Fine-mapping of muscle weight QTL in LG/J and SM/J intercrosses

A. Lionikas, R. Cheng, J. E. Lim, A. A. Palmer, and D. A. Blizard

1School of Medical Sciences, College of Life Sciences and Medicine, University of Aberdeen, Foresterhill, Aberdeen, United Kingdom; 2Center for Developmental and Health Genetics, Pennsylvania State University, University Park, Pennsylvania; Departments of 3Human Genetics and 4Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, Illinois; and 5Department of Pharmacology and Cancer Biology, Duke University Medical Center, Durham, North Carolina

Submitted 17 May 2010; accepted in final form 9 July 2010

Lionikas A, Cheng R, Lim JE, Palmer AA, Blizard DA. Fine-mapping of muscle weight QTL in LG/J and SM/J intercrosses. Physiol Genomics 42A: 33–38, 2010. First published July 13, 2010; doi:10.1152/physiolgenomics.00100.2010.—Genetic variation plays a substantial role in variation in strength, but the underlying mechanisms remain poorly understood. The objective of the present study was to examine the mechanisms underlying variation in muscle mass, a predictor of strength, between LG/J and SM/J strains, which are the inbred progeny of mice selected, respectively, for high and low body weight. We measured weight of five hindlimbs muscles in LG/J and SM/J males and females, in F1 and F2 intercrosses, and in an advanced intercross (AI), F34, between the two. F2 mice were genotyped using 162 SNPs throughout the genome; F34 mice were genotyped at 3,015 SNPs. A twofold difference in muscle mass between the LG/J and SM/J mouse strains was observed. Integrated genome-wide association analysis in the combined population of F2 and AI identified 22 quantitative trait loci (QTL); genome-wide P < 0.05) affecting muscle weight on Chr 2 (2 QTL), 4, 5, 6 (7 QTL), 7 (4 QTL), 8 (4 QTL), and 11 (3 QTL). The LG/J allele conferred greater muscle weight in all cases. The 1.5-LOD QTL support intervals ranged between 0.3 and 13.4 Mb (median 3.7 Mb) restricting the list of candidates to between 5 and 97 genes. Selection for body weight segregated the alleles affecting skeletal muscle, the most abundant tissue in the body. Combination of analyses in an F2 and AI was an effective strategy to detect and refine the QTL in a genome-wide manner. The achieved resolution facilitates further elucidation of the underlying genetic mechanisms affecting muscle mass.

skeletal muscle size; quantitative trait loci; fine mapping

SKELETAL MUSCLE IS THE MOST abundant tissue in the body and is critical for diverse functions, including locomotion, glucose homeostasis, thermoregulation, and protection of bones and viscera. Muscle strength is an important factor of health and fitness as it inversely associated with the risk for coronary heart disease and stroke (33). Because of the world’s growing geriatric population, age-related loss of strength and muscle mass (sarcopenia) is an increasingly pressing problem. Understanding the mechanisms underlying skeletal muscle mass and its function is therefore of increasing importance.

Genetic variation accounts for around half of the variation in strength in humans (1, 5, 11, 30, 32, 34, 37), presumably reflecting various aspects of muscle structure and function. Discovery of the relevant gene variants affecting muscle mass may identify the biological pathways affecting skeletal muscle development and atrophy thereby facilitating identification of novel pharmacological targets. Nevertheless, very few specific genes underlying genetic variation in muscle mass have been identified to date (8, 25, 41, 42).

Animal models such as mice (25, 35), sheep (6), and pigs (41) have been used to study the genetics of muscle mass. In these studies, traits that are largely monogenic have been examined, facilitating discovery of the underlying genes. However, variation in skeletal muscle phenotypes is a polygenic trait in all species and attempts to study the polygenic architecture of muscle variation have been initiated in mice (4, 20, 24), pigs (26, 43), and chicken (18, 27, 29, 40). A number of quantitative trait loci (QTL) have been mapped; however, refinement of these QTL and identification of the underlying genes is a challenging problem.

We used a classical F2 intercross and an advanced intercross (AI) population, which allows mapping of QTL to very small intervals (9), in an effort to accelerate progress in this area. By integrating studies of F2 and AI mice, we have combined the power of an F2 population and the resolution of an AI population. The inbred strains used to make this AI were divergently selected for body size (13, 22) and are thus ideally suited for this experiment. Studies in pigs (26, 43), sheep (8), and mice (19, 20) have demonstrated that genetic variation can exert muscle-specific effects. Therefore, we examined five different muscles in the hindlimb of mice in our populations.

METHODS

Animals and phenotype. The study was carried out on males and females of the LG/J and SM/J inbred strains, LGSMF1, LGSMF2 intercrosses, and a population of the 34th filial generation of AI, LGSMF34, (see Table 1 for details). This set of mice, recently described by Cheng and colleagues (7), was involved in a behavioral study prior to the analyses of muscle phenotypes. Animals were killed at the age of 94 ± 4 days and their carcasses frozen. All procedures through death were conducted at the University of Chicago and approved by the Institutional Animal Care and Use Committee.

Carcasses were defrosted and tibialis anterior (TA), extensor digitorum longus (EDL), gastrocnemius (Gastroc), soleus, and quadriceps femoris (QF) were dissected under a dissection microscope and weighed on a Mettler AE 50 balance (Mettler Toledo, OH) to the closest 0.1 mg.

Aggregate genetic effects. Broad sense heritability estimates, H2, were obtained as follows: \( H^2 = (V_T - V_E)/V_T \) (10), where phenotypic variance, \( V_T \), is variance of genetically mixed population (LGSMF2 or LGSMF34) and environmental variance, \( V_E \), is variance of LGSMF1 population.

QTL mapping. Mice were genotyped using SNP markers that were approximately evenly distributed across the genome at 162 (F2) SNPs or 3,015 (AI) single nucleotide polymorphisms (SNPs). We performed genome-wide association analysis in the combined population of the F2 and F34 intercrosses using software developed at the University of Chicago (QTLRel; http://www.palmerlab.org) with some custom modifications. This software allowed us to account for
the complex relationships (e.g., sibling, half-sibling, cousins) among the AI mice by using a mixed model as described by Cheng and colleagues (7). Because of sex differences in muscle mass and also the discovery of sex-specific QTL in other studies (19, 20) we explored genetic models where sex was included as either an additive or an interactive covariate.

Significance thresholds. Muscle-specific thresholds were derived using a gene dropping procedure, which generates genotypic data by simulating meiosis involving individuals in a pedigree, as described in Cheng et al. (7). Briefly, a genotypic data set was simulated, and a genome-wide scan was performed using the simulated genotypic data and the original phenotypic data and using the model for which a significance threshold was being sought. The maximum of the test statistic was recorded. This process was repeated 1,000 times, and the corresponding maxima of the test statistic were pooled to estimate the significance thresholds on a per-chromosome basis.

There is substantial difference in muscle size between males and females. Also, a number of loci affecting skeletal growth in a sex-specific manner were reported in this lineage (28). Therefore, to test for the sex-specific effects on muscle weight we compared the model with QTL-by-sex interaction (alternative hypothesis) and the model without the interaction (null hypothesis). The testing procedure was the same as described above, and the testing method was also gene dropping.

In instances of multiple peaks on the chromosome the following criteria were used to discriminate between linked QTL; the peak must be above the genome-wide P < 0.05 threshold and must drop by 3 LOD on each side of that peak (except for the end of the chromosome).

We defined the confidence interval for each QTL as the 1.5-LOD drop off on either side of the peak marker. This interval was expressed in physical map position (Mb) by using the nearest genotyped SNP that flanked the support interval.

Database analyses. Mouse phenome database (3) was used to identify nonsynonymous SNPs in the QTL regions. The SNP genotypes of the LG/J and SM/J strains were selected from the CGD1 imputed data set (http://phenome.jax.org/dbg?tn=projects/details&syn=CGD1) (36).

PolyPhen web-based tool (31) was utilized to predict the possible effects of amino acid substitution on the function of a protein (http://genetics.bwh.harvard.edu/pph/). These predictions are based on multiple sequence alignments and functional and structural characterization of the substitution site.

Mouse genome informatics (MGI) database (http://www.informatics.jax.org/) and gene expression omnibus (http://www.ncbi.nlm.nih.gov/geo/) were used to examine tissue expression of the genes. The MGI database was also utilized to search for transgenic models influencing skeletal muscle.

Other statistical analyses. Muscle weight phenotypes approximated normality in both F2 and F3 populations. Strain effect on muscle weight in progenitors was assessed using two-way ANOVA (strain and sex as independent factors). Pearson’s product-moment correlations were calculated between different muscle weights within males and females. First principal component (PC1) was extracted from the five muscles and used as additional phenotype in the QTL analysis.

RESULTS

Phenotypic analyses. There was a more than twofold difference in muscle weight between the LG/J and SM/J strains (P << 0.0001; Table 1), indicating that selection for body size at 60 days of age (13, 22) resulted in segregation of alleles affecting skeletal muscle weight at 90 days. Phenotypic correlations (within-sex) between weights of various muscles were comparable in the LGSMF3 and LGSMF34 populations and ranged between 0.70 and 0.90 (all P < 0.01) in males and 0.69 and 0.91 (P < 0.01) in females.

Broad sense heritability ranged between 0.60 and 0.82, was similar among the five muscles and tended to be somewhat higher in males than females (Table 2). Heritability estimates derived from the LGSMF34 variances were comparable with those based on the LGSMF2.

QTL analysis. Multiple linked QTL emerged as a characteristic feature of the genetic architecture of muscle weight in the cross of LG/J and SM/J strains. Association analysis of individual muscles and of PC1 in the integrated LGSMF2 and LGSMF34 population mapped 22 QTL (genome-wide P < 0.05) to Chromosomes 2 (2 loci), 4, 5, 6 (7 loci), 7 (4 loci), 8 (4 loci), and 11 (3 loci) (Fig. 1). Characteristics of the genetic architecture of muscle weight are summarized in Table 3. Size of the QTL effect did not differ between males and females statistically significantly. Among the QTL, two (Skmw28, Skmw37) explained >10% of phenotypic variance, and 10 loci explained 5% or more of phenotypic variation in one or more muscles and/or in PC1. The LG/J allele always conferred greater muscle weight. The QTL support interval (1.5-LOD

| Table 1. Phenotypes of the LG/J and SM/J strains and their intercross populations |
|-----------------------------------|-------|--------|----------|---------|---------|---------|---------|
|     | Sex  | Age, days | TA, mg  | EDL, mg | Gastroc, mg | Soleus, mg | QF, mg   |
| LG/J | ♂ (n = 18) | 93.9 ± 2.6 | 68.5 ± 4.3 | 12.7 ± 1.0 | 164.8 ± 9.8 | 10.1 ± 0.9 | 321.0 ± 18.6 |
|      | ♀ (n = 19) | 95.1 ± 1.7 | 56.7 ± 3.0 | 10.0 ± 0.5 | 122.3 ± 7.1 | 8.0 ± 0.8 | 241.4 ± 12.4 |
| SM/J | ♂ (n = 15) | 90.5 ± 0.9 | 31.2 ± 2.2 | 6.4 ± 0.6 | 80.8 ± 5.7 | 4.8 ± 0.5 | 150.2 ± 9.2 |
|      | ♀ (n = 10) | 90.7 ± 0.9 | 26.6 ± 2.0 | 5.1 ± 0.5 | 64.1 ± 5.8 | 3.8 ± 0.5 | 117.0 ± 11.0 |
| LGSMF1 | ♂ (n = 11) | 93.2 ± 1.2 | 58.9 ± 2.9 | 10.2 ± 0.6 | 138.1 ± 7.0 | 8.1 ± 0.5 | 255.3 ± 12.8 |
|      | ♀ (n = 5) | 91.4 ± 1.3 | 46.5 ± 3.4 | 8.1 ± 0.6 | 100.1 ± 6.0 | 5.9 ± 0.5 | 184.4 ± 13.3 |
| LGSMF3 | ♂ (n = 252) | 95.3 ± 2.5 | 49.4 ± 6.5 | 9.2 ± 1.2 | 122.8 ± 14.4 | 7.0 ± 1.1 | 227.8 ± 27.9 |
|      | ♀ (n = 245) | 95.6 ± 2.2 | 38.3 ± 5.9 | 7.0 ± 1.0 | 87.3 ± 11.2 | 5.2 ± 0.8 | 166.6 ± 22.4 |
| LGSMF34 | ♂ (n = 254) | 93.0 ± 5.4 | 51.8 ± 6.8 | 9.7 ± 1.3 | 128.2 ± 15.8 | 7.7 ± 1.3 | 233.9 ± 29.3 |
|      | ♀ (n = 238) | 94.1 ± 4.4 | 40.4 ± 5.6 | 7.3 ± 1.0 | 92.4 ± 12.8 | 5.9 ± 1.0 | 172.8 ± 23.1 |

Values are means ± SD. TA, tibialis anterior; EDL, extensor digitorum longus; Gastroc, gastrocnemius; QF, quadriceps femoris.
drop-off) of individual traits ranged between 0.3 and 13.4 Mb (median 3.7 Mb).

Consistent with the positive phenotypic correlation between different muscles, the majority of the QTL exerted pleiotropic effects on more than one muscle. Weight of the TA muscle was affected by 12 (together accounting for 54% of phenotypic variance, based on single QTL model estimates), EDL by 12 (45%), Gastroc by 7 (42%), soleus by 9 (19%), and QF by 6 (39%) QTL. The genetic architecture of the PC1 largely overlapped with that of the individual muscles. In addition, PC1 analysis permitted refinement of three additional loci (Table 3). A total of 14 QTL were identified for PC1.

Candidate genes. The mapping resolution provided by our analyses guided us to a manageable number of genes in many loci; for instance, Skmw27 and Skmw30 harbored fewer than 10 genes and three loci; Skmw23, Skmw32, and Skmw40 harbored fewer than 20. The SNP analysis of all QTL identified 30 genes with nonsynonymous polymorphisms between the two strains that were clustered within 10 QTL regions. We examined these genes using the PolyPhen tool (http://genetics.bwh.harvard.edu/pph/), which relies on the sequence and phylogenetic and structural information characterizing the substitution (31). Most of these polymorphisms were predicted to have no effect on the protein function. However, polymorphisms in the zinc finger protein 341, Zfp341, (Skmw21); lipopolysaccharide binding protein, Lbp, phospholipase γ 1, Plcγ1 (Skmw22); tubulin tyrosine ligase-like family member 3, Tll3; interleukin 17 receptor C, Il17rc; Fanconi anemia complementation group D2, Fancd2; interleukin-1 receptor-associated kinase 2, Ira2 (Skmw29); and WD repeat containing protein 17, Wdrl7 (Skmw38) genes are predicted to alter protein function. Based on the available microarray data, transcripts of all eight genes are detected in mouse skeletal muscle (GDS2840, GDS2698) providing support for further scrutiny of the candidacy of these genes.

We examined available transgenic models for possible effects on muscle tissue using the MGI database and found that allelic variations (due to spontaneous or induced mutations) in a number of genes from the QTL regions can affect the skeletal muscle (summarized in Table 3). There were no nonsynonymous polymorphisms between the two strains in these genes. However, polymorphisms causing expression variation could be the underlying mechanism of the QTL effects.

DISCUSSION

We carried out the genome-wide mapping of QTL affecting skeletal muscle mass in an F2 and AI population of mice derived from strains selected for high and low body weight at 60 days of age. We hypothesized that such selection would have segregated alleles affecting skeletal muscle weight, the most abundant tissue in the body. Indeed, the LG/J and SM/J strains exhibited a twofold difference in weight of five hindlimb muscles and heritability estimates ranged between 0.60 and 0.88 for the weight of different muscles. We mapped 22 QTL affecting weight of one or more muscles. The LG/J allele conferred an increasing effect in all QTL. The analyses in the integrated F2 and AI population permitted us to achieve a genome-wide fine mapping of the loci to median resolution of 3.7 Mb. In four muscles, the aggregate effect of the identified QTL accounted for between 38 and 59% of phenotypic variation and somewhat less, 19%, in soleus. The underlying mechanism of this contrast between soleus and other muscles is not clear but could be partially due to the fiber type composition differences. Mouse soleus consists of a mixture of type I and IIA fibers, whereas four other tested muscles predominantly consist of type IIA and IIB fibers (47, 49).

Analyses of body and organ weights as well as skeletal and body growth have proven the LG/J × SM/J lineage a suitable
We were able to refine nine loci to <3 Mb size. The strength of the classical QTL mapping in an F2 intercross lies in its ability to identify the entire genetic architecture contributing to phenotypic variation. The refinement of the QTL, however, poses a significant hurdle. Typically, the fine mapping is carried out via congenic strain approach on a locus by locus basis aiming to narrow down the region of interest to a manageable list of candidate genes (38). An advanced intercross is an appealing alternative, particularly for the phenotype with rich genetic architecture, because it allows screening of the entire genome in one study (9). However, the power to detect QTL in AI appeared more limited than in F2S. For instance, none of the QTL affecting TA weight on Chr 6, 7, or 8 in LGSMF2 intercross was replicated in a separate analysis within the LGSMF34 population (Fig. 2). We hypothesize that this is an outcome of the partitioning of QTL into several linked loci. The effect size of fractioned QTL might have been beyond detection power of the available sample size. Partitioning of the QTL has been reported previously (45) and suggested as a possible cause of replication failure in addition to the false positives and the Beavis effect (2). A QTL scan in the
integrated F2 and AI population of the same lineage offers a superior strategy of data analysis by increasing the sample size, combining the detection power of the former and accuracy of the latter. To illustrate the aspect of accuracy, we continue an example presented in Fig. 2. A 1.5-LOD drop off interval of the TA weight QTL on Chr 6 spanned between 54.1 and 104.8 Mb (50.7 Mb long) in the LGSMF2 population. The QTL broke down into six smaller loci in the combined analysis of the LGSMF2 and LGSMF34 populations (Supplemental Table S1). The size of the refined loci ranged between 0.3 and 5.4 Mb, a >10-fold reduction. Importantly, three out of six loci were outside of the LGSMF2 1.5-LOD drop off interval, and thus, candidacy of the genes within those regions would not even be considered in the analysis of solely F2 population.

Mapping resolution led us to a manageable list of the candidate genes in most of the loci. Among them we identified eight genes, Zfp341, Lhp, Plcg1, Tll3, Il17rc, Fancd2, Ira2, and Wdr17, with nonsynonymous polymorphisms that are likely to influence function of the proteins between the two strains. These genes are expressed in muscle tissue and thus are good candidates for follow-up analysis. Several other genes in the QTL regions are likely functional candidates as well. For instance, mutation in mitochondrial protease Omi coding gene, Htra2 (Skmw26), leads to muscle wasting (16), overexpression of caveolin-3, Cav3, gene results in fiber phenotypes resembling Duchenne muscle dystrophy (12), and disruption of type 1 IGF receptor gene, Igf1r (Skmw32), causes severe suppression of muscle growth (21), whereas Tead1 (Skmw34) (39), Tead4 (Skmw30) (48), and Mef2a (Skmw32) (48) genes are transcription factors abundantly expressed in skeletal muscle. The genomic sequence of the LG/J and SM/J strains and the transcriptome analysis will provide additional leverage for further nomination of candidate genes underlying those QTL.

The overall genetic architecture was richer than that of any individual muscle indicating that not all muscles were affected (or at least were not affected equally) by the same locus. Six out of 22 loci were found influencing only EDL (three loci), soleus (two loci), or TA (one locus) muscle. Specificity of the Skmw22 and Skmw42 loci to soleus might be related to the distinct fiber type composition of this muscle (a mixture of type I and type 2A fibers) compared with the other examined muscles (predominantly a mixture of type 2A and type 2B fibers) as well as its function (soleus is a frequently used postural muscle). The effects specific to EDL (Skmw23 and Skmw24 loci) but not, for instance, to its synergist TA were also reported in the intercross between the C57BL/6J and DBA/2 strains (19, 20), although the underlying mechanisms of the specificity between muscles of similar type remain unclear. Specificity of the loci affecting TA (Skmw27) and EDL (Skmw36) muscles most likely is an artifact arising from inability to discriminate between the linked loci according to the selected criteria, i.e., 3 LOD drop between the peaks, in other muscles (Fig. 1). Muscle-specific effects were observed earlier in various species (15, 19, 20, 23, 26, 43), and a gene underlying variation of such nature was recently identified in sheep (46). Identification of the genes and pathways affecting different muscles in a variable manner will help in understanding myopathies that often affect some but not all muscles (14).

In conclusion, we identified and refined the genetic architecture consisting of 22 QTL affecting variation in muscle weight in the LG/J × SM/J lineage. The resolution and effect size of the QTL are favorable for a fruitful pursuit of the underlying genes.

ACKNOWLEDGMENTS
We acknowledge the valuable technical and scientific assistance of Greta Sokoloff, Katilin Samocha, Tanya Cebollero, Michaelanne Munoz, Claudia Wing, Sima Lionikaïte, and Viktė Lionikaïtė. We also thank the laboratory of Dr. James Cheverud for making the F33 mice available to us.

GRANTS
National Institute of Health Grants AR-052879 to A. Lionikas, AR-056280 to D. A. Blizard, DA-024845 to A. A. Palmer, MH-020065 to A. A. Palmer. Schweppes Foundation Career Development Award to A. A. Palmer.

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

REFERENCES
MUSCLE WEIGHT IN LG/J AND SM/J STRAINS


