Haplotype Analysis in Multiple Crosses to Identify a QTL Gene

Xiaosong Wang,1 Ron Korstanje, David Higgins, and Beverly Paigen

The Jackson Laboratory, Bar Harbor, Maine 04609, USA

Identifying quantitative trait locus (QTL) genes is a challenging task. Herein, we report using a two-step process to identify Apo2 as the gene underlying Hdlq5, a QTL for plasma high-density lipoprotein cholesterol (HDL) levels on mouse chromosome 1. First, we performed a sequence analysis of the Apo2 coding region in 46 genetically diverse mouse strains and found five different APOA2 protein variants, which we named APOA2α to APOA2ε. Second, we conducted a haplotype analysis of the strains in 21 crosses that have so far detected HDL QTLs; we found that Hdlq5 was detected only in the nine crosses where one parent had the APOA2β protein variant characterized by an Ala61-to-Val61 substitution. We then found that strains with the APOA2β variant had significantly higher (P ≤ 0.002) plasma HDL levels than those with either the APOA2α or the APOA2ε variant. These findings support Apo2 as the underlying Hdlq5 gene and suggest the Apo2 polymorphisms responsible for the Hdlq5 phenotype. Therefore, haplotype analysis in multiple crosses can be used to support a candidate QTL gene.

[Supplemental material is available online at www.genome.org. The following individuals kindly provided reagents, samples, or unpublished information as indicated in the paper: D. Puppione.]

Most common human diseases, such as atherosclerosis, diabetes, and obesity, are complex traits determined by many genetic and environmental factors. The genetic factors are usually studied in animal models, most commonly mice, and frequently through a process known as quantitative trait locus (QTL) analysis, which has the advantage of finding novel key genes in a metabolic pathway. To date, more than 1800 mouse QTLs have been found (Mouse Genome Informatics, http://www.informatics.jax.org); however, identifying the genes underlying these QTLs has been an extremely challenging task (Nadeau and Frankel 2000; Korstanje and Paigen 2002).

The level of plasma high-density lipoprotein cholesterol (HDL), although not a disease, is also a complex trait. It has been intensely studied because, in humans, it is inversely correlated with the risks of coronary artery disease, and therapies that raise plasma HDL levels may significantly reduce these risks (Boden and Pearson 2000). QTL analysis has identified many genomic regions that regulate HDL levels, both in mice and in humans—to date, 37 mouse and 29 human HDL QTLs have been identified (update of review by Wang and Paigen 2002). One of the mouse HDL QTLs, Hdlq5 (Wang et al. 2003), on distal chromosome 1 (cM 92), has been repeatedly identified in nine of those crosses (update of review by Wang and Paigen 2002). An obvious candidate for the Hdlq5 gene was Apo2a (cM 92.6), because its encoded protein, apolipoprotein A-I (APOA2), helps maintain plasma HDL levels in mice (for review, see Blanco-Vaca et al. 2001).

To determine whether Apo2 was the Hdlq5 gene, we took the approach of haplotype analysis. Recent single nucleotide polymorphism (SNP) maps indicate that the genome of common inbred mouse strains is defined by 1–2 Mb haplotype blocks (Wade et al. 2002; Wiltshire et al. 2003), which can be used to narrow a QTL, because its underlying gene should be in subregions where the parental strains have different haplotypes (Park et al. 2003; Manenti et al. 2004). We extended this haplotype analysis further by using it to identify a QTL gene. In doing so, we performed a sequence analysis of the coding region of Apo2 in 46 genetically diverse mouse strains, and a haplotype analysis of strains both in the nine crosses that detected Hdlq5 and in the 12 crosses that failed to detect Hdlq5. We found that not only did the haplotype analysis support Apo2 as the Hdlq5 gene, but it also suggested that the Apo2 mutation is responsible for the Hdlq5 phenotype. Thus, haplotype analysis in multiple crosses can be used to support a QTL gene.

RESULTS

Inbred Mouse Strains Have Five APOA2 Protein Types

Polymorphisms of mouse Apo2 coding sequence have been reported before; a total of 15 SNPs cause eight amino acid changes in APOA2 (Doolittle et al. 1990; Higuchi et al. 1991; Purcell-Huynh et al. 1995; Suto et al. 1999; Kitagawa et al. 2003). For our purpose of testing Apo2 as the candidate gene for Hdlq5, we sequenced all four Apo2 exons in 46 genetically diverse and widely used inbred mouse strains, including 43 whose plasma HDL concentrations were known (Mouse Phenome Database, http://www.jax.org/phenome) and three that were parents in crosses to map HDL QTLs (CASA, MRL, and NZO). We found that the four exons of Apo2 had a total of 16 SNPs, resulting in nine amino acid changes and producing five APOA2 protein variants, which we named type “a” (APOA2α) to “e” (APOA2ε; see Table 1 and Supplemental Fig. 1). The “b” strains were distinct with regard to two features: Their APOA2 sequence had an Ala61-to-Val61 substitution, and none of them were wild-derived strains (except those having APOA2α), humans, chimpanzees, monkeys, horses, cattle, pigs, and rats (Fig. 1), suggesting that Val61 is a mutation in mice with APOA2ε.

Ala61 Is Conserved in Eight Mammalian Species

By comparing the APOA2 peptide sequences of eight mammal species, we found that APOA2 Ala61 was conserved in mice (except those having APOA2α), humans, chimpanzees, monkeys, horses, cattle, pigs, and rats (Fig. 1), suggesting that Val61 is a mutation in mice with APOA2ε.
Hdlq5 Was Only Detected in Crosses Whose Parental Strains Differed at APOA2 Amino Acid 61

Of the 21 different mouse crosses used to identify HDL QTLs, Hdlq5 was found only in the nine where one parent carried APOA2b and the second parent carried other APOA2 types (Table 2). Because mice with APOA2b have a unique Val61, our results clearly showed that Hdlq5 was found only when the two parental strains differed at amino acid 61, that is, one should carry Ala61 and the other should carry Val61. Hdlq5 may not be detected even though the two parental strains have different APOA2 types; for example, Hdlq5 was not found when C57BL/6 (having APOA2a) and SPRET (having APOA2e) were crossed (Table 2).

Haplotype analysis of the strains carrying APOA2 excluded the possibility that an unknown gene in linkage disequilibrium with Apoa2 was underlying Hdlq5 in these nine crosses: The APOA2b strains do not share a common haplotype block near Apoa2 (Table 3). The nearby SNPs reduced the region the strains shared to one containing only two genes, Apoa2 and Fcer1g (Fc receptor, IgE, high affinity I, polypeptide). There is no evidence that Fcer1g is involved in lipoprotein metabolism.

Mice with APOA2b Had Higher Plasma HDL Concentrations Than Did Those With Either APOA2a or APOA2c

To determine whether any of the APOA2 variants were associated with variant plasma HDL levels, we analyzed the plasma HDL concentrations of the 43 inbred mouse strains from the Mouse Phenome Database. In male mice, HDL levels of strains having APOA2b (103 ± 6 mg/dL, mean ± SEM, 12 strains) were significantly higher than those of strains having either APOA2a (67 ± 3 mg/dL, n = 24) or APOA2c (62 ± 4 mg/dL, n = 4) (P = 9 × 10⁻⁷ and P = 0.002, respectively; Fig. 2A). Similar results were found in female mice: HDL levels of strains having APOA2b (81 ± 8 mg/dL, n = 12) were significantly higher than those of strains having either APOA2a (52 ± 2 mg/dL, n = 24) or APOA2c (46 ± 2 mg/dL, n = 4) (P = 3 × 10⁻⁵ and P = 0.02, respectively; Fig. 2B).

The distinguishing feature of the APOA2b is the presence of the amino acid substitution of alanine to valine at position 61. HDL levels of strains having Val61 were significantly higher than those having Ala61 in both male and female mice (103 ± 6 vs. 69 ± 3 mg/dL, P = 3 × 10⁻⁶, and 81 ± 8 vs. 53 ± 3 mg/dL, P = 6 × 10⁻³, respectively; Fig. 2C, D).

DISCUSSION

Identifying the gene underlying a QTL is a challenging task. Whereas considerable success has been achieved in identifying genes responsible for Mendelian traits—more than 1400 genes for them have been found (Page et al. 2003), fewer than 50 have been identified for polygenic or quantitative traits (Glazier et al. 2002; Korstanje and Paigen 2002).

Although there is no “gold standard” for positively identifying a QTL gene, the Complex Trait Consortium recently suggested that an identified gene should meet more than one of the following eight criteria (Biola et al. 2003): (1) polymorphisms in either its coding or regulatory regions have been found; its function has been (2) linked to the quantitative trait being analyzed,
Table 2: Hdlq5 Was Only Detected When the Two Parental Strains in a Cross Differed at AA61 of APOA2

<table>
<thead>
<tr>
<th>Cross APOA2 Type</th>
<th>Cross APOA2 AA</th>
<th>Hdlq5 detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>B6 × b</td>
<td>Val61</td>
<td>Yes</td>
</tr>
<tr>
<td>B6 × c</td>
<td>Ala61</td>
<td>No</td>
</tr>
<tr>
<td>B6 × e</td>
<td>Ala61</td>
<td>No</td>
</tr>
<tr>
<td>b × c</td>
<td>Ala61</td>
<td>No</td>
</tr>
<tr>
<td>a × e</td>
<td>Ala61</td>
<td>No</td>
</tr>
<tr>
<td>c × c</td>
<td>Ala61</td>
<td>No</td>
</tr>
</tbody>
</table>

Examples: B6 × H11538

of between 10 and 40 mice (24 CASA, MRL, and NZO; Mouse Phenome Database). Each group consisted of males (B) fasted for 4 h (all the strains in Table 1 except (mean APOA2 positions (see Table 1). Plasma HDL concentrations were compared. Second, among the 21 strains with APOA2b had significantly higher plasma HDL levels than those with either APOA2a or APOA2c. Second, among the 21 crosses used to map HDL QTL, Hdlq5 was only detected in the nine where the parental strains had different amino acid at position 61 in APOA2 protein. We excluded the possibility that a gene in linkage disequilibrium with Hdlq5 was underlying Hdlq5 because Fcer1g, the only gene that was in linkage disequilibrium with APOA2, has no known functions in lipoprotein metabolism. Haplotype analysis enabled us not only to identify Apoa2 as the gene for Hdlq5 but also to pinpoint the Ala61-to-Val61 substitution as the possible causal change for this QTL. First, mice with Val61 had significantly higher HDL levels than did those with Ala61. Second, Hdlq5 was only found in the crosses in which one parental strain had Val61 and the other had Ala61 in APOA2. Third, in the nine crosses that detected Hdlq5, the strain with Val61 always had the high allele of this QTL. This method may not be limited to finding a QTL gene with amino acid differences among the parental strains; it may also be used to find the key SNPs in the regulatory elements of a candidate gene.

Although Apoa2 is mostly likely the gene underlying Hdlq5, and Ala61-to-Val61 substitution in Apoa2 is most probably the causal polymorphism, other possibilities exist. First, polymorphisms in the promoter region of Apoa2 may lead to mRNA differences and thereby differences in plasma HDL levels. The follow-
ing results, however, suggest that there is no correlation between Apoa2 mRNA and plasma HDL levels and that Apoa2 promoter polymorphisms are unlikely to affect plasma HDL levels. In the B6xNZB cross in which Hdlq5 was found, B6 mice (having Apoa2b) have higher levels of Apoa2 mRNA but lower levels of plasma HDL levels, compared with NZB mice (having Apoa2a) (P < 0.05) (Wang et al. 2003). In the B6xC3H cross in which Hdlq5 was found, B6 mice had higher levels of both Apoa2 mRNA but lower levels of plasma HDL levels, compared with C3H mice (having Apoa2b) (P < 0.05; X. Wang and B. Paigen, unpubl.). On the other hand, B6 mice have similar levels of Apoa2 mRNA but lower levels of plasma HDL levels, compared with BALB/c mice (Doolittle et al. 1990). Second, a polymorphic enhancer in the same haplotype as Apoa2 coding sequence may regulate another gene, and this gene could regulate HDL levels. Although unlikely, we cannot exclude this possibility.

It has been reported that compared with C57BL/6 mice (having Apoa2b), BALB/c mice (having Apoa2a) have similar liver Apoa2 mRNA levels, but higher Apoa2 protein synthesis rate in hepatocytes (Doolittle et al. 1990). This suggests that a change in messenger sequence, presumably the nucleotide change that causes Ala61-to-Val61 substitution, increases the efficiency of Apoa2 messenger translation, leading to more rapid translation of the Apoa2 mRNA.

### Table 3. Haplotype Analysis of Apoa2 Region in Mice with APOA2b

<table>
<thead>
<tr>
<th>Marker</th>
<th>Gene</th>
<th>Start (bp)</th>
<th>End (bp)</th>
<th>129 BALB</th>
<th>C3H</th>
<th>CBA</th>
<th>FVB</th>
<th>KK</th>
<th>LP</th>
<th>NON</th>
<th>NZB</th>
<th>NZW</th>
<th>RF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wm3</td>
<td>Nr1i3</td>
<td>175048841</td>
<td>175048962</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Wm1</td>
<td>Apoa2</td>
<td>175081943</td>
<td>175082139</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Wr1</td>
<td>Fcer1g</td>
<td>175087300</td>
<td>175092009</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Wrs3</td>
<td>Ndufs2</td>
<td>175092583</td>
<td>175104835</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

1Microsatellite markers that we designed. We used PCR to amplify them. PCR products were scored according their sizes, with 1 < 2 < 3 in size. The sequences of the PCR primers for these markers are: Wm3: forward: 5'-TGCAGCATTTTCTTGTTG-3'; reverse: 5'-AAGGAATGGGGGGTGATTAGAAG-3'; Wm1: forward: 5'-TCATATACCTCAGATGCCACATGC-3'; Wrs3: forward: 5'-CCAGATTTGGTTGGGAG-3'; Wrs3: forward: 5'-AGTCCCTCTCATCACCC-3'; reverse: 5'-CACCCCAAGTCTCCTGTC-3'.

2Genes that are closest to Apoa2, according to Ensembl Mouse Genome Server (http://www.ensembl.org/Mus_musculus/), Build 32.

3Positions were retrieved from Ensembl Mouse Genome Server (http://www.ensembl.org/Mus_musculus), Build 32. The distance from the neighboring genes to Apoa2: Nr1i3: 6244 bp; Fcer1g: 3207 bp; Ndufs2: 8490 bp.
production of APOA2 protein and therefore larger HDL particle and higher plasma HDL concentrations. We found no sequence changes in the 3′- or 5′-UTR regions between C57BL/6 and BALB/c mice (Supplemental Fig. 1).

Our findings are important for future HDL research for two reasons. First, to avoid Hdp5p, whose strength may mask other HDL QTLs (and hence their discovery), future HDL QTL mapping efforts in mice should cross parental strains with the same amino acid 61 in APOA2. Second, a human QTL for plasma HDL levels was identified in the APOA2 region (Elbein and Hasstedt 2002). To determine whether APOA2 regulates human HDL concentrations, APOA2 SNPs should be analyzed in human populations.

Obviously, the analysis we conducted will be more powerful when the same QTL is detected in multiple crosses. Now that about 1500 QTLs, many of which were detected in multiple mouse crosses, have been mapped in the mouse genome (http://informatics.jax.org) and a wide variety of phenotypic data from many commonly used and genetically diverse mouse strains are easily accessible in the Mouse Phenome Database (http://www.jax.org/phenome), many QTL genes can be discovered quickly by analyzing their SNPs and haplotypes in multiple crosses, as exemplified in this study.

METHODS

Obtaining and Genotyping Genomic DNA

We obtained genomic DNA from Mouse DNA Resource at the Jackson Laboratory (http://www.jax.org/dnares/index.html). PCR genotyping of the four microsatellite markers (Wm3, Wm1, Wrs1, and Wrs3) in the APOA2 region was carried out for 35 cycles under the following conditions: 94°C, 30 sec; 55°C, 30 sec; 72°C, 1 min. Polymorphisms were detected by electrophoresing the PCR products on 4% Nusieve 3:1 agarose gels in 1 X Tris-borate–EDTA running buffer for 2 h at 190 volts. Gels were then stained with ethidium bromide and photographed under ultraviolet light.

Sequencing APOA2

To define APOA2 haplotypes, we sequenced the four exons of mouse APOA2. We amplified each of the four exons with PCR with the same protocol as shown above except using an annealing temperature of 59°C instead of 55°C, and checked the product sizes on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced using Big Dye Terminator Cycle Sequencing Chemistries on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced with the same protocol as shown above except using an annealing temperature of 59°C instead of 55°C, and checked the product sizes on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced using Big Dye Terminator Cycle Sequencing Chemistries on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced with the same protocol as shown above except using an annealing temperature of 59°C instead of 55°C, and checked the product sizes on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced using Big Dye Terminator Cycle Sequencing Chemistries on 4% Nusieve 3:1 agarose gels.

Sequencing APOA2

To define APOA2 haplotypes, we sequenced the four exons of mouse APOA2. We amplified each of the four exons with PCR with the same protocol as shown above except using an annealing temperature of 59°C instead of 55°C, and checked the product sizes on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced using Big Dye Terminator Cycle Sequencing Chemistries on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced with the same protocol as shown above except using an annealing temperature of 59°C instead of 55°C, and checked the product sizes on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced using Big Dye Terminator Cycle Sequencing Chemistries on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced with the same protocol as shown above except using an annealing temperature of 59°C instead of 55°C, and checked the product sizes on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced using Big Dye Terminator Cycle Sequencing Chemistries on 4% Nusieve 3:1 agarose gels.

Sequencing APOA2

To define APOA2 haplotypes, we sequenced the four exons of mouse APOA2. We amplified each of the four exons with PCR with the same protocol as shown above except using an annealing temperature of 59°C instead of 55°C, and checked the product sizes on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced using Big Dye Terminator Cycle Sequencing Chemistries on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced with the same protocol as shown above except using an annealing temperature of 59°C instead of 55°C, and checked the product sizes on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced using Big Dye Terminator Cycle Sequencing Chemistries on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced with the same protocol as shown above except using an annealing temperature of 59°C instead of 55°C, and checked the product sizes on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced using Big Dye Terminator Cycle Sequencing Chemistries on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced with the same protocol as shown above except using an annealing temperature of 59°C instead of 55°C, and checked the product sizes on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced using Big Dye Terminator Cycle Sequencing Chemistries on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced with the same protocol as shown above except using an annealing temperature of 59°C instead of 55°C, and checked the product sizes on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced using Big Dye Terminator Cycle Sequencing Chemistries on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced with the same protocol as shown above except using an annealing temperature of 59°C instead of 55°C, and checked the product sizes on 4% Nusieve 3:1 agarose gels. The PCR products were sequenced using Big Dye Termina


Received April 9, 2004; accepted in revised form July 7, 2004.