Identification of genetic determinants of IGF-1 levels and longevity among mouse inbred strains

Magalie S. Leduc, **Rachael S. Hageman**, **Qingying Meng**, **Ricardo A. Verdugo**, **Shirng-Wern Tsaih**, **Gary A. Churchill**, **Beverly Paigen**[*], and **Rong Yuan**

The Jackson laboratory, Bar Harbor, Maine, 04609

Summary

The IGF-1 signaling pathway plays an important role in regulating longevity. To identify the genetic loci and genes that regulate plasma IGF-1 levels, we intercrossed MRL/MpJ and SM/J, inbred mouse strains that differ in IGF-1 levels. Quantitative trait loci (QTL) analysis of IGF-1 levels of these F2 mice detected four QTL on chromosomes (Chrs) 9 (48 Mb), 10 (86 Mb), 15 (18 Mb) and 17 (85 Mb). Haplotype association mapping of IGF-1 levels in 28 domesticated inbred strains identified three suggestive loci in females on Chrs 2 (13 Mb), 10 (88 Mb) and 17 (28 Mb) and in males on Chrs 1 (159 Mb), 3 (52 and 58 Mb) and 16 (74 Mb). Except for the QTL on Chr 9 and 16, all loci co-localized with IGF-1 QTL previously identified in other mouse crosses. The most significant locus was the QTL on Chr 10, which contains the *Igf1* gene and which had a LOD score of 31.8. Haplotype analysis among 28 domesticated inbred strains revealed a major QTL on Chr 10 overlapping with the QTL identified in the F2 mice. This locus showed three major haplotypes; strains with haplotype 1 had significantly lower plasma IGF-1 and extended longevity (*P* < 0.05) than strains with haplotype 2 or 3. Bioinformatic analysis, combined with sequencing and expression studies, showed that *Igf1* is the most likely QTL gene, but that other genes may also play a role in this strong QTL.

Keywords

IGF-1; QTL; longevity; mouse; haplotype analysis

Introduction

Insulin-like growth factor 1 (IGF-1) is a polypeptide involved in a wide range of biological effects, including regulation of glucose metabolism, cell proliferation, differentiation and survival (Stewart & Rotwein 1996; Yakar *et al.* 2004). Previous studies demonstrated that reduced activity of the IGF-1 signaling pathway increases lifespan in nematodes, fruit flies and mice reviewed in (Kuningas *et al.* 2008). IGF-1 levels are genetically determined and highly heritable in humans (h^2 ~ 0.59) and other species (van Heemst *et al.* 2005). Polymorphisms in the *Igf1* gene itself and in the IGF-1 signaling pathway — such as IGF binding proteins 1 and 3 (*IGFBP1* and *3*), *IGF-1*, IGF-1 receptor (*IGF1R*), and phosphatidylinositol 3-kinase catalytic beta polypeptide (*PIK3CB*) — are associated with IGF-1 levels (Bonafe *et al.* 2003; Patel *et al.* 2008). Furthermore, some of the same genes

[*]Corresponding author: Beverly Paigen, The Jackson Laboratory, 600 Main Street, Bar Harbor, Maine, US 04609, Phone: 207-288-6388, Fax: 207-288-6078, bev.paigen@jax.org.

have also been associated with aging and longevity in yeast, worms, fruit flies, rodents and human (Longo & Finch 2003; Albani *et al.* 2009). Recently, we found that circulating IGF-1 levels were inversely correlated with longevity among 31 inbred mouse strains (Yuan *et al.* 2009). Since the environment was controlled in this study, the variation in IGF-1 level and longevity was mostly attributed to genetic factors. Thus, identifying the genetic regulation of IGF-1 levels will help elucidate the genetic regulation of aging.

Inbred mouse strains are excellent models for identifying complex trait genes. It has been suggested that loci regulating both human and mouse complex traits are often concordant (Wang & Paigen 2005). QTL analysis is an unbiased statistical analysis that has been used for more than 20 years to study the genetic regulation of complex traits. This type of analysis helps identify genomic loci that are linked to variation of a particular trait. In addition, genome-wide association analysis has become accessible in mice through the development of high density SNP databases, high throughput phenotyping, and improved statistical methods. Meanwhile, the identification of the QTL gene has become easier. Our group previously developed a mouse bioinformatics toolbox to add further evidence to each of the genes located under a QTL (DiPetrillo *et al.* 2005; Burgess-Herbert *et al.* 2008). The criteria for the genes located under the QTL include haplotype block analysis, difference in expression between the QTL strains, and non-synonymous coding polymorphisms. The combination of these various genomic analyses and bioinformatic tools available in the mouse allows for the identification of candidate genes within each QTL.

Previously, three crosses identified QTL for IGF-1 levels; these crosses involved C57BL/6J (B6), BALB/cJ (BALB), DBA/2J (D2), Du6i and C3H/HeJ (C3H) (Brockmann *et al.* 2000; Rosen *et al.* 2000; Harper *et al.* 2003; Hanlon *et al.* 2006). The three crosses identified 23 QTL on 14 chromosomes. Eight of these QTL co-localize at four different locations. However, bioinformatic methods to narrow a QTL become more powerful when the QTL is replicated in additional studies with different strains. In this report, we used two different approaches to identify loci influencing IGF-1 levels: a QTL cross between MRL/MpJ (MRL) and SM/J (SM), and a haplotype association mapping (HAM) analysis using 28 inbred strains. We identified 11 QTL, nine of which replicated QTL found previously, and we tentatively identified *Igf1* as the most likely QTL gene for a major locus on Chr 10.

Results

Genetic loci that regulate plasma IGF-1 in the MRL×SM intercross

IGF-1 levels of MRL, SM, F1 and F2 mice—To find the genetic loci that regulate IGF-1, we intercrossed MRL females and SM males, two strains with significantly different IGF-1 levels as previously reported (Yuan *et al.* 2009). First, we confirmed that MRL had a significantly higher ($P < 0.0001$) plasma IGF-1 level than SM at 10 weeks old (626 ± 11 *vs.* 153 ± 19 ng/ml in females and 594 ± 10 *vs.* 234 ± 9 ng/ml in males) (Table 1; Fig. S1, Supporting Information). We compared IGF-1 in the F1 mice: (MRL×SM) F1 mice were created by crossing MRL females with SM males; reciprocal (SM×MRL) F1 mice were created by crossing SM females with MRL males. Female F1s did not differ, but (MRL×SM) F1 males had significantly lower IGF-1 compared to (SM×MRL) F1 males ($P < 0.001$; Table 1; Fig. S1, Supporting Information). This result suggests that IGF-1 level may be partially inherited through genes located on the sex chromosomes or mitochondrial DNA, imprinted genes or maternal intrauterine environment. We crossed (MRL×SM) F1 mice to produce 136 female and 235 male F2 mice; we collected more males so that we could also use this cross for a kidney phenotype expressed preferentially in males. The IGF-1 levels of F2 mice were normally distributed, varying between 195 and 817 ng/ml in females and 212 and 744 ng/ml in males (Fig. S1, Supporting Information).

QTL analysis—Using R/qtl, we performed a three-step QTL analysis for plasma IGF-1 of the 371 F2 mice. The QTL, LOD score and 95% confidence interval (CI) are represented in Fig. 1 and summarized in Table 2. First, we used sex as an additive covariate and identified three main effect QTL: one significant QTL on Chr 10 at 43 cM (86 Mb) (*Igf1q4*) and two suggestive QTL on Chr 9 at 26 cM (48 Mb) (*Igf1q3*);) and Chr 17 at 10 cM (18 Mb) (*Igf1q5*) (Fig. 1 and 2A–C, Table 2). All QTL were named by following the rules of QTL nomenclature (Maltais *et al.* 2002). Second, we used sex as an interactive covariate; the difference between the first and second models is a test for a sex-specific interactive QTL. If the difference shows a LOD score higher than 2, we judged that QTL to be sex specific. We identified one female-specific QTL on Chr 15 at 40 cM (85 Mb) (*Igf1q6*) (Fig. 1 and 2D, Table 2). For *Igf1q3*, we determined that the MRL allele conferring high levels of IGF-1 is recessive because mice homozygous for MRL had higher IGF-1 than heterozygous or homozygous SM mice (Fig. 2E). For two loci, *Igf1q4* and *5*, we determined that the mode of inheritance is additive because mice with heterozygous alleles of MRL and SM had intermediate IGF-1 levels compared with mice homozygous MRL or SM (Fig. 2F and G). For the sex-specific QTL, *Igf1q6*, female mice carrying at least one SM allele had higher IGF-1 levels compared to homozygous female MRL mice (Fig. 2H). Third, we performed a pairscan analysis to identify interactive QTL, but found no significant interactions.

Igf1q4 had an extremely high LOD score of 31.8, suggesting that it is the major determinant of IGF-1 levels in this F2 population. Regression analysis estimated that *Igf1q3, 4, 5,* and *6* explained 3.3%, 31.4%, 3.4% and 1.6% of the variation, sex explained 4.6%, and interaction between sex and *Igf1q6* explained 1.2% (Table S1, Supporting Information). To reveal other regions that could be hidden by the strong effect of the *Igf1q4* locus, we carried out the QTL analysis again but adjusted for the *Igf1q4* locus by adding its nearest SNP (rs6394370) as an additive covariate. This recalculation confirmed the previously identified main effect QTL on Chrs 9 and 17, but failed to confirm the Chr 15 female-specific QTL, most likely due to the lower number of females compared to males.

Genetic loci that regulate IGF-1 among inbred strains

HAM analyses for IGF-1 levels—Plasma IGF-1 levels, measured at 6 months in 28 domesticated inbred strains, published previously (Yuan *et al.* 2009) and available online (http://www.jax.org/phenome), were used to carry out HAM analyses. The mean value of IGF-1 levels and the abbreviations of strain names are listed in Table S2, Supporting Information. HAM analyses, which are similar to genome-wide association studies in humans, were conducted separately in males and females. We identified seven loci associated with IGF-1 levels: three in females (*Igf1q7,8,9*) and four in males (*Igf1q10,11,12,13*) (Table 2). *Igf1q8* on Chr 10 at 89 Mb was the locus most significantly associated with IGF-1 variation ($-\log_{10}P = 4.12$) (Table 2); this is the same locus (*Igf1q4*) found in the MRL×SM intercross, but it requires a different name according to nomenclature rules.

Co-localization of IGF-1 QTL detected by HAM and QTL crosses—Among the four major QTL identified in the (MRL×SM) F2 population, three (*Igf1q4, 5, 6*) overlap with previously reported QTL identified in other crosses (Fig. 3) (Brockmann *et al.* 2000;Rosen *et al.* 2000;Brockmann *et al.* 2001;Harper *et al.* 2003). All three loci that were identified by the HAM analysis in females (*Igf1q7, 8, 9*) co-localized with QTL identified in previous crosses: Chr 2 (153 Mb), Chr 10 (89 and 92 Mb), Chr 17 (25 Mb) (Fig. 3). In males, three of the four QTL identified by HAM analysis also co-localized with QTL identified in previous crosses: Chr 1 (159 Mb) and Chr 3 (52 and 68 Mb). Two of the three previous QTL studies also identified a locus on Chr 10 around 87 Mb associated with IGF-1

levels (Brockmann *et al.* 2000;Rosen *et al.* 2000;Brockmann *et al.* 2001;Harper *et al.* 2003;Hanlon *et al.* 2006).

Igf1q8, IGF-1 levels, and lifespan of inbred strains—According to the SNPs located at *Igf1q8,* the 28 inbred strains can be subdivided into five haplotype groups (Table S3, Supporting Information). Strains in haplotype group 1 have a significantly lower IGF-1 level than strains in haplotype groups 2 and 3 (Fig. 4); haplotype groups 4 and 5 were excluded from this analysis because these two groups have only two and three strains respectively. ANOVA analysis revealed that haplotypes 1–3 could explain 63.6% ($P < 0.001$) of the variation in IGF-1 level among these inbred strains.

Because of the important role of IGF-1 in regulating aging and longevity, we tested if this locus also affects longevity. We examined the lifespan of all inbred strains having haplotypes 1–3. The strains in haplotype group 1 (low IGF-1 level) had the longest lifespans. In females, the median lifespans of strains in haplotype groups 1, 2, and 3 were 748 (95% confidence interval: : 692–773), 623 (556–659) and 659 (624–683) days. In males, the median lifespans of strains in haplotype groups 1, 2, and 3, were 798 (764–812), 550 (504–581) and 679 (637–701) days. In both sexes, the 95% CI of haplotype group 1 did not overlap with the 95% CI of haplotype groups 2 and 3, indicating that this difference was significant ($P < 0.05$). In males, the 95% CI of the median lifespan of haplotype group 2 did not overlap with that of group 3, indicating that the difference between these two group was also significant (Fig. 4). The LogRank test shows that for both males and females, the survival curves of strains in haplotype groups 1, 2, and 3 are significantly different ($P < 0.001$).

Bioinformatics analysis identified candidate genes in the QTL on Chr 10

Chr 10 QTL—*Igf1q4* and *Igf1q8* overlapped between our QTL and HAM analysis. This locus has also been found in two other crosses (Fig. 3): the B6×C3H cross, with the C3H allele accounting for high IGF-1 (Rosen *et al.* 2000); and the 4-way cross among strains BALB, B6, C3H, and D2, with the C3H allele accounting for high IGF-1 and the D2 allele accounting for low IGF-1 (Harper *et al.* 2003). A cis expression QTL has also been identified for *Igf1* in a B6×D2 cross (S-W Tsaih, personal communication) (Mehrabian *et al.* 2005). However, D2 was the high IGF-1 allele and B6 was the low IGF-1 allele in this cross. These results are contradictory to the direction of the alleles in the 4-way cross where D2 was the low IGF-1 allele and B6 the high IGF-1 allele. Thus we decided to exclude the 4-way cross from our analysis because we could not determine with certainty if D2 should be considered a high or low allele strain at this locus. We also excluded the D2×Dui6 QTL because its peak is located at 72 Mb, almost 20 Mb upstream of *Igf1q4* and *Igf1q8* (Brockmann *et al.* 2001). Therefore, our analysis focused on the MRL×SM and B6×C3H crosses.

Haplotype analysis—We performed a haplotype analysis between the parental strains that led to the QTL on Chr 10. The confidence interval of *Igf1q4* is very narrow (85–88) because of the high LOD score (Table 2). The suggestive interval for *Igf1q8* is very narrow as well (88.7–89.1) (Table 2). To prevent us from missing any potential candidate genes in this region, we extended our search 4 Mb to each side of the *Igf1q8* peak between 85 and 93Mb. As explained previously, we looked for regions within this QTL that were identical between the strains with low IGF-1 alleles (B6 and SM) and identical between the strains with high IGF-1 alleles (MRL and C3H), but different between the high-allele and low-allele strains. Using the mouse diversity array dataset (Yang *et al.* 2009) and its web tool, available at http://cgd.jax.org/straincomparison, we identified several regions of interest. These genomic regions contain a total of 21 genes, which are indicated in Fig. 5. Among

them, the *Igf1* gene itself was identified, and we first investigated the gene for additional bioinformatic evidence.

Igf1 as a candidate gene for the Chr 10 IGF-1 QTL locus

The molecular basis of a QTL gene could be a change between the parental strains in an amino acid that affects the function of a protein or a difference in expression of the gene.

Igf1 transcripts—*Igf1* possesses at least eight transcripts that are reported in the Ensembl database (Version 55.37h). The list of transcripts, exons and positions were extracted from Ensembl and are available in Table S4, Supporting Information. However, it is likely that the number of annotated transcripts is incomplete (Dr. Joel Graber, personal communication). The transcription of five transcripts (ENSMUST00000095360, ENSMUST00000105300, ENSMUST00000122386, ENSMUST00000121952, ENSMUST00000075330) starts at exon1, while the transcription of three transcripts (ENSMUST00000062862, ENSMUST00000122100, ENSMUST00000121161) starts at exon 2. It was previously reported that circulating IGF-1 is produced mostly in liver and transcribed from exon 2 (Adamo *et al.* 2006).

Sequencing—We first searched for a polymorphic difference between MRL and SM that could account for this QTL, either a difference in a coding region sequence or a difference in the regulatory regions. We reasoned that the C3H allele should be identical to the MRL allele and that the B6 allele should be identical to the SM allele. We resequenced exons, intron-exon splice sites, and 5', 3' UTR in all four strains (MRL, C3H, B6, SM) based on the eight transcripts reported in Ensembl. We also resequenced D2 to investigate any polymorphism between this strain and B6. In addition, we resequenced an additional 3Kb in the 3' region of the gene because larger transcripts in the 3' UTR region have been reported in RefSeq. We did not identify any synonymous or non-synonymous coding polymorphisms. Thus, there is no evidence for a functional change in the protein. However, we did identify a specific insertion, TGCTGC, in D2 in the 5' UTR of ENSMUST00000062862 (Table 3). We also identified a polymorphism (C/A). This SNP (rs29342496) is located in the 5' UTR of two transcripts (ENSMUST00000062862 and ENSMUST00000121161) and upstream of a third transcript (ENSMUST00000122100) (Table S4, Supporting Information). The presence of the C allele (SM and B6) creates a potential binding site for MyoD. No polymorphism was identified in the 3' UTR annotated in Ensembl (www.ensembl.org). However, we identified two T indels in the additional 3 Kb resequenced in the 3' region of the gene. We investigated the genomic loci surrounding the 2 SNPs in the 3' region for potential miRNA binding sites but did not identify any binding site at these genomic locations. The single nucleotide polymorphisms, rs29342496 in 5' UTR and the two polymorphisms in 3' region segregate high IGF-1 strains C3H and MRL from low IGF-1 strains SM and B6; based on these polymorphisms, D2 resembles the high allele strains.

Expression of the Igf1 gene—To measure expression we used a microarray that assessed three IGF-1 transcripts (ENSMUST00000075330, ENSMUST00000095360, ENSMUST00000062862). The three transcripts were tagged by a total of 43 probes (Table S5, Supporting Information) for *Igf1*. We verified that none of the polymorphisms we identified through re-sequencing was located at probe binding sites. We first assessed expression in the parental strains MRL and SM using livers of three males and three females from each strain at 13 weeks of age. For all three transcripts, MRL mice had a higher expression level than SM, which mirrored the difference in protein levels. The difference was statistically significant in females ($P < 0.05$) and close to significant in males for one transcript ($P = 0.07$) (Table 4). The difference in expression was in the same direction in

both sexes (+ 1.39 and + 1.19 fold change in females and males respectively in MRL compared to SM mice).

We then investigated if *Igf1* expression was *cis* or *trans* regulated using RNA from livers of 282 F2 mice on the same microarray platform. The expression QTL analysis (using sex as an additive covariate) showed that none of the three transcripts was cis regulated (LOD score < 1) (Table 4). Since the level of expression of a transcript was calculated as the average expression of all probes mapping to the transcript, we then investigated the expression level of each of the 43 probes that tag the *Igf1* gene (probe location indicated in Table S5, Supporting Information); seven of these were cis-regulated (LOD score > 3) (Table 4). Considering that MRL was the high-allele strain for the QTL as defined by circulating IGF-1 protein, we had expected a higher expression in this strain. However, four of the seven probes, including the three highest LOD scores (1054540, 317441 and 732594), tagged exons 1 and 2, and had SM as high allele. Three of the seven had much lower LOD scores (533871, 18947, 231840), tagged exons 4 and 6, and had MRL as high allele. Exons 1 and 2 are in the UTR; exons 4 and 6 are in the mature protein. These results suggest that transcriptional regulation is not responsible for the QTL observed between MRL and SM and that regulation may occur at the translational level. The C/A SNP (rs29342496), located in the 5' UTR of two transcripts that express circulating IGF-1, segregates between the high IGF-1 strains and the low IGF-1 strains and may be the causal polymorphism.

Bioinformatics analyses identify candidate genes in the QTL on Chr 10

Since *Igf1q4* had a LOD score of 31.4 and we did not find any molecular evidence for the *Igf1* gene itself, we also carried out bioinformatics analysis in the entire QTL to determine if other viable candidate genes were present.

Non synonymous polymorphism—For the 20 candidate genes identified by haplotype analysis (after removing *Igf1*), we first searched for a non-synonymous polymorphism. Using the 7.8 million SNP database of the Center for Genome Dynamic (CGD) (Szatkiewicz *et al.* 2008), we compared the high IGF-1 (MRL and C3H) and low IGF-1 (SM and B6) strains for the candidate region. We identified only two genes, *Nr1h4, (nuclear receptor subfamily 1, group H, member 4)* and *4930547N16Rik,* carrying known segregating, non synonymous polymorphisms (Table 5). None of the polymorphisms were predicted to change the function of the protein using SIFT and Polyphen.

Expression analysis—Among the 20 genes that were not identical by descent (IBD), microarray analysis showed that 10 genes were differentially expressed. Two of them, *Timp3 (tissue inhibitor of metalloproteinase 3)* and *Slc17a8 (solute carrier family 17 (sodium-dependent inorganic phosphate cotransporter), member 8),* were differentially expressed in both sexes ($P < 0.001$). The expression of both genes was significantly increased in SM mice compared to MRL mice (Table 5; Table S6, Supporting Information). We then used our microarray results in the 282 F2 mice to investigate if any of the expression differences observed between parental strains in the 10 genes were cis or trans regulated. Using the expression level as phenotypic trait, we carried out expression QTL analysis. We identified four genes that were cis regulated in the F2 mice: *Timp3, Slc17a8, Stab2 (Stabilin 2) and Pah (Phenylalanine Hydroxylase)* (Table 5). The difference in expression between the parental strains was not significant for *Stab2*, but we did identify a cis expression QTL, and this result may be due to the low number of mice in the parental group. We present evidence for six additional genes, two with a non-synonymous polymorphism and four with a cis-regulated expression difference.

Discussion

Here, we report the results of two mapping approaches and the follow-up of one major QTL for IGF-1 level. We performed a QTL analysis between two inbred strains with a highly varied IGF-1 level (MRL and SM) and a HAM analysis in females from 28 inbred strains. Using this strategy, we insured that either new QTL would be found or that known QTL would be confirmed and narrowed. We identified a total of 11 QTL (four in the MRL×SM cross and seven in the HAM, three in females and four in males). Two of the loci (*Igf1q4 and Igfq8* on Chr 10, and *Igf1q5* and *Igf1q9* on Chr 17) overlapped in both analyses, leaving us with nine different loci on Chrs 1, 2, 3, 9, 10, 15, 16 and 17, with two loci on Chr 3. The strongest QTL was found on Chr 10 (*Igf1q4*) with a LOD score of 31.8, overlapping with *Igf1q8* from the HAM analysis. This QTL had already been identified in one previous cross (B6×C3H), not only confirming our results but also helping in narrowing the locus. Therefore, we investigated the strongest QTL on Chr 10 and found that the haplotypes at this locus also segregated with lifespan. Using our bioinformatic tool, we searched for candidate genes in the region. To our surprise, we could not find completely convincing evidence for the *Igf1* gene itself located at 87.3 Mb. We added some evidence to six other potential candidate genes located in this region of Chr 10 based on sequence and expression differences in the parental strains.

Using the bioinformatic toolbox to identify candidate genes for circulating IGF-1 level

This present study complements and confirms previous QTL results for plasma IGF-l. The QTL on Chrs 2, 3, 10, 15 and 17 overlapped with previously observed QTL, while the QTL on Chrs 9 and 16 were new (Brockmann *et al.* 2000; Rosen *et al.* 2000; Harper *et al.* 2003; Hanlon *et al.* 2006). Although QTL analysis is a powerful method for identifying genetic loci that associate with a particular quantitative trait, identifying the underlying QTL genes is challenging. Recently, Dr. Paigen's laboratory developed a series of bioinformatic tools to add evidence to each gene located within a QTL (DiPetrillo *et al.* 2005; Burgess-Herbert *et al.* 2008). We showed here that the availability and development of new bioinformatics resources and tools, such as gene expression databases and haplotype analysis of combined crosses, helps narrow QTL and identify potential candidate genes. While we followed-up our strongest QTL (*Igf1q4*), additional QTL that are currently being investigated using our bioinformatic tool are *Igf1q5* and *Igf1q6* on Chrs 15 and 17 because QTL from other crosses have been found at these locations.

Using strain survey results to improve the experimental design of QTL study

Although previous QTL studies have successfully identified several QTL (Rosen *et al.* 2000; Harper *et al.* 2003; Hanlon *et al.* 2006), these studies were designed without considering the variation of IGF-1 levels in the whole mouse family. For example, B6 and C3H are significantly different in IGF-1 levels and are used in two of three previous IGF-1 QTL analyses. However, based on our strain survey study, the C3H IGF-1 level is only 25% higher than B6, while the biggest differences of IGF-1 levels were found in NOD.B10Sn-*H2^b*/J (NOD.B10) *vs.* SM, and MRL *vs.* SM, for females and males respectively (Yuan *et al.* 2009). While intercrossing two strains with large differences in IGF-1 level increases the likelihood of identifying QTL, this approach does not guarantee the detection of unknown QTL; rather it assures the identification of the most important QTL with the largest impact on regulating IGF-1 level.

IGF-1 and lifespan

Previous studies have shown that IGF-1 level is negatively correlated with longevity among mouse inbred strains (Yuan *et al.* 2009). These results in mice are consistent with results in other species: reduced IGF-1 signaling extends lifespan in worms and flies reviewed in

(Kuningas *et al.* 2008). In our QTL study, Chr 10 was a major contributor of plasma IGF-1 level. Our analysis confirms a relationship between a locus regulating IGF-1 level (*Igf1q4* and *Igf1q8*) and longevity: the haplotype associated with lower IGF-1 was also associated with an increase in median lifespan and a lower mortality rate. These results suggest that this locus may play a role in delaying aging and extending longevity by reducing IGF-1 level at young age. It also suggests that identifying the underlying genes of the IGF-1 QTL will help to understand the genetic regulation of lifespan.

IGF-1 regulation is complex

Although we did not obtain any convincing evidence for *Igf1* using our bioinformatic tools, we believe that *Igf1* may be regulated by other translational and post-translational events. Adamo et al. have shown that an increase in transcription from exon 2 leads to an increased level of circulating IGF-1 (Adamo *et al.* 2006). In liver, exon 2 was up-regulated in C3H mice compared to B6, which contributed to the increase in circulating IGF-1 level. The expression analysis revealed that for all transcripts, *Igf1* was differentially expressed between MRL and SM in females but not in males, which was consistent with the HAM analysis, since the *Igf1q8* was identified only in the females of inbred strains. However, the difference in expression was not cis regulated. IGF-1 level is known to be influenced by other biological mechanisms, including at the *Igf1* mRNA expression level (Adamo *et al.* 2006), by protein degradation and the interaction with IGF-1 binding proteins (Rui *et al.* 2001; Patel *et al.* 2008), and by feedback regulation within the IGF-1 signaling pathway (Salminen & Kaarniranta 2009). We identified a polymorphism (rs29342496) in the 5' UTR of two transcripts starting at exon 2 that segregated between the high allele strains and the low allele strains. Since the expression of *Igf1* was not cis regulated, we hypothesize that this polymorphism may play a role in the rate of translation of the transcripts in which the SNP is located in the 5' UTR. Interestingly, human population studies have also suggested that polymorphisms between the promoter sequence and intron 2, including exon 2, are associated with IGF-1 level, birth weight and postnatal growth (Johnston *et al.* 2003). Combined with our results in mice, these studies suggest that IGF-1 level is regulated by a mechanism evolutionarily conserved between mice and humans. In addition, these results imply that under the hypothesis that the *Igf1* gene is the QTL gene in the MRL×SM cross, a simple gene expression analysis of the *Igf1* transcripts is not sufficient and that clarifying its regulation will require further investigation.

Potential candidate genes influencing IGF-1 levels

A very high LOD score often indicates that more than one gene is responsible for the QTL. To our knowledge, no such example exists for the IGF-1 trait, but this phenomenon has been observed in mice and rats for other complex traits such as epilepsy, bone mineral density phenotype and blood pressure (Legare *et al.* 2000; Saad *et al.* 2001; Beamer *et al.* 2007). In these studies, subsequent congenic mouse work helped dissect each locus thought to be due to a single gene into multiple QTL and multiple genes. We are currently developing a series of overlapping congenic mice carrying the C3H allele at the Chr 10 locus on a B6 background. The size of the C3H chromosomal segment varies, some contain the *Igf1* gene while others do not. In the future, these new sub-congenic mice will help determine whether *Igf1* is the QTL gene for *Igf1q4* and responsible for extending longevity, or whether other genes are involved as well.

Among the six additional candidate genes for which we found evidence within *Igf1q4*, four of them were cis regulated (*Timp3, Stab2, Pah* and *Slc17a8*) and two had amino acid changes segregating between the parental strains (*4930547N16Rik* and *Nr1h4*). Two of the six candidate genes may potentially play a role in IGF-1 regulation, but further experiments will be needed. *Nr1h4* is a nuclear hormone receptor involved in regulating cholesterol

homeostasis (Sinal *et al.* 2000). However, the knock-out shows a weight loss phenotype after birth when the diet is supplemented with 1% cholic acid, suggesting that this gene might be involved in the GH/IGF-1 pathway and hence, regulate body growth (Sinal *et al.* 2000). *Timp3* is present in all tissues, especially in the placenta. TIMP3, unlike the other TIMP family members, is associated to the extracellular matrix. IGF-1 is mostly transported in the bloodstream as a complex composed of IGFBP3, the acid-labile subunit and IGF-1. (Loechel *et al.* 2000)) have shown that TIMP3 inhibits ADAM 12-S, which cleaves IGFBP-3 and IGFBP-5. ADAM12-S is found only during pregnancy, and TIMP proteins are also inhibitors of metalloproteases. Therefore, our results suggest that TIMP3 may play a role in the determination of circulating IGF-1 level.

To conclude, this study models how to use the strain survey data to guide the genetic analysis of aging-related phenotypes. Based on the strain survey results, we crossed strains that had the most significant differences in the IGF-1 level. We identified a major genetic determinant of IGF-1 level variation on Chr 10 that was also associated with longevity. Based on bioinformatics evidence, we reduced the list of candidate genes to six genes in addition to *Igf1*. No conclusive evidence has so far been found for the *Igf1* gene, and further investigation will be necessary to reveal its regulation. This report also adds further evidence for the important role of the *Igf1* locus in longevity.

Experimental procedures

Mice and Husbandry

MRL/MpJ (MRL) and SM/J (SM) mice, obtained from The Jackson Laboratory, were used to generate (MRL× SM) F1 by crossing MRL females with SM males and reciprocal (SM×MRL) F1 mice by crossing SM females with MRL males. The (MRL female × SM male) F1 mice were intercrossed to generate 371 F2 mice (136 females and 235 males). Over 4 months, F2 mice at wean age were incorporated into the study in staggered cohorts. All mice were housed in a specific pathogen free facility with *ad libitum* access to acidified water (pH 2.8–3.1) and an autoclaved diet with 6% fat (LabDiet® 5K52, PMI Nutritional International, Bentwood, Mo) as described previously (Yuan *et al.* 2009). All experiments were approved by The Jackson Laboratory Animal Care and Use Committee.

DNA isolation and genotyping

F2 mice from the MRL×SM cross were tail-tipped at 2 weeks of age. DNA was extracted using the Gentra kit that utilizes a phenol/chloroform extraction method. DNA was genotyped by the High Throughput Sequenom and Illumina Genotyping facility (http://www.hpcgg.org/), using a 760-SNP array; 258 of these SNPs were polymorphic between MRL and SM for an average distance between markers of 5.5 cM, ranging from 0 to 25.1cM. Rs33585432 on Chr 17 was additionally genotyped through resequencing. Physical marker positions were determined in NCBI build 37, and genetic positions were estimated from the newly calculated mouse genetic map (Cox *et al.* 2009).

Tissue collection

Blood, collected at 10 weeks of age, was centrifuged at 14,000 rpm for 5 minutes, and plasma was saved at −20 °C prior to the IGF-1 assay. At 13 weeks of age, liver collection was performed on three males and three females from each parental strain (MRL and SM), three F1 males and three F1 females of the MRL×SM and SM×MRL intercrosses, and 282 of the 371 F2 offspring. Each mouse was housed individually for 3 days prior to the tissue collection and fasted from 8:00 am until the tissue collection took place, between 12pm and 1pm. Liver samples were preserved in RNAlater (Ambion, Applied Biosystems, Foster City, CA) and saved at −80 °C prior to the gene expression study.

IGF-1 assay

IGF-1 levels were measured by a direct radioimmunoassay (RIA, ALPCO Diagnostics, Salem, NH) by the MeCORE Laboratory, Saint Joseph Hospital, Bangor, ME (Rosen *et al.* 2000).

Microarray analysis for liver gene expression

RNA and microarray processing—RNA was labeled and hybridized to the Mouse Gene 1.0 ST microarray (1M) following the manufacturer's protocols (Affymetrix, Santa Clara, CA). Raw data were imported in the R language/environment version 2.7.2 for data analyses (www.rqtl.org). Quality control and quantiles normalization (Bolstad *et al.* 2003) were performed with the *affy* V 1.20.0 and *preprocessCore* V1.6 packages from Bioconductor (www.bioconductor.org). The transcript analysis was performed with a custom CDF file (Dai *et al.* 2005) for Ensembl transcripts (ENST package V.11, 37,264 probesets) from the BrainArray (U. of Michigan) website (http://brainarray.mbni.med.umich.edu). All probes mapping to exons in a transcript were grouped into one probeset, and their average expression was calculated by the median polish on the log2 scale.

Microarray Analysis—Differential expression was tested by fitting a linear model, including genotype and sex, to normalized data in the R/Maanova package. The following linear model was fitted (using the R/maanova software version 1.13 (Wu *et al.* 2003)) to: $y_{ijk} = \mu + G_i + S_j + I(GxS)_{ij} + e_{ijk}$ where, y_{ijk} = normalized $\log_2$-transformed gene expression, μ = overall mean, G_i = effect of ith genotype, S_j = effect of jth sex, $I(GxT)_{ij}$ = effect of genotype by sex interaction and e_{ijk} = residual effect. *P* values for the G_i, S_j, and $I(GxS)_{ij}$ terms were calculated by permuting the sample labels and refitting the model 1,000 times. *Post-hoc* tests were performed on a cell-means model using $y_{ij} = \mu_i + e_{ij}$ where, y_{ij} = normalized $\log_2$-transformed gene expression, μ_i = mean for i^{th} group and e_{ij} = residual effect. The term μ_i is the mean intensity of samples from six experimental groups: MRL males (M.m), MRL females (M.f), SM males (S.m), SM females (S.f), F1 males (F1.m), F1 females (F1.f). *Post hoc* tests were performed as follows: additive_female (a_f = M.f−S.f), additive males (a_m = M.m−S.m), dominance females (d_f = F1.f−(M.f + S.f)/2), and dominance males (d_m = F1.m−(M.m + S.m)/2). *F* values were calculated using shrinkage estimates of error variance (Cui *et al.* 2005). *P* values were computed from 1000 permutations of the data (Yang & Churchill 2007) and corrected for multiple comparisons with the q-value method (Storey 2002).

QTL analysis

QTL analysis was performed for plasma IGF-1, gene expression level of the three transcripts (ENSMUST00000075330, ENSMUST00000095360, ENSMUST00000062862) and the 43 probes that tagged *Igf1* using R/qtl (v1.09-43) (www.rqtl.org) (Broman *et al.* 2003). We performed a 3-step QTL analysis. First, phenotypes were analyzed for main-effect QTL using sex as an additive covariate. Second, sex was then added as an interactive covariate into our previous model; the difference between the two models provided a test for sex-specific QTL (i.e., one genotype increases expression of the trait in males while lowering the expression of the same trait in females). Third, an epistatic effect was investigated using the pairscan function of R/qtl that tests for interacting QTL. Alternatively, we added rs6394370 on Chr 10 as an additive covariate in each model to control for the Chr 10 locus. Thresholds for significant LOD scores (*P* value < 0.05) and suggestive LOD scores (*P* value < 0.63) were based on 1000 permutations of the observed data. The 95% CI was calculated using the Bayesian method. All suggestive and significant QTL in addition to sex were incorporated into a multiple regression model. The proportion of variance of IGF-1 level explained by sex

and each of the QTL was estimated. For the QTL analysis of IGF-1, we used the standard interval mapping expectation maximization (EM) method with a 1 cM increment; for the QTL analysis of the transcripts, we used the Haley-Knott (HK) regression method with a 2 cM increment.

Haplotype association mapping (HAM)

Mean value of IGF-1 levels—IGF-1 levels of inbred mouse strains at 6 months were reported earlier (Yuan *et al.* 2009), but wild-derived inbred strains were excluded because they are usually phenotypic and genotypic outliers. Within each strain and sex, we used the extreme studentized deviate method to determine outliers of IGF-1 level. After removing two female and three male outliers, the mean values of IGF-1 for each strain/sex were calculated (Table S2, Supporting Information).

Genotype data set—The genotype data of these 28 domesticated inbred strains were downloaded from the SNP database of the Center for Genome Dynamics at The Jackson Laboratory (http://cgd.jax.org/datasets/popgen/70KSNP.shtml).

HAM analysis—Mean values of IGF-1 were input as vectors, and genotype data across multiple inbred mouse strains were input as a matrix. The Hidden Markov Model (HMM) was applied to fit five states at each SNP, for the primary purpose of imputing missing genotypes and for the secondary purpose of haplotype identification (Szatkiewicz *et al.* 2008). At each SNP, we determined the strain distribution pattern using the HMM smoothed haplotype states. Regression-based test statistics were computed to measure the strength of association between genotype and phenotype. Because the segregation of strains into genotypic groups varies widely over haplotype blocks, *P* values, rather than the test statistics, were compared between haplotype blocks. All *P* values were transformed using $-\log_{10}(P \text{ value})$ in the scan plots. We controlled the type I error rate for multiple testing, which results from a genome-wide search, using family-wise error rate control (Westfall & Young 1993). We shuffled the strain labels in the phenotype data and kept the genotype data intact. The minimum *P* value was recorded on each permutation, and the distribution of these *P* values provided approximate, multiple, test-adjusted thresholds. All analyses, except the imputation of missing genotypes, were done in the MATLAB computing environment (The Mathworks, http://www.mathworks.com).

Bioinformatics

Re-defining the positions of previously identified QTL—The locations of previously identified QTL were defined by searching the position of the markers in NCBI Mouse Build 37. If the confidence interval of the QTL was not reported, we assigned 20 Mb on each side of the QTL peak as the confidence interval.

Haplotype analysis—We used the new mouse diversity array datasets and a webtool (http://cgd.jax.org/straincomparison/) to perform the haplotype analysis. In the presence of overlapping QTL, we used a haplotype analysis of combined crosses to identify the genes not in IBD by combining the strains that carry the allele associated with the higher phenotype and comparing them to the strains that carry the allele associated with the lower phenotype. If a polymorphism not in IBD was located within 10 Kb of a gene, we considered the gene not to be in IBD as well.

Identification of non-synonymous polymorphisms in candidate genes—To identify the non-synonymous coding polymorphisms between parental strains of each cross, we used the imputed 7.8 million SNPs from the Center for Genome Dynamics at The Jackson Laboratory (http://cgd.jax.org/datasets/popgen/imputed.shtml) (Szatkiewicz *et al.*

2008). We used genotyped SNPs and imputed SNPs with a high threshold of confidence (>0.9).

Sequencing

We used resequencing to confirm the non-synonymous polymorphims (rs29381032 and rs3675893) in *Nr1h4*. We fully sequenced the known exons of IGF-1 in addition to 20 bp in the intron-exon junction. Since the 3' UTR of IGF-1 is not well defined, we added 3 Kb at the 3' end of the gene. Primers were designed using Primer 3 (http://frodo.wi.mit.edu/). Primer pair sequences are available in supplementary materials (Table S4, Supporting Information). PCR product was purified using ExoSAP-IT® (USB Corporation, Cleveland, OH) and run on an ABI3730XL capillary-based sequencing machine (Applied Biosystems, Foster City, CA), available through The Jackson Laboratory Scientific Services. Sequence analysis was performed using Sequencher (version 4.8, Gene Codes Technology, Ann Arbor, MI).

Evaluation of the functionality of the polymorphisms

For the non-synonymous polymorphisms, we evaluated their potential functionality using PolyPhen (http://genetics.bwh.harvard.edu/pph/) and SIFT (http://sift.jcvi.org). If either of the two tools estimated the amino acid change as being deleterious (PolyPhen) or likely to affect protein function (SIFT), we concluded that the amino acid change is functional. For polymorphisms located in the 5' UTR, we searched for a potential binding site at the location of the SNP using the TRANSFAC database and the webtool available at http://www.gene-regulation.com. For the SNPs located in the 3' region of the *Igf1* gene, we looked for potential binding site of miRNA using the database available at http://www.mirbase.org/cgi.

Statistical analyses

Except for the microarray, QTL and HAM analyses, all other statistical analyses were performed in JMP 7.0 (The SAS institute, Cary, NC, USA). ANOVA was used to compare IGF-1 levels among different groups. Lifespan data for the 28 domesticated inbred strains were previously published (Yuan *et al.* 2009) and updated on July 31 2009. We used the Kaplan–Meier method to draw the survival curves and calculate median lifespans and 95% CI. We tested for differences between survival curves using the log-rank method.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Adamo ML, Ma X, Ackert-Bicknell CL, Donahue LR, Beamer WG, Rosen CJ. Genetic increase in serum insulin-like growth factor-I (IGF-I) in C3H/HeJ compared with C57BL/6J mice is associated with increased transcription from the IGF-I exon 2 promoter. Endocrinology 2006;147:2944–2955. [PubMed: 16527837]

Albani D, Batelli S, Polito L, Vittori A, Pesaresi M, Gajo GB, De Angeli S, Zanardo A, Gallucci M, Forloni G. A polymorphic variant of the insulin-like growth factor 1 (IGF-1) receptor correlates with male longevity in the Italian population: a genetic study and evaluation of circulating IGF-1 from the "Treviso Longeva (TRELONG)" study. BMC Geriatr 2009;9:19. [PubMed: 19460140]

Beamer WG, Shultz KL, Ackert-Bicknell CL, Horton LG, Delahunty KM, Coombs HF 3rd, Donahue LR, Canalis E, Rosen CJ. Genetic dissection of mouse distal chromosome 1 reveals three linked BMD QTLs with sex-dependent regulation of bone phenotypes. J Bone Miner Res 2007;22:1187–1196. [PubMed: 17451375]

Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 2003;19:185–193. [PubMed: 12538238]

Bonafe M, Barbieri M, Marchegiani F, Olivieri F, Ragno E, Giampieri C, Mugianesi E, Centurelli M, Franceschi C, Paolisso G. Polymorphic variants of insulin-like growth factor I (IGF-I) receptor and phosphoinositide 3-kinase genes affect IGF-I plasma levels and human longevity: cues for an evolutionarily conserved mechanism of life span control. J Clin Endocrinol Metab 2003;88:3299–3304. [PubMed: 12843179]

Brockmann GA, Haley CS, Wolf E, Karle S, Kratzsch J, Renne U, Schwerin M, Hoeflich A. Genome-wide search for loci controlling serum IGF binding protein levels of mice. FASEB J 2001;15:978–987. [PubMed: 11292658]

Brockmann GA, Kratzsch J, Haley CS, Renne U, Schwerin M, Karle S. Single QTL effects, epistasis, and pleiotropy account for two-thirds of the phenotypic F(2) variance of growth and obesity in DU6i × DBA/2 mice. Genome Res 2000;10:1941–1957. [PubMed: 11116089]

Broman KW, Wu H, Sen S, Churchill GA. R/qtl: QTL mapping in experimental crosses. Bioinformatics 2003;19:889–890. [PubMed: 12724300]

Burgess-Herbert SL, Cox A, Tsaih SW, Paigen B. Practical applications of the bioinformatics toolbox for narrowing quantitative trait loci. Genetics 2008;180:2227–2235. [PubMed: 18845850]

Cox A, Ackert-Bicknell CL, Dumont BL, Ding Y, Bell JT, Brockmann GA, Wergedal JE, Bult C, Paigen B, Flint J, Tsaih SW, Churchill GA, Broman KW. A new standard genetic map for the laboratory mouse. Genetics 2009;182:1335–1344. [PubMed: 19535546]

Cui X, Hwang JT, Qiu J, Blades NJ, Churchill GA. Improved statistical tests for differential gene expression by shrinking variance components estimates. Biostatistics 2005;6:59–75. [PubMed: 15618528]

Dai M, Wang P, Boyd AD, Kostov G, Athey B, Jones EG, Bunney WE, Myers RM, Speed TP, Akil H, Watson SJ, Meng F. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. Nucleic Acids Res 2005;33:e175. [PubMed: 16284200]

DiPetrillo K, Wang X, Stylianou IM, Paigen B. Bioinformatics toolbox for narrowing rodent quantitative trait loci. Trends Genet 2005;21:683–692. [PubMed: 16226337]

Hanlon P, Lorenz WA, Shao Z, Harper JM, Galecki AT, Miller RA, Burke DT. Three-locus and four-locus QTL interactions influence mouse insulin-like growth factor-I. Physiol Genomics 2006;26:46–54. [PubMed: 16782841]

Harper JM, Galecki AT, Burke DT, Pinkosky SL, Miller RA. Quantitative trait loci for insulin-like growth factor I, leptin, thyroxine, and corticosterone in genetically heterogeneous mice. Physiol Genomics 2003;15:44–51. [PubMed: 12865502]

Johnston LB, Dahlgren J, Leger J, Gelander L, Savage MO, Czernichow P, Wikland KA, Clark AJ. Association between insulin-like growth factor I (IGF-I) polymorphisms, circulating IGF-I, and pre- and postnatal growth in two European small for gestational age populations. J Clin Endocrinol Metab 2003;88:4805–4810. [PubMed: 14557458]

Kuningas M, Mooijaart SP, van Heemst D, Zwaan BJ, Slagboom PE, Westendorp RG. Genes encoding longevity: from model organisms to humans. Aging Cell 2008;7:270–280. [PubMed: 18208581]

Legare ME, Bartlett FS 2nd, Frankel WN. A major effect QTL determined by multiple genes in epileptic EL mice. Genome Res 2000;10:42–48. [PubMed: 10645948]

Loechel F, Fox JW, Murphy G, Albrechtsen R, Wewer UM. ADAM 12-S cleaves IGFBP-3 and IGFBP-5 and is inhibited by TIMP-3. Biochem Biophys Res Commun 2000;278:511–515. [PubMed: 11095942]

Longo VD, Finch CE. Evolutionary medicine: from dwarf model systems to healthy centenarians? Science 2003;299:1342–1346. [PubMed: 12610293]

Maltais LJ, Blake JA, Chu T, Lutz CM, Eppig JT, Jackson I. Rules and guidelines for mouse gene, allele, and mutation nomenclature: a condensed version. Genomics 2002;79:471–474. [PubMed: 11944975]

Mehrabian M, Allayee H, Stockton J, Lum PY, Drake TA, Castellani LW, Suh M, Armour C, Edwards S, Lamb J, Lusis AJ, Schadt EE. Integrating genotypic and expression data in a segregating mouse population to identify 5-lipoxygenase as a susceptibility gene for obesity and bone traits. Nat Genet 2005;37:1224–1233. [PubMed: 16200066]

Patel AV, Cheng I, Canzian F, Le Marchand L, Thun MJ, Berg CD, Buring J, Calle EE, Chanock S, Clavel-Chapelon F, Cox DG, Dorronsoro M, Dossus L, Haiman CA, Hankinson SE, Henderson BE, Hoover R, Hunter DJ, Kaaks R, Kolonel LN, Kraft P, Linseisen J, Lund E, Manjer J, McCarty C, Peeters PH, Pike MC, Pollak M, Riboli E, Stram DO, Tjonneland A, Travis RC, Trichopoulos D, Tumino R, Yeager M, Ziegler RG, Feigelson HS. IGF-1, IGFBP-1, and IGFBP-3 polymorphisms predict circulating IGF levels but not breast cancer risk: findings from the Breast and Prostate Cancer Cohort Consortium (BPC3). PLoS One 2008;3:e2578. [PubMed: 18596909]

Rosen CJ, Churchill GA, Donahue LR, Shultz KL, Burgess JK, Powell DR, Ackert C, Beamer WG. Mapping quantitative trait loci for serum insulin-like growth factor-1 levels in mice. Bone 2000;27:521–528. [PubMed: 11033447]

Rui L, Fisher TL, Thomas J, White MF. Regulation of insulin/insulin-like growth factor-1 signaling by proteasome-mediated degradation of insulin receptor substrate-2. J Biol Chem 2001;276:40362–40367. [PubMed: 11546773]

Saad Y, Garrett MR, Rapp JP. Multiple blood pressure QTL on rat chromosome 1 defined by Dahl rat congenic strains. Physiol Genomics 2001;4:201–214. [PubMed: 11160999]

Salminen A, Kaarniranta K. Insulin/IGF-1 paradox of aging: Regulation via AKT/IKK/NF-kappaB signaling. Cell Signal. 2009

Sinal CJ, Tohkin M, Miyata M, Ward JM, Lambert G, Gonzalez FJ. Targeted disruption of the nuclear receptor FXR/BAR impairs bile acid and lipid homeostasis. Cell 2000;102:731–744. [PubMed: 11030617]

Stewart CE, Rotwein P. Growth, differentiation, and survival: multiple physiological functions for insulin-like growth factors. Physiol Rev 1996;76:1005–1026. [PubMed: 8874492]

Storey JD. A direct approach to false discovery rates. Journal of the Royal Statistical Society Series B 2002;64:479–498.

Szatkiewicz JP, Beane GL, Ding Y, Hutchins L, Pardo-Manuel de Villena F, Churchill GA. An imputed genotype resource for the laboratory mouse. Mamm Genome 2008;19:199–208. [PubMed: 18301946]

van Heemst D, Beekman M, Mooijaart SP, Heijmans BT, Brandt BW, Zwaan BJ, Slagboom PE, Westendorp RG. Reduced insulin/IGF-1 signalling and human longevity. Aging Cell 2005;4:79–85. [PubMed: 15771611]

Wang X, Paigen B. Genetics of variation in HDL cholesterol in humans and mice. Circ Res 2005;96:27–42. [PubMed: 15637305]

Westfall PH, Young SS. On Adjusting P-Values for Multiplicity. Biometrics 1993;49:941–945.

Wu, H.; Kerr, MK.; Cui, X.; Churchill, GA. MAANOVA: A Software package for the analysis of spotted cDNA microarray experiments. In: P, G.; G, ES.; I, RA.; Z, SL., editors. The analysis of gene expression data: methods and software. NY: Springer; 2003.

Yakar S, Setser J, Zhao H, Stannard B, Haluzik M, Glatt V, Bouxsein ML, Kopchick JJ, LeRoith D. Inhibition of growth hormone action improves insulin sensitivity in liver IGF-1-deficient mice. J Clin Invest 2004;113:96–105. [PubMed: 14702113]

Yang H, Churchill G. Estimating p-values in small microarray experiments. Bioinformatics 2007;23:38–43. [PubMed: 17077100]

Yang H, Ding Y, Hutchins LN, Szatkiewicz J, Bell TA, Paigen BJ, Graber JH, de Villena FP, Churchill GA. A customized and versatile high-density genotyping array for the mouse. Nat Methods 2009;6:663–666. [PubMed: 19668205]

Yuan R, Tsaih SW, Petkova SB, Marin de Evsikova C, Xing S, Marion MA, Bogue MA, Mills KD, Peters LL, Bult CJ, Rosen CJ, Sundberg JP, Harrison DE, Churchill GA, Paigen B. Aging in inbred strains of mice: study design and interim report on median lifespans and circulating IGF1 levels. Aging Cell 2009;8:277–287. [PubMed: 19627267]
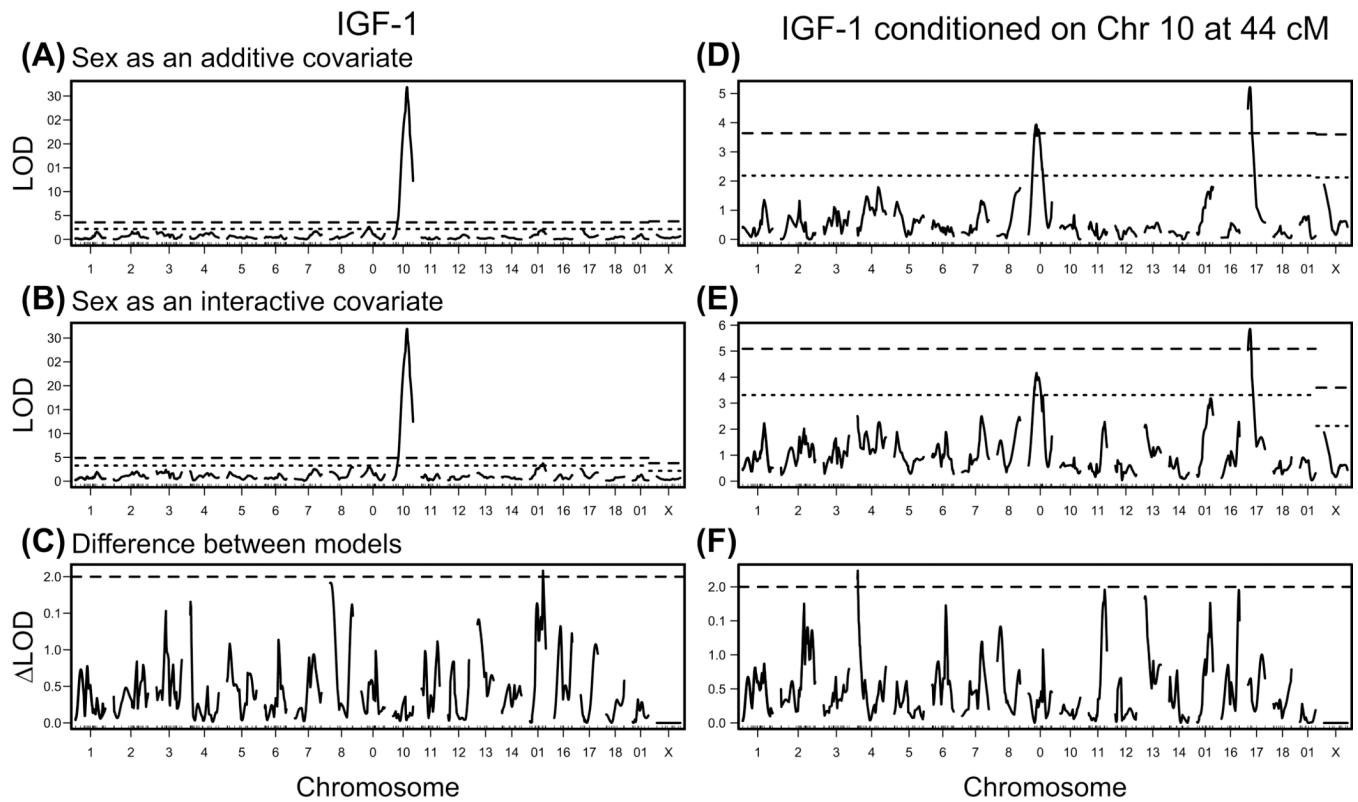
Fig. 1.
Genome-wide scan for IGF-1 level in 371 F2 mice. Analysis was performed with sex as an
(A) additive covariate and (B) interactive covariate. The difference between both models (C)
with a delta LOD > 2 indicates a sex-specific QTL. In (D), (E) and (F), we added rs6394370
as an additive covariate to the previous models to adjust for the Chr 10 *Igf1* locus.
For each model, data were permuted 1000 times to determine the genome-wide level of
significance. The threshold of significance for the Chr X was determined with 17,820
permutations. The dashed line represents the threshold of significance (*P* = 0.05), and the
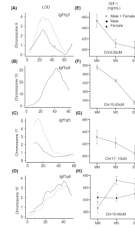dotted line represents the threshold for suggestive QTL (*P* = 0.63).

Fig. 2.
QTL and allele effect plot of the main effect and sex-specific QTL. The QTL plot is represented in (A), (B), (C) and (D) for Chrs 10, 9, 17 and 15 respectively. For Chrs 10, 9 and 17, the QTL plot is represented with sex as an additive covariate (solid line). For Chrs 9 (B) and 17 (C), rs6394370 was added as an additive covariate to adjust for the *Igf1* locus (dotted line). For the sex-specific QTL on Chr 15 (D), sex was added as an additive and interactive covariate (solid line); rs6394370 was then added an additive covariate to adjust for the *Igf1* locus (dotted line). For the allele effect plots represented in (E), (F), (G) and (H), homozygous mice for the MRL allele are represented as MM, the heterozygous mice as MS and the homozygous mice for the SM allele as SS. IGF-1 means and SE were calculated in males and females together for the main effect QTL on Chrs 10, 9 and 17 in (E), (F), (G) and in males and females separately for the female-specific QTL on Chr 15 (H). IGF-1 levels are indicated as mean +/−SE in ng/mL.
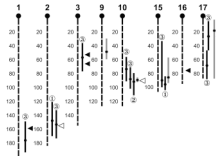
Fig. 3.
Comparison of QTL in this study with those found previously. Newly identified QTL (wide gray bars), HAM in females (white triangular arrows) and in males (black triangular arrows) overlap with previously identified IGF-1 QTL (thin black bars). QTL identified in ① Du6i×D2, ② B6×C3H, ③ (BALB×B6)×(C3H×D2). Bars represent the confidence intervals of QTL. Dots in the bars indicate the locations of peaks of the QTL. Numbers on chromosomes are in Mb.
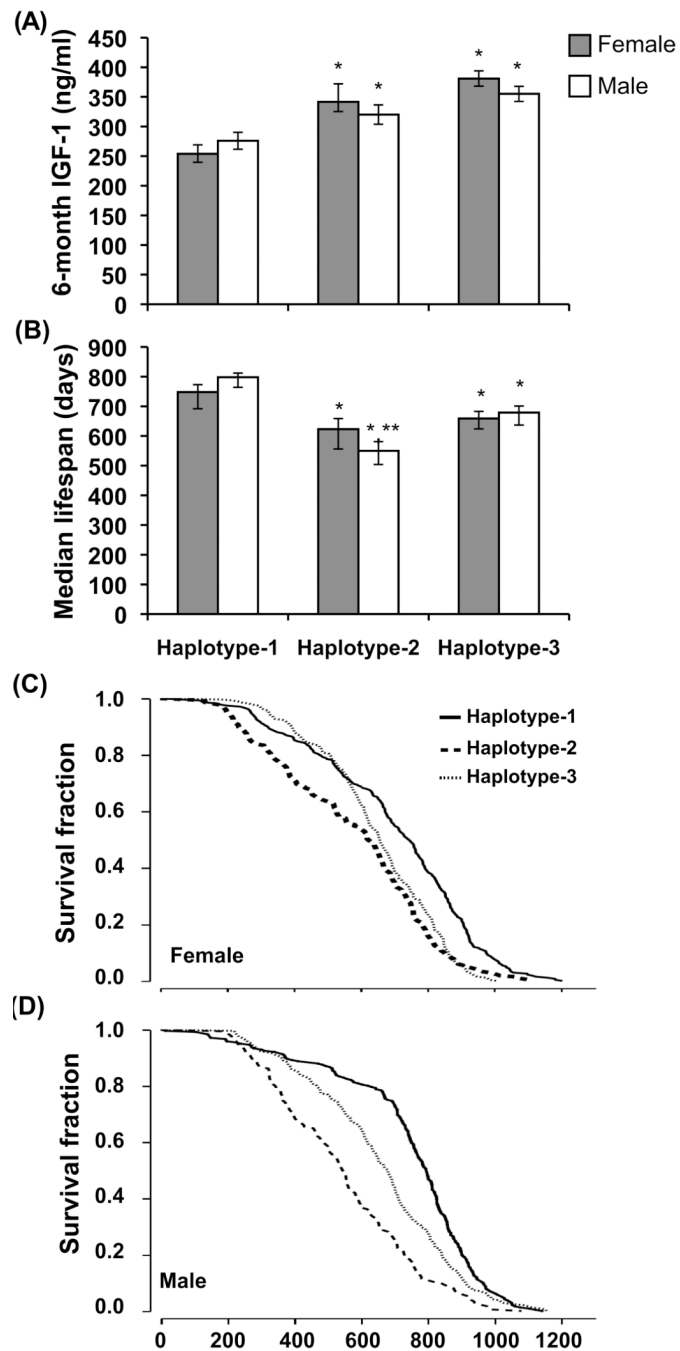
Fig. 4.
Haplotypes of Chr 10 HAM associated with IGF-1 levels and with longevity. Haplotypes are defined in Table S3, Supporting Information. Strains with haplotype-1 have significantly lower IGF-1 levels (A) and longer median lifespans (B) than strains with haplotype-2 or -3. Bars in (A) indicate the standard error of mean; bars in (B) indicate the 95% confidence interval. * indicates a significant ($P < 0.05$) difference with the haplotype-1 group; ** indicates a significant ($P < 0.05$) difference with haplotype-3 group. Lifespan analysis for females (C) and males (D) shows that the survival curves are significantly different among Haplotype-1, -2 and -3 groups (Log Rank test, $P = 0.0001$).
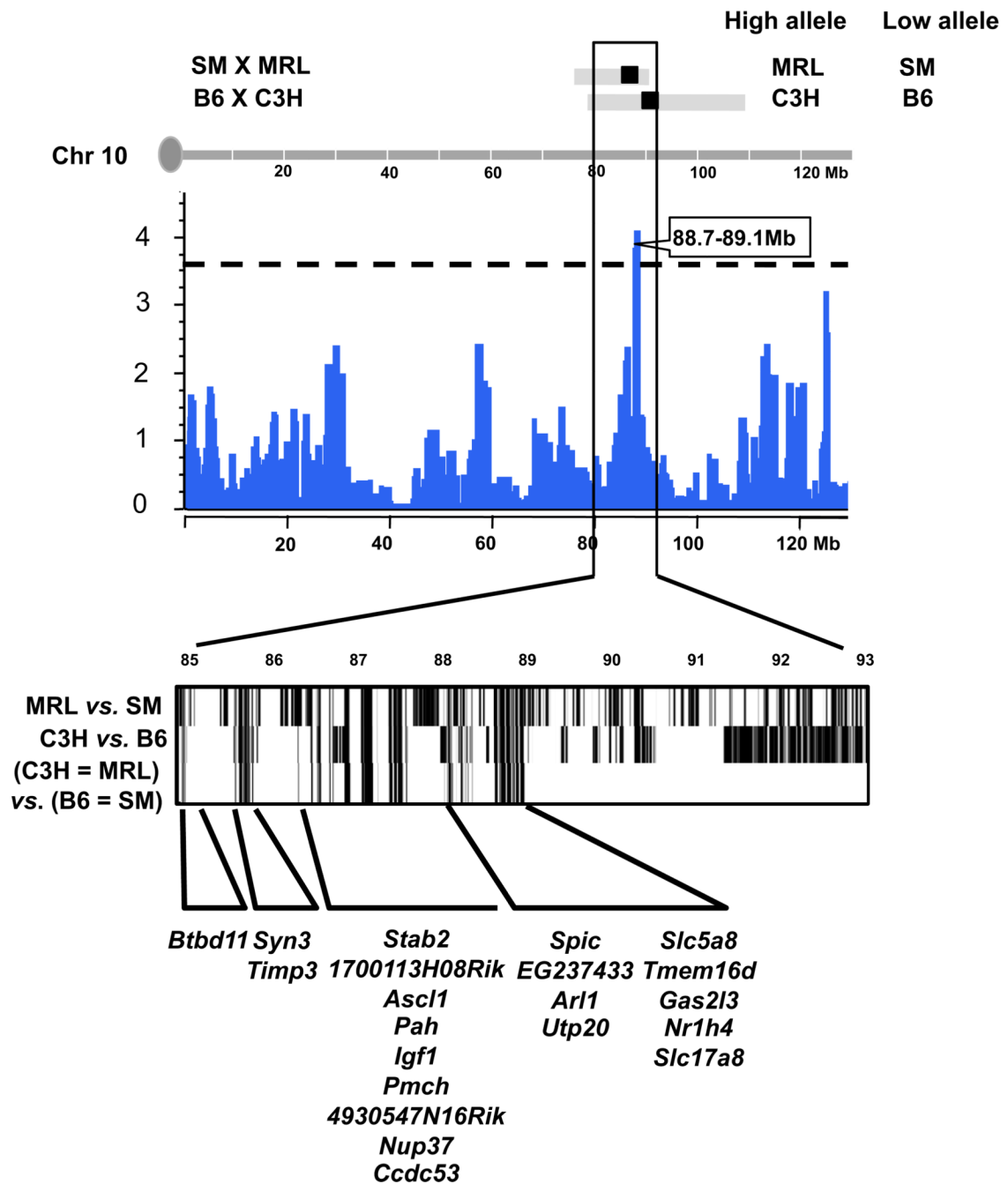
Fig. 5.
Using the bioinformatics tools to narrow the Chr 10 locus and identify candidate genes. The top part of the figure shows the overlap among QTL in the region of Chr 10. The middle of the figure shows the *P* values of the HAM analysis in the same region (the dashed line indicates the level of $P = 0.63$). In the lower part of the figure, black regions indicate those not identical by descent between strains. We combined the haplotypes as follows: we selected regions that were identical between the high allele strains (C3H and MRL), identical between the low allele strains (B6 and SM), and different from each other. The final candidate region contains 21 genes.

Table 1

IGF-1 levels (mean ± SEM) in MRL, SM, F1 and F2 mice

Mice	Males		Females	
	N	IGF-1 (ng/ml)	N	IGF-1 (ng/ml)
MRL	5	594 ± 10	5	626 ± 11
SM	5	234 ± 9[*]	5	153 ± 19[*]
(MRL×SM) F1	12	324 ± 30[*†]	5	483 ± 20[*†]
(SM×MRL) F1	5	525 ± 65[†‡]	5	489 ± 37[*†]
(MRL×SM) F2	229	409 ± 6	135	438 ± 10

[*]$P < 0.001$ vs. MRL

[†]$P < 0.001$ vs. SM

[‡]$P < 0.001$ vs. (MRL×SM) F1

Table 2

QTL in MRL × SM F2 mice and HAM QTL for IGF-1 levels of 6-month-old females and males

MRL×SM QTL	Chr	Peak (cM) (95% CI)[1]	Peak (Mb) (95% CI)[1]	LOD score (adjusted LOD)[2]	Closest marker	High strain; mode of inheritance
Igf1q3	9 [3,4]	26 (18–36)	48 (26–121)	2.6 (**4.1**)	rs13480173	M; recessive
Igf1q4	10 [3]	43 (42–44)	86 (85–88)	**31.8** (ns)	rs6394370	M; additive
Igf1q5	17 [3,4]	10 (8–58)	18 (12–89)	2.2 (**5.8**)	rs6384940	M; additive
Igf1q6	15 [5]	40 (22–46)	85 (55–91)	3.7 (3.2)	rs3681229	M; recessive

Sex	HAM QTL	Chr	Position (bp)[6] Start	End	Score (−log$_{10}$P)	SNP Start	End
Female	Igf1q7	2	153,602,815	153,638,312	3.67	2-153840428	rs6385706
	Igf1q8	10	88,735,006	89,031,602	4.12	rs29369977	rs29384008
	Igf1q9	17	28,830,413	28,853,776	3.77	17-27287107	17-273106092
Male	Igf1q10	1	158,911,588	159,733,488	5.14	1-156981448	1-157803580
	Igf1q11	3	52,448,663	52,821,781	4.05	NES10804633	rs6335414
	Igf1q12	3	67,787,738	69,548,776	4.44	3-52873658	rs6335414
	Igf1q13	16	74,123,988	74,157,921	3.84	16-74474798	16-74508912

Abbreviations: Chr, chromosome; CI, confidence interval; ns, non-significant; M, MRL/MpJ; S, SM/J.

[1] 95% CI. Genome-wide significance levels were determined by permuting the observed data 1000 times. CI was calculated with the Bayesian method.

[2] LOD scores were calculated with sex as an additive covariate except for the sex-specific QTL on Chr 15, where sex was added as an interactive covariate. Bold LOD score indicates significant QTL ($P < 0.05$); other LOD scores indicate suggestive QTL ($P < 0.63$). IGF-1 was further adjusted for the IGF-1 locus on Chr 10 by adding rs6394370 as an additive covariate. The LOD scores in parentheses indicate the LOD score for IGF-1 after adjustment.

[3] Main effect QTL.

[4] Significant QTL after adjustment for IGF-1 locus.

[5] Sex-specific (female) QTL (Δ LOD > 2).

[6] NCBI build 37.

Table 3

Polymorphisms within the *Igf-1* exons, UTRs and 3' region in the MRL, C3H, SM, B6 and D2 strains of inbred mice

Transcript (location)	SNP	Position	High-allele strains		Low-allele strains			Ref.
			MRL	C3H	B6	SM	D2	
ENSMUST00000062862 (5' UTR)	Indel	−1,036	-	-	-	-	(TGC)$_2$	new
ENSMUST00000121161	A/C	−86	A	A	C	C	A	rs29342496
ENSMUST00000062862 (5' UTR)								
3' region	Indel	+6,003	-	-	T	T	-	new
3' region	Indel	+6,034	-	-	T	T	-	new

All *Igf-1* exons were resequenced in addition to 3 Kb in the 3' region of the gene in all five strains. +1 position indicates the starting codon (ATG) of ENSMUST00000062862 for the SNPs located in the 5' UTR. +1 indicates the starting codon of ENSMUST00000095360 for the SNPs located in the 3' region.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 4

Expression evidence of the transcripts and probes within *Igf1*

Transcript or probe ID[*]	Exon tagged[*]	Gene expression difference, MRL vs. SM[†]: fold change and (*P* value)		QTL analysis[‡]	
		Females	Males	LOD[§]	High allele
Transcripts:					
ENSMUST00000075330	1b, 3, 4, 5a, 6d	+1.39 (0.002)	+1.19 (0.072)	0.4	-
ENSMUST00000095360	1b, 3, 4, 6d	+1.42 (0.001)	+1.16 (0.106)	0.3	-
ENSMUST00000062862	2a, 3, 4, 5b	+1.27 (0.024)	+1.10 (0.349)	0.7	-
Probes:					
1054540	1b	-	-	22.6	SM
317441	1b	-	-	20.8	SM
732594	1b	-	-	17.5	SM
817986	2	-	-	3.4	SM
533871	4	-	-	3.9	MRL
18947	6	-	-	3.2	MRL
231840	6d	-	-	3.2	MRL

[*] The probe IDs and the exons are referenced in Tables S4 and S5, Supporting Information.

[†] Fold changes are represented as the mean in the MRL mice compared to the mean in the SM mice in males and females separately. *P* values were calculated by ANOVA analysis and are indicated in parentheses for the transcripts. No calculations were performed for the probes.

[‡] The microarray and QTL analysis was performed on 282 liver samples from the 371 F2 mice. The results of the QTL analysis in this sub-population was similar to what we found in the QTL analysis for the entire cohort. QTL analysis in the 282 F2 mice confirmed the Chr 10 QTL at 44 cM (LOD = 25.2) (*Igf1q4*), which explain about 30% of the IGF-1 variation.

[§] LOD scores were calculated using the expression level as the phenotype. Sex was added as an additive covariate in the model. Max LOD scores are reported at the *Igf1* locus.

Table 5

Evidence for the 21 candidate genes at the Chr 10 locus

Gene	Mb*	Gene expression difference MRL vs SM† Fold change (P value)		Expression QTL analysis		Non-synonymous polymorphism
		Female	Male	LOD score‡	High allele	
Btbd11	84.9	−1.01 (0.940)	−1.21 (0.040)	-	-	-
Syn3	85.7	+1.21 (0.003)	+1.03 (0.550)	-	-	-
Timp3	85.7	−1.41 (0.004)	−1.25 (0.044)	18.1	SM	-
Stab2	86.3	+1.12 (0.201)	+1.17 (0.082)	53.2	MRL	-
1700113H08Rik	86.6	+1.15 (0.437)	−1.15 (0.439)	-	-	-
Ascl1	86.9	+1.17 (0.089)	−1.06 (0.508)	-	-	-
Pah	87.0	−1.09 (0.001)	−1.19 (0.056)	6.8	SM	-
Igf1	87.3	+1.39 (0.002)	+1.19 (0.072)	-	-	-
Pmch	87.5	+1.16 (0.151)	+1.00 (0.962)	-	-	-
4930547N16Rik	87.5	−1.04 (0.538)	−1.08 (0.229)	-	-	E152A
Nup37	87.6	−1.17 (0.007)	+1.03 (0.543)	-	-	-
Ccdc53	87.6	−1.05 (0.567)	+1.14 (0.188)	-	-	-
Spic	88.1	−1.07 (0.581)	+1.11 (0.410)	-	-	-
EG237433	88.1	+1.07 (0.583)	−1.15 (0.280)	-	-	-
Arl1	88.1	+1.02 (0.723)	−1.13 (0.052)	-	-	-
Utp20	88.2	−1.22 (0.032)	+1.10 (0.275)	-	-	-
Slc5a8	88.3	+1.10 (0.110)	−1.18 (0.016)	-	-	-
Tmem16d	88.5	+1.13 (0.125)	+1.02 (0.760)	-	-	-
Gas2l3	88.8	+1.03 (0.665)	+1.21 (0.665)	-	-	-
Nr1h4	88.9	−1.19 (0.005)	+1.10 (0.085)	-	-	E28D, K199R
Slc17a8	89.0	−2.52 (<0.001)	−2.11 (<0.001)	50.4	SM	-

*
Build 37.

†
Fold changes are given as MRL mice compared to SM mice in males and females separately. P values were calculated by ANOVA analysis and are indicated in parentheses for the transcripts. Some of the genes have more than one transcript. We report here the most significant transcript. All other transcripts are indicated in Table S6, Supporting Information. Numbers in bold indicate significant difference between MRL and SM (P < 0.05).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

‡ Results from QTL analysis for Chr 10 using the expression level of the transcript as the phenotype. The LOD score of the Chr 10 peak is indicated.