous pleiotropic effects. Although the modern synthesis generally emphasized the importance of alleles with relatively small effects for phenotypic evolution in the wild, the potential importance of GOMEs has long been appreciated in the world of artificial selection and domestication, and the potential importance of GOMEs in the wild has also been emphasized in recent years (Stinchcombe *et al.* 2009). Theoretical studies have shown that GOMEs can alter genetic architecture dramatically for brief periods of time and hence change the rate and direction of multivariate evolution in important ways (Agrawal *et al.* 2001). Various empirical studies have also suggested the importance of GOMEs for evolution in the wild (*e.g.*, Voss and Shaffer 1997; Voss and Smith 2005).

The mini-muscle phenotype, the underlying *Myh4^{Minimsc}* mutation discussed here, and its potential evolutionary significance are fairly unique in the context of experimental evolution (Garland and Rose 2009). GOMEs as a product of experimental evolution are typically detected after fixation, inhibiting the examination of their effect on selection, variation, or covariation of traits (Lande 1979). In the current model, we have the ability to "watch" this mutation in both a historical and a current context, as it has disappeared in one control line, gone to fixation in one selection line, and remains polymorphic in an additional selection line. Although other examples exist (see Stinchcombe *et al.* 2009), we feel the discovery and subsequent identification of this mutation will provide an opportunity to more thoroughly examine the role of GOMEs in the response to selection. Another benefit of identifying a GOME in the context of this particular selection experiment is that we have a plethora of data on traits other than the selected phenotype (voluntary wheel running) (see references in Garland *et al.* 2011b; Burniston *et al.* 2013). This will enable us to better understand how the mutation has affected and might affect variation in related suborganismal traits related to voluntary wheel running. Significant interactions have previously been identified between *Myh4^{Minimsc}* and loci underlying voluntary wheel running and tissue mass traits (Nehrenberg *et al.*

2010). Future work will build upon these results and continue to determine how *Myh4*^{Minimsc} interacts with additional minor-effect loci to determine phenotypes observed in individuals with the mini-muscle phenotype (*e.g.*, Papa *et al.* 2013). Additionally, it will be a priority to determine what genetic mechanism is facilitating these interactions (pleiotropy, physical linkage/linkage disequilibrium, or epistasis).

Methodological significance of narrowing a broad locus to two candidate SNPs

Here we combined the benefits of experimental evolution with the tools of modern genomics to discover an adaptive phenotype inherited in a Mendelian recessive nature; to map the underlying genetic architecture to a 3.1-Mb interval on MMU11, a region containing ~100 expressed or predicted genes; and to identify the causal SNP responsible for the mini-muscle phenotype.

Typically, identification of the causative variant responsible for a given phenotype is limited by the presence of extensive linkage disequilibrium in both human and animal models and by our incomplete knowledge of how genetic variants affect the function of nearby genes. However, it is possible to identify the causative variant if these mutations are recent and within within a known haplotype. Here we combined standard linkage mapping with the identification and deep sequencing of two samples with discordant phenotypes with identical haplotypes in the candidate region to identify two candidate SNPs in a 3.1-Mb interval. Finally, we were able to exclude these two variants by performing a simple association analysis in a relatively small set of additional samples. Using next-generation sequencing, this approach can be extended to identify causative mutations arising in mouse inbred lines and thus offers a great avenue to overcome one of the most challenging steps in quantitative genetics. Note that exome sequencing, a standard approach to identifying GOMEs (*e.g.*, Ng *et al.* 2010), would have failed to identify the causative SNP.

Despite the success of our approach, the mechanism by which an intronic SNP affects the function of the *Myh4* gene remains unclear. The location of the SNP, the recessive nature of the allele, and the expression data strongly suggest a regulatory variant. However, future investigations will seek to determine if there is an association between the variant and lower levels of protein expression, which would lend further support to the regulatory variant hypothesis. Additionally, validation will require generation and analysis of mice with specific genotype combinations at *Myh4* during embryonic and early postnatal development.

Acknowledgments

We thank T. Meek and V. Careau for assistance in obtaining muscle samples and the DNA sequencing production team led by Lucinda Fulton and Robert Fulton at the Genome Institute at Washington University for sample management and data production. This work was partially supported by National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) grant DK076050 (to D.P.). Some phenotypes were collected using the Animal Metabolism Phenotyping core facility within the University of North Carolina's Nutrition Obesity Research Center (funded by NIDDK grant DK056350). Additional support was provided by National Science Foundation grants IOS-1121273 and IOB-0543429 (to T.G.). Support for the development of genomic tools and pipelines for array and whole-genome sequence analysis was provided by National Institute of Mental Health/National Human Genome Research Institute Center of Excellence for Genome Sciences grants P50MH090338 and P50HG006582 (to F.P.-M.d.V.).

Literature Cited

Acakpo-Satchivi, L. J. R., W. Edelmann, C. Sartorius, B. D. Lu, P. A. Wahr *et al.*, 1997 Growth and muscle defects in mice lacking adult myosin heavy chain genes. J. Cell Biol. 139: 1219–1229.

Agrawal, A. F., E. D. Brodie, and L. H. Rieseberg, 2001 Possible consequences of genes of major effect: transient changes in the g-matrix. Genetica 112–113: 33–43.

Allen, D. L., B. C. Harrison, C. Sartorius, W. C. Byrnes, and L. A. Leinwand, 2001 Mutation of the IIB myosin heavy chain gene results in muscle fiber loss and compensatory hypertrophy. Am. J. Physiol. Cell Physiol. 280: C637–C645.

Aylor, D. L., W. Valdar, W. Foulds-Mathes, R. J. Buus, R. A. Verdugo *et al.*, 2011 Genetic analysis of complex traits in the emerging Collaborative Cross. Genome Res. 21: 1213–1222.

Burniston, J. G., T. H. Meek, S. N. Pandey, G. Broitman-Maduro, M. F. Maduro *et al.*, 2013 Gene expression profiling of gastrocnemius of 'Mini-Muscle' mice. Physiol. Genomics 45: 228–236.

Careau, V. C., O. R. P. Bininda-Emonds, G. Ordonez, and T. Garland, Jr., 2012 Are voluntary wheel running and open-field behavior correlated in mice? Different answers from comparative and artificial selection approaches. Behav. Genet. 42: 830–844.

Careau, V., M. E. Wolak, P. A. Carter, and T. Garland, Jr., 2013 Limits to behavioral evolution: the quantitative genetics of a complex trait under directional selection. Evolution DOI: 10.1111/evo.12200.

Carriére, Y., J. P. Deland, D. A. Roff, and C. Vincent, 1994 Life history costs associated with the evolution of insecticide resistance. Proc. Biol. Sci. 258: 35–40.

Didion, J. P., H. Yang, K. Sheppard, C. P. Fu, L. McMillan *et al.*, 2012 Discovery of novel variant in genotyping arrays improves genotype retention and reduces ascertainment bias. BMC Genomics 13: 34.

Ferrer, I., and M. Olive, 2008 Molecular pathology of myofibrillar myopathies. Expert Rev. Mol. Med. 10: e25.

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

Garland, T. Jr., M. T. Morgan, J. G. Swallow, J. S. Rhodes, I. Girard *et al.*, 2002 Evolution of a small-muscle polymorphism in lines of house mice selected for high activity levels. Evolution 56: 1267–1275.

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

Garland, T. Jr., H. Schutz, M. A. Chappell, B. K. Keeney, T. H. Meek *et al.*, 2011b The biological control of voluntary exercise,

spontaneous physical activity and daily energy expenditure in relation to obesity: human and rodent perspectives. J. Exp. Biol. 214: 206–229.

Girard, I., J. G. Swallow, P. A. Carter, P. Koteja, J. S. Rhodes *et al.*, 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.

Goebel, H. H., and H. D. Muller, 2006 Protein aggregate myopathies. Semin. Pediatr. Neurol. 13: 96–103.

Guderley, H., P. Houle-Leroy, G. M. Diffee, D. M. Camp, and T. Garland, Jr., 2006 Morphometry, ultrastructure, myosin isoforms, and metabolic capacities of the "mighty mini muscles" favoured by selection for high activity in house mice. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 144: 271–282.

Guderley, H., D. R. Joanisse, S. Mokas, G. M. Bilodeau, and T. Garland, Jr., 2008 Altered fiber types in gastrocnemius muscle of high wheel-running selected mice with mini muscle phenotypes. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 149: 490–500.

Hannon, R. M., S. A. Kelly, K. M. Middleton, E. M. Kolb, D. Pomp *et al.*, 2008 Phenotypic effects of the "mini-muscle" allele in a large HR × C57BL/6J mouse backcross. J. Hered. 99: 349–354.

Harrison, B. C., M. L. Bell, D. l. Allen, W. C. Byrnes, and L. A. Leinwand, 2002 Skeletal muscle adaptations in response to voluntary wheel running in myosin heavy chain null mice. J. Appl. Physiol. 92: 313–322.

Harrison, B. C., D. L. Allen, and L. A. Leinwand, 2011 IIB or not IIB? Regulation of myosin heavy chain gene expression in mice and men. Skelet. Muscle 1: 5.

Hartmann, J., T. Garland, Jr., R. M. Hannon, S. A. Kelly, G. Muñoz *et al.*, 2008 Fine mapping of "Mini-Muscle," a recessive mutation causing reduced hind-limb muscle mass in mice. J. Hered. 99: 679–687.

Horton, M. J., C. A. Brandon, T. J. Morris, T. W. Braun, K. M. Yaw *et al.*, 2001 Abundant expression of myosin heavy-chain IIB RNA in a subset of human masseter muscle fibres. Arch. Oral Biol. 46: 1039–1050.

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

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

Keane, T. M., L. Goodstadt, P. Danecek, M. A. White, K. Wong *et al.*, 2011 Mouse genomic variation and its effect on phenotypes and gene regulation. Nature 477: 289–294.

Kelly, S. A., P. P. Czech, J. T. Wight, K. M. Blank, and T. Garland, Jr.., 2006 Experimental evolution and phenotypic plasticity of hindlimb bones in high-activity house mice. J. Morphol. 267: 360–374.

Kolb, E. M., S. A. Kelly, K. M. Middleton, L. S. Sermsakdi, M. A. Chappell *et al.*, 2010 Erythropoietin elevates VO2,max but not voluntary wheel running in mice. J. Exp. Biol. 213: 510–519.

Kurapati, R., C. McKenna, J. Lindqvist, D. Williams, M. Simon *et al.*, 2012 Myofibrillar myopathy caused by a mutation in the motor domain of mouse MyHC IIb. Hum. Mol. Genet. 21: 1706–1724.

Lande, R., 1979 Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. Evolution 33: 402–416.

McGillivray, D. G., T. Garland, Jr., E. M. Dlugosz, M. A. Chappell, and D. A. Syme, 2009 Changes in efficiency and myosin expression in the small-muscle phenotype of mice selectively bred for high voluntary running activity. J. Exp. Biol. 212: 977–985.

Meek, T. H., J. C. Eisenmann, and T. Garland, Jr., 2010 Western diet increases wheel running in mice selectively bred for high voluntary wheel running. Int. J. Obes. 34: 960–969.

Middleton, K. M., S. A. Kelly, and T. Garland, Jr., 2008 Selective breeding as a tool to probe skeletal response to high voluntary locomotor activity in mice. Integr. Comp. Biol. 48: 394–410.

Nehrenberg, D. L., S. Wang, R. M. Hannon, T. Garland, Jr., and D. Pomp, 2010 QTL underlying voluntary exercise in mice: interactions with the "mini muscle" locus and sex. J. Hered. 101: 42–53.

Ng, S. B., K. J. Buckingham, C. Lee, A. W. Bigham, H. K. Tabor *et al.*, 2010 Exome sequencing identifies the cause of a Mendelian disorder. Nat. Genet. 42: 30–35.

Papa, R., D. D. Kapan, B. A. Counterman, K. Maldonado, D. P. Lindstrom *et al.*, 2013 Multi-alleic major effect genes interact with minor effect QTLs to control adaptive color pattern variation in *Heliconius erato*. PLoS ONE 8: e57033.

Rivero, A., A. Magaud, A. Nicot, and J. Veziliér, 2011 Energetic cost of insecticide resistance in *Culex pipiens* mosquitoes. J. Med. Entomol. 48: 694–700.

Schroder, R., and B. Schoser, 2009 Myofibrillar myopathies: a clinical and myopathological guide. Brain Pathol. 19: 483–492.

Stinchcombe, J. R., C. Weinig, K. D. Heath, M. T. Brock, and J. Schmitt, 2009 Polymorphic genes of major effect: consequences for variation, selection and evolution in *Arabidopsis thaliana*. Genetics 182: 911–922.

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

Swallow, J. G., J. S. Rhodes, and T. Garland, Jr., 2005 Phenotypic and evolutionary plasticity of organ masses in response to voluntary exercise in house mice. Integr. Comp. Biol. 45: 426–437.

Swallow, J. G., J. P. Hayes, P. Koteja, and T. Garland, Jr., 2009 Selection experiments and experimental evolution of performance and physiology, pp. 301–333 in *Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments*, edited by T. Garland Jr., and M. R. Rose. University of California Press, Berkeley, CA.

Syme, D. A., K. Evashuk, B. Grintuch, E. L. Rezende, and T. Garland, Jr., 2005 Contractile abilities of normal and "mini" triceps surae muscles from mice (*Mus domesticus*) selectively breed for high voluntary wheel running. J. Appl. Physiol. 99: 1308–1316.

Tajsharghi, H., and A. Oldfors, 2013 Myosinopathies: pathology and mechanisms. Acta Neuropathol. 125: 3–18.

Tasca, G., E. Ricci, S. Penttila, M. Monforte, V. Giglio *et al.*, 2012 New phenotype and pathology features in MYH7-related distal myopathy. Neuromuscul. Disord. 22: 640–647.

Voss, S. R., and H. B. Shaffer, 1997 Adaptive evolution via a major gene effect: paedomorphosis in the Mexican axolotl. Proc. Natl. Acad. Sci. USA 94: 14185–14189.

Voss, S. R., and J. J. Smith, 2005 Evolution of salamander life cycles: a major-effect quantitative trait locus contributes to discrete and continuous variation for metamorphic timing. Genetics 170: 245–281.

Yang, H., Y. Ding, L. N. Hutchins, J. Szatkiewicz, T. A. Bell *et al.*, 2009 A customized and versatile high-density genotyping array for the mouse. Nat. Methods 6: 663–666.

Yang, H., J. R. Wang, J. P. Didion, R. J. Buus, T. A. Bell *et al.*, 2011 Subspecific origin and haplotype diversity in the laboratory mouse. Nat. Genet. 43: 648–655.

Communicating editor: D. W. Threadgill

GENETICS

Supporting Information
http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.113.154476/-/DC1

A Novel Intronic Single Nucleotide Polymorphism in the *Myosin heavy polypeptide 4* Gene Is Responsible for the Mini-Muscle Phenotype Characterized by Major Reduction in Hind-Limb Muscle Mass in Mice

Scott A. Kelly, Timothy A. Bell, Sara R. Selitsky, Ryan J. Buus, Kunjie Hua, George M. Weinstock, Theodore Garland Jr., Fernando Pardo-Manuel de Villena, and Daniel Pomp

DOI: 10.1534/genetics.113.154476

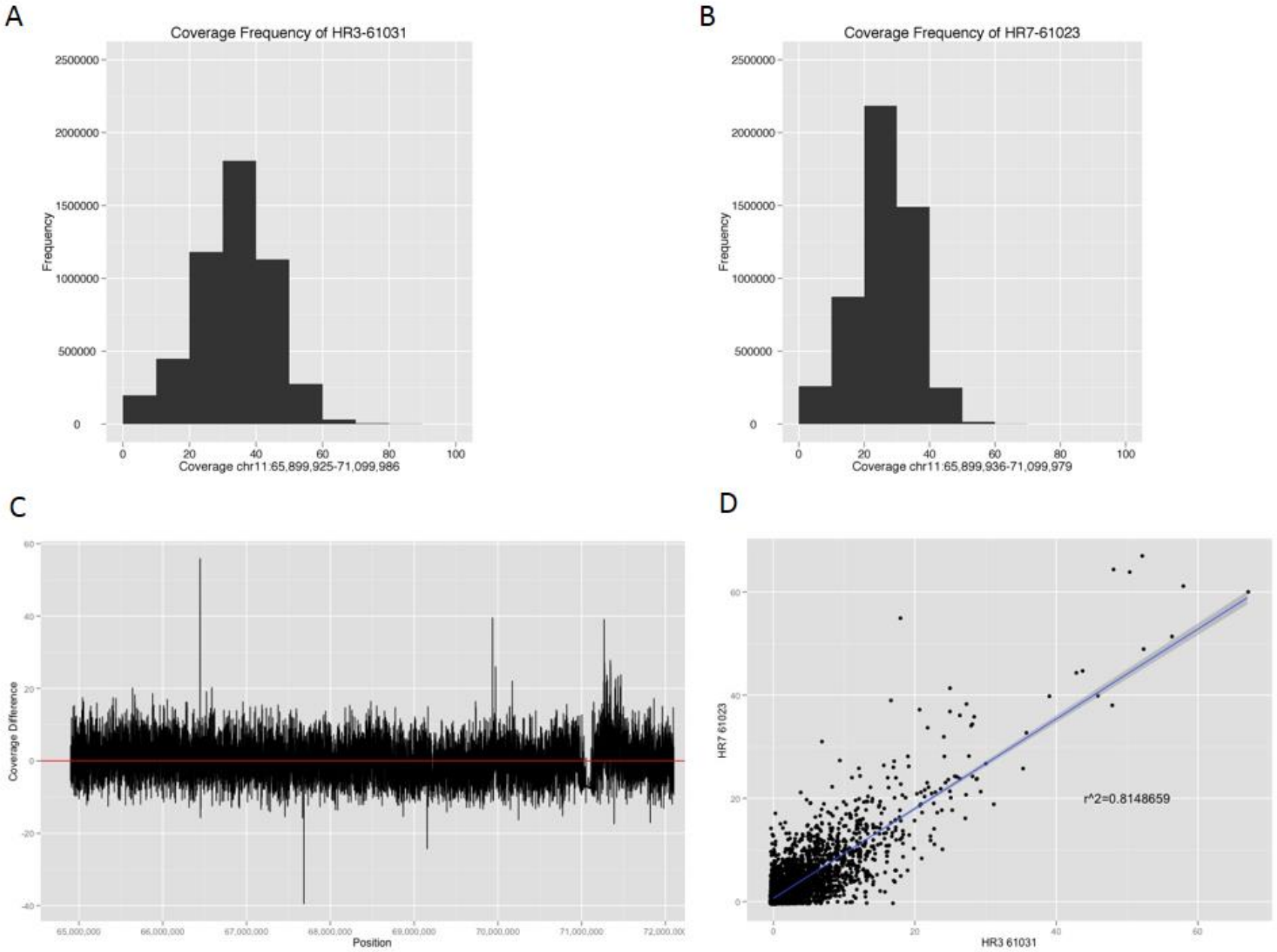


Figure S1 Coverage and sequence consistency in the two HR samples for which whole-genome sequence was generated in this study. (A) and (B) show the frequency distribution in pileup high in the updated candidate interval ± 1 Mb proximal and distal. (C) shows the coverage difference between the two samples along the entire interval calculated as: (HR3-HR7) - average (HR3-HR7). (D) shows the correlation between the number of consistent calls in the two samples in 200 bp windows across the candidate interval ± 1 Mb proximal and distal.

Files S1-S2

Available for download at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.113.154476/-/DC1

File S1

Whole-genome sequence of HR3-61031 utilized to identify candidate mutations for *Minimsc*

File S2

Whole-genome sequence of HR7-61023 utilized to identify candidate mutations for *Minimsc*

File S3

SAMtools pileup format

Example:

chr11	66823583	t	17	.c...,.,,..,.,..,	H-H9HH5CAHEEHBFH9
chr11	66823584	t	17	.,...,.,,..,.,..,	H(HEH@A;<FE?F7GH0
chr11	66823585	c	17	.,...,.,,..,.,..,	F9HCHH@DAHBEF@CHE
chr11	66823586	c	17	.,...,.,,..,.,..,	CFHEHFBIEHDGF?GHE
chr11	66823587	t	17	.,...,.,,..,.,..,	B=FEHF@B<HEBE@BH9
chr11	66823588	c	17	.,...,.,,..,.,..,	CG@>DH<FDHDGFFFHE
chr11	66823589	c	17	.,...,.,,..,.,..,	GEH8HHBECH@GHEGHB
chr11	66823590	c	17	.,...,.,,..,.,..,	?FG2CH3EAG@E@FEH=
chr11	66823591	c	19	.,...,.,,..,.,..,^~.^~.	GDD0DH<ECH@GFCEHB9=

The lines of the pileup consist of chromosome, position, reference base, coverage, read bases, and base quality. The read bases are which bases align to that specific position. A period represents a match of the reference base on the forward strand and a comma represents a match on the reverse. The "^" symbol represents the start of a new read and the ASCII of the character minus 33 following represents mapping quality. The "$" represents the end of a read. An "A", "G", "C", or "T" represents a mismatch to the reference on the forward strand and "a", "g", "c", or "t" represents a mismatch to the reference on the reverse strand. The last column represents the mapping Phred quality, which is the ASCII value minus 33.

Helpful links:

SAMtools pileup format explanation:

http://samtools.sourceforge.net/pileup.shtml

http://en.wikipedia.org/wiki/Pileup_format

ASCII Wikipedia page:

http://en.wikipedia.org/wiki/ASCII

Table S1 Genotypes at the two candidate SNPs for the mini-muscle phenotype. Samples were genotyped for *Myosin, heavy polypeptide 4, skeletal muscle* gene (*Myh4*) and *Netrin 1* (*Ntn1*). The origin reflects the taxonomical origin of the haplotype present in the *Minimsc* candidate region as described in Yang *et al.* 2011. Most laboratory inbred strains genotyped have not been directly phenotyped.

Name	Type	Origin	*Ntn1*	*Myh4*	Expected Phenotype	Actual Phenotype[a]
129P2	Classical	*M. m. domesticus*	T	C	normal	nd
129S1/SvImJ	Classical	*M. m. domesticus*	T	C	normal	nd
129S5	Classical	*M. m. domesticus*	T	C	normal	nd
A/J	Classical	*M. m. domesticus*	T	C	normal	nd
AKR/J	Classical	*M. m. domesticus*	T	C	normal	nd
BALB/cJ	Classical	*M. m. domesticus*	T	C	normal	nd
BULS	Wild-derived	*M. m. musculus*	T	C	normal	nd
BUSNA	Wild-derived	*M. m. musculus*	T	C	normal	nd
C3H/HeJ	Classical	*M. m. domesticus*	T	C	normal	nd
C57BL/6J	Classical	*M. m. domesticus*	T	C	normal	normal
CALB/RkJ	Wild-derived	*M. m. domesticus*	T	C	normal	nd
CAST/EiJ	Wild-derived	*M. m. castaneus*	T	C	normal	nd
CBA/J	Classical	*M. m. domesticus*	T	C	normal	nd
CZECHI/EiJ	Wild-derived	*M. m. musculus*	T	C	normal	nd
CZECHII/EiJ	Wild-derived	*M. m. musculus*	T	C	normal	nd
DBA/2J	Classical	Switches[b]	T	C	normal	nd
DDY/JclSidSeyFrkJ	Classical	*M. m. domesticus*	T	C	normal	nd
FVB/NJ	Classical	*M. m. domesticus*	T	C	normal	nd
JF1/Ms	Wild-derived	*M. m. musculus*	T	C	normal	nd
KK/HlJ	Classical	*M. m. domesticus*	T	C	normal	nd
LG/J	Classical	*M. m. domesticus*	T	C	normal	nd
LP/J	Classical	*M. m. domesticus*	T	C	normal	nd
LEWES/EiJ	Wild-derived	*M. m. domesticus*	T	C	normal	nd
MOLD/RkJ	Wild-derived	*M. m. musculus*	T	C	normal	nd
MOLF/EiJ	Wild-derived	*M. m. musculus*	T	C	normal	nd
NOD/ShiLtJ	Classical	*M. m. domesticus*	T	C	normal	nd

NON/ShiLtJ	Classical	Switches[b]	T	C	normal	nd
NZW/LacJ	Classical	*M. m. domesticus*	T	C	normal	nd
NZO/HILtJ	Classical	*M. m. domesticus*	T	C	normal	nd
PERC/EiJ	Wild-derived	*M. m. domesticus*	T	C	normal	nd
POHN/Deh	Wild-derived	*M. m. castaneus*	T	C	normal	nd
PWD/PhJ	Wild-derived	*M. m. musculus*	T	C	normal	nd
PWK/PhJ	Wild-derived	*M. m. musculus*	T	C	normal	nd
RBF/DnJ	Wild-derived	*M. m. domesticus*	T	C	normal	nd
RIIIS/J	Classical	*M. m. domesticus*	T	C	normal	nd
SKIVE/EiJ	Wild-derived	*M. m. musculus*	T	C	normal	nd
SEB	Wild-derived	*M. spretus*	T	C	normal	nd
SEG	Wild-derived	*M. spretus*	T	C	normal	nd
SFM	Wild-derived	*M. spretus*	T	C	normal	nd
SPRET/EiJ	Wild-derived	*M. spretus*	T	C	normal	nd
SMON	Wild-derived	*M. spretus*	T	C	normal	nd
STLT	Wild-derived	*M. m. domesticus*	T	C	normal	nd
STUF	Wild-derived	*M. m. musculus*	T	C	normal	nd
WMP	Wild-derived	*M. m. domesticus*	T	C	normal	nd
WSB/EiJ	Wild-derived	*M. m. domesticus*	T	C	normal	nd
XBS	Wild-derived	*M. macedonicus*	T	C	normal	nd

[a]nd, sample not phenotyped.

[b]The phylogenetic origin of the haplotype switches between different *Mus musculus* subspecies in this region.