Fine Mapping and Identification of Candidate Pulmonary Adenoma Susceptibility 1 Genes Using Advanced Intercross Lines

Min Wang, William J. Lemon, Gongjie Liu, Yan Wang, Fuad A. Iraqi, Alvin M. Malkinson, and Ming You

Division of Human Cancer Genetics, The Ohio State University Comprehensive Cancer Center, Columbus, Ohio 43210 [W. G. L., M. Y.]; Department of Surgery and The Alvin J. Siteman Cancer Center, Washington University School of Medicine, St. Louis, Missouri 63110 [W. J. L., Y. W., M. Y.]; International Livestock Research Institute, Nairobi, Kenya [F. A. I.]; and Department of Pharmaceutical Sciences, University of Colorado Cancer Center and Health Sciences Center, Denver, Colorado 80262 [A. M. M.]

ABSTRACT

In the present study, we used newly developed F11 generation mouse advanced intercross lines (AIL) to fine map Pas1–3 quantitative trait loci (QTL). The (A/J × C57BL/6) F1, AIL mouse population was created by crossing lung tumor-resistant C57BL/6 mice with lung tumor-susceptible A/J mice. By selectively genotyping 30% of the population, we have confirmed the Pas1 QTL and narrowed it to an interval of ~1.0 cM or 1.3 Mb in the vicinity of the Kras2 gene. The Pas2 QTL was detected by both ANOVA and regression analysis but not by MapMaker EXP/QTL software. In addition, an interaction between the Pas1 and Pas2 QTLs was revealed. However, the Pas3 QTL was not confirmed in this study. It was either lost during the development of the AIL or too weak to be detected using AIL. The Pas1 locus is now sufficiently fine-mapped that candidate gene(s) for the Pas1 locus can be characterized by positional cloning. In this study, all 27 of the known or predicted genes located in the Pas1 candidate region were characterized as possible candidate Pas1 genes. Six genes were selected for additional analyses because of their relevant function in tumorigenesis or allelic changes between A/J and C57BL/6 mice. The Lrmp gene bears amino acid polymorphisms among various mouse strains that are highly correlated with the Pas1 allele status. The Pas1e1 gene (RIKEN Ak016641), encoding an intermediate filament tail domain-containing protein, produces alternatively spliced transcripts across inbred strains of mice, and its splicing pattern cosegregates with the Pas1 allele. The genetic and expression data support these two genes as strong candidates for the Pas1 locus. Of the other four genes (Eca39, RIKEN Ak015530, mHoj-1, and Krag), no functional polymorphisms or differential gene expression were found in Eca39, mHoj-1, and Krag between lung tumor-susceptible and -resistant strains. The Ak015530 carries an amino acid polymorphism, but this polymorphism does not cosegregate with mouse lung tumor susceptibility. Thus, these 4 genes are less likely candidates for the Pas1 locus.

INTRODUCTION

Epidemiological studies have indicated that ~85% of all lung cancer deaths in the United States are associated with tobacco smoking (1). Tobacco smoking increases the relative risk for lung cancer in smokers by 13-fold and in passive smokers by 1.5-fold (2). Although the majority of lung cancer cases are associated with cigarette smoking, increasing evidence suggests that individuals differ in their susceptibility to lung cancer. An increased familial risk for lung cancer was observed among relatives of lung cancer probands (3–5). Additional segregation analyses provided evidence that susceptibility to human lung cancer follows a pattern of autosomal-dominant Mendelian inheritance (3). To date, there have been no reports on the localization and identification of human lung tumor susceptibility genes.

Different inbred mouse strains show widely different susceptibilities to both spontaneous and chemically induced lung tumor formation, and, thus, serve as models for research in lung cancer genetics (6). The multiplicity of mouse lung tumors are a quantitative trait controlled by multiple genetic loci (7). Recent linkage studies have been conducted to identify Pas3 and Par loci. A major susceptibility locus was mapped in (A/J x C3H/HeJ) F2 mice to distal chromosome 6, and was termed the Pas1 locus (8). This locus produced a maximum LOD score of 9. Consistent results were obtained in comprehensive linkage studies using (A/J × C57BL/6) F2, (A/J × C57BL/6J) × C57BL/6J, (A/J × Mus Spretus) × C57BL/6J, and A × B, and B × A recombinant inbred mice (9–12). Additional loci shown to modulate the effect of Pas1 were mapped to chromosomes 17 and 19. Linkage to a locus on chromosome 17, the site of the putative Pas2 locus, was observed in (A/J × C57BL/6) F2 mice (9). The location of the Pas2 locus is homologous to human chromosome 6p21; potential candidates at this location are the genes for tumor necrosis factors α and β. Similarly, linkages to lung tumor susceptibility were also seen at markers on chromosome 19 (Pas3) using (A/J × C57BL/6J) × C57BL/6J mice and (A/J × C57BL/6) F2 mice (10). Other potential lung tumor susceptibility loci and their complex interactions have been identified using recombinant congenic strains (13–16).

Most QTL detection and mapping studies in mice have been carried out using F2 intercross and backcross designs. Because of limited numbers of recombination events in small chromosomal regions, the mapping resolution has been restricted to relatively large CI. However, precise localization of QTL is required for positional cloning. To address this, a number of population designs have been proposed that increase recombination frequency in QTL-segregating populations (17). The AIL design is one approach (18). An advantage of this design over congenic approaches is its suitability for simultaneous refinement of multiple loci (19). In addition, the AIL design lends itself to situations where information on the number of QTLs in a particular region is unavailable, because data obtained with AILs potentially allow linked QTLs to be dissected into constituent loci (19).

In this report, we applied the AIL design of higher resolution mapping to additionally characterize Pas1–3 using the F1 generation of the (A × C57BL/6) AIL mouse population. Our objective was to fine map the Pas QTLs to a sufficiently small size (<1 cM) for candidate identification. Our results revealed that the Lrmp and Ak016641 genes, located in the now more refined Pas1 candidate region, bear amino acid polymorphisms. Ak016641 also produces alternatively spliced transcripts. These amino acid polymorphisms and the Ak016641 strain transcript pattern highly cosegregate with the mouse Pas1 allele, providing genetic evidence for these two genes as strong Pas1 candidates.
MATERIALS AND METHODS

AILs

The F11 generation of the (A/J × C57BL/6) AIL mice developed at the International Livestock Research Institute (Nairobi, Kenya) was used in the present study. Original parental strains, C57BL/6JOlaHsd and A/JOlaHsd (Harlan United Kingdom Ltd., Oxon, United Kingdom), produced >35 F1 litters. F2 mice were generated by crossing 50 pairs of F1 mice; there were no duplicate matings. Sibling pairing was avoided, but otherwise pairing was random. For each generation after F2 to F10, 50 litters were produced, and 65 breeding pairs selected to produce the following generation. The same strategy that was used in producing the F2 was adopted in all of the subsequent generations. More than 200 pairs of F10 were selected to produce ~1120 F11 mice.

Lung Tumorigenesis Study

The F11 AIL mice were used in a mouse lung tumor bioassay. Four to 7-week-old mice received a single i.p. injection of urethane (1 g/kg), and 6 months later, mice were sacrificed by CO2 asphyxiation, and their lungs removed and fixed in 10% buffered formalin. Surface lung tumors were counted with the use of a dissecting microscope.

Genotype Data Collection

In total, 1120 (A/J × C57BL/6) F11 AIL mice were used in the lung tumor bioassay. After counting lung tumors, high molecular weight genomic DNA of F11, and parental strain mice was prepared from tails. A tail clipping from each F11 mouse was homogenized and incubated overnight at 37°C. The DNA was isolated using TRIzol reagent according to the manufacturer’s protocol (Life Technologies, Inc., Gaithersburg, MD). The quality of the isolated RNA was assessed by absorbance at 260 nm, the A260/A280 ratio, and electrophoresis on 1% agarose/formaldehyde gels that indicated the intensity and integrity of the 28S and 18S bands. Two μg of total RNA was used in a reverse-transcription reaction to synthesize the first strand cDNA using oligoethyleneoxythymidylic acid primer.

The mRNA sequences of Eca39, Lmp1, Ak015530, Ak016641, and Krug genes were retrieved from the National Center for Biotechnology Information GenBank (GI: 6680771, 6678713, 12853911, 12855487, and 437199, respectively). The mRNA sequence of mHoj-1 was derived through blasting its human homologue HOJ-1 (GI: 4009349) against the RIKEN and Celera mouse genome databases. The mouse mRNA sequence was confirmed by RT-PCR analysis and has been deposited in GenBank (accession no. AF241572; GI: 23338217).

For coding region nucleotide polymorphism analyses, primers were designed based on exon-flanking sequences. Using mouse lung DNA, PCR reactions were performed to amplify coding sequence fragments for each gene. PCR products were resolved on 1.2% ethidium bromide-stained agarose gels, and purified using QiAquick gel extraction kits (Qiagen, Hilden, Germany). Automated sequencing was performed using dyeoxy terminator cycle sequencing kits (Applied Biosystems) and Applied Biosystems model 377 DNA sequencers (Perkin-Elmer, Foster City, CA). Both directions were sequenced for each fragment to assure sequence fidelity.

For gene expression analyses, the following specific primer sets were designed: Eca39, forward 5’-ATG AAC GAG TGC AGT AAT GG-3’, reverse 5’-AGT TCC ACA GCG TAT GGC ACG-3’; Lmp1, forward 5’-AAG AGG GTG AAC CTT GAA GAG G-3’, reverse 5’-TCT AAC CTC TCG CTA GC-3’; Ak015530, forward 5’-GCC AAT TCG TTA CGA GGA G-3’, reverse 5’-TCA GAC TTT GGT ATC TGA ATA G-3’; Ak016641, forward 5’-GGA AGT AGA GAT TGG AAA CCA C-3’, reverse 5’-CAA CTC TAC AAT AGT TAC TCA GTA G-3’; mHoj-1, forward 5’-TCT GAA GCC AGC GCC ATG AC-3’, reverse 5’-CAC CTG CTC GAT CAA AGC-3’; Krag, forward 5’-TAC CAA GTT GAC GAG AGC-3’, reverse 5’-TTA GGC CTT TGG TAG CTG GC-3’. For β-actin, forward 5’-TGA CAT CCG TAA AGC AGT CTA TGC C-3’, reverse 5’-AAG CAC TCT CGG TGC TGC ACG ATG GAG-3’. The linear amplification region for each gene was predetermined by plot experiments. The reaction profile generally consisted of one cycle at 95°C for 2 min, followed by 30 cycles at 94°C, 55°C, and 72°C each for 30 s. Twenty-two cycles of amplification was used for the β-actin control. Thirty cycles of amplification was used for both the Eca39 and Ak015530 genes, and 32 cycles and 1 min extension time for the Ak016641 gene. The PCR products were resolved in 1.2% ethidium bromide-stained agarose gels and visualized under UV.

RESULTS

Fine Mapping of Pas QTLs. Fig. 1 shows results from the MapMaker EXP/QTL analysis with markers placed according to their physical position based on the Celera CDS 3.6 database. For the Pas1 QTL, a peak was found between the D6Osus6 and D6Mit294 markers with the highest LOD score at D6Mit7 (Fig. 1a). Extremely high LOD scores (>150) were observed, which can be attributed to the AIL breeding method, a large gene effect of Pas1 QTL, and/or the selective genotyping strategy. Iraqui et al. (19) also found large LOD scores in some regions when using AIL to fine map trypansomiasis resistance genes. They used a 2-LOD supporting interval with “95% confidence” citing this as conservative estimation (19). However, the use of 1- or 2-LOD supporting intervals as CIs was found not highly informative in the present study, because the interval would have been placed very close to a single marker, i.e., D6Mit57. ANOVA analysis of Pas1 QTL has shown that whereas the markers proximal to D6Mit57 are highly linked, D6Mit15 is also associated with mouse lung tumor susceptibility (Table 1), revealing some complexity within the Pas1 QTL region. Finally, single term linear regression analysis revealed that only the D6Mit57 marker was significantly associated with lung tumor susceptibility, suggesting that the CI for Pas1 could be conservatively defined by the D6Osus6 and D6Mit294 markers (Fig. 2a). In the Celera mouse genome map, these two markers are...
adjacent to Eca39 and Krag genes, respectively. The genetic distance between these two markers is <1 cM according to Mouse Genome Database, and the physical distance is ~1.3 Mb.

For the Pas2 QTL, inclusion of all of the animals showed only an insignificant peak by MapMaker Exp/QTL analysis. However, when only mice heterozygous at D6Osu6 were included, the peak near D17Mit16 exceeded LOD = 2, suggesting interaction between the Pas1 and Pas2 QTLs, and that dominance of Pas1 could mask the Pas2 effect (Fig. 1b). This interacting relationship has been reinforced by both ANOVA (Table 1) and regression result (Fig. 2b) analyses. The Pas2 QTL lies in the 7.6 Mb region between D17Mit23 and D17Mit231 with the most likely candidate region localized at D17Mit16 and not extending to D17Mit23.

No marker at Pas3 was linked to lung tumor susceptibility (Fig. 1c; Table 1; Fig. 2c). It is possible that high-resolution AIL design missed the Pas3 gene because of too large intervals of genotyping, such as between D19Mit132 and D19Mit19. However, the fact that we did not detect any significant linkage at D19Mit42 and D19Mit19 markers (significantly linked to Pas3 in ref. 10) suggests that the Pas3 locus either is not present or has a very small genetic effect. The highly non-Mendelian distribution of Pas3 alleles (1:4:4 on average) suggests that the selection process produced selective pressure against the AA genotype. It is not clear how to interpret this pressure, because the entire region surveyed was affected.

**Coding Region Nucleotide Polymorphism Analyses**. In the newly refined Pas1 region, there are ~27 known or predicted genes. After initial screening using oligonucleotide arrays (20) and computer-assisted single nucleotide polymorphism analysis (21), we select 6 mostly likely candidates (Lrmp, RIKEN Ak016641, RIKEN Ak015530, Eca39, mHoj-1, and Krag) for additional analyses because of their relevant function in tumorigenesis or allelic changes between A/J and C57BL/6 mice. The physical position of each gene relative to K-ras is indicated in Fig. 3. The coding regions of the Lrmp gene in A/J and C57BL/6J mice were first amplified by RT-PCR and directly sequenced. By comparing the sequences from these two distinct strains, we found a total of eight nucleotide polymorphisms located at codons 31, 56, 58, 60, 243, 343, 438, and 537 (Fig. 4). Among them, five polymorphisms are missense polymorphisms that give rise to amino acid alterations, namely, codons 31 (GA/GC), 56 (GG/AC), 58 (TTC/G), 438 (A/GGG), and 537 (CC/TG; Fig. 4). At codon 31, the hydrophilic, negatively charged aspartic acid in A/J mice has changed to hydrophobic neutral glycine in C57BL/6J mice. At codon 56, glycine in A/J mice is substituted by aspartic acid in C57BL/6J mice. At codon 58, phenylalanine in A/J mice has been substituted by leucine in C57BL/6J mice. Codon 438 encodes a hydrophobic, positively charged amino acid, arginine, in A/J mice and a hydrophobic, neutral amino acid, glycine, in C57BL/6J mice. Finally, codon 537 is located on the COOH terminus, which encodes for proline in A/J mice and leucine in C57BL/6J mice. The strain distribution pattern for these amino acid polymorphisms was established by sequencing PCR fragments amplified from different strains of mouse DNA. We found a high correlation between these polymorphisms and the Pas1 allele status (Table 2). Strain SM/J is resistant to lung tumor development, possibly because it carries multiple Pas1 candidate genes.

Table 2 lists amino acid substitutions in the Ak016641 and Ak015530 genes. For the Ak016641 gene, we found two amino acid-changing polymorphisms at codons 218 and 258 in the entire coding region. These polymorphisms and its mRNA transcript pattern (see below) are highly associated with strain Pas1 allele status. Because the Ak016641 gene is a strong Pas1 candidate, we rename it the Pas1 candidate 1 gene (Pas1c1). The open reading frame of gene Ak015530 (from bp123 to bp383) carries one polymorphism at codon 28; A/J strain carries GGC (Gly), whereas C57BL/6J strain has GAC (Asp). Strain distribution patterns show that Pas1-susceptible strains 129/SvJ and BALB/cJ exhibit GGC at this codon, similar to A/J, but SWR/J and CBA/J carry GAC at this codon as C57BL/6J and other lung tumor-resistant strains. For three other genes,
Eca39, mHoj-1, and Krag, no amino acid-changing polymorphisms were detected in their coding sequences.

**Gene Expression Analyses.** Semiquantitative RT-PCR was used to evaluate the relative expression level of each of the above 6 genes across several inbred strains (Fig. 5). Lrmp is expressed in mouse lung tissues without significant difference between lung tumor-susceptible and -resistant strains. This result is not consistent with a previous report (23), which showed by Northern blotting that Lrmp is not expressed in mouse or human lung tissues. This discrepancy is most likely because the use of RT-PCR, which is a more sensitive technique than Northern blotting.

Eca39, Ak015530, mHoj-1, and Krag are expressed in similar amounts in lung tissues from all strains of mice examined.

When amplifying its COOH-terminal coding region, the Pas1c1 gene exhibited different isoforms of transcripts in A/J and C57BL/6J strains of mice. A/J mice carried two Pas1c1 transcripts, whereas C57BL/6J mice carried only the large transcript. Additional sequencing disclosed that one exon (from bp 894 to bp 1013, encoding 40 amino acids) is spliced out of the smaller transcript without changing the entire open reading frame (Fig. 6a). RT-PCR data on the various inbred strains revealed that the mRNA splicing pattern of each strain highly cosegregated with Pas1 allele status (Fig. 5).

<table>
<thead>
<tr>
<th>QTLs</th>
<th>Markers</th>
<th>AA</th>
<th>n</th>
<th>AB</th>
<th>n</th>
<th>BB</th>
<th>n</th>
<th>Main Effect</th>
<th>× D6Osu6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Pas1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D6Mit59</td>
<td>5.4</td>
<td>107</td>
<td>3.8</td>
<td>163</td>
<td>1.7</td>
<td>129</td>
<td>&lt;2.2e-16*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pklk3c2g</td>
<td>5.5</td>
<td>108</td>
<td>3.9</td>
<td>153</td>
<td>1.7</td>
<td>138</td>
<td>&lt;2.2e-16*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sox5</td>
<td>5.8</td>
<td>102</td>
<td>4.4</td>
<td>130</td>
<td>1.6</td>
<td>167</td>
<td>&lt;2.2e-16*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D6Osu6</td>
<td>5.8</td>
<td>109</td>
<td>4.4</td>
<td>133</td>
<td>1.3</td>
<td>157</td>
<td>&lt;2.2e-16*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D6Mit57</td>
<td>5.7</td>
<td>120</td>
<td>4.3</td>
<td>124</td>
<td>1.3</td>
<td>155</td>
<td>&lt;2.2e-16*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D6Mit294</td>
<td>5.7</td>
<td>116</td>
<td>4.3</td>
<td>133</td>
<td>1.3</td>
<td>150</td>
<td>0.106</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D6Mit15</td>
<td>5.7</td>
<td>117</td>
<td>4.3</td>
<td>132</td>
<td>1.3</td>
<td>150</td>
<td>0.005*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D6Mit201</td>
<td>5.7</td>
<td>118</td>
<td>4.3</td>
<td>131</td>
<td>1.3</td>
<td>150</td>
<td>0.159</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D6Mit373</td>
<td>5.7</td>
<td>119</td>
<td>4.2</td>
<td>131</td>
<td>1.3</td>
<td>149</td>
<td>0.948</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D6Mit26</td>
<td>5.7</td>
<td>106</td>
<td>4.4</td>
<td>141</td>
<td>1.3</td>
<td>152</td>
<td>0.745</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D6Mit372</td>
<td>5.7</td>
<td>105</td>
<td>4.4</td>
<td>142</td>
<td>1.3</td>
<td>152</td>
<td>0.075</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D6Mit390</td>
<td>5.7</td>
<td>105</td>
<td>4.3</td>
<td>145</td>
<td>1.3</td>
<td>149</td>
<td>0.306</td>
<td></td>
</tr>
<tr>
<td><strong>Pas2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D17Mit23</td>
<td>3.9</td>
<td>128</td>
<td>3.4</td>
<td>193</td>
<td>3.3</td>
<td>78</td>
<td>1.6e-06*</td>
<td>0.003*</td>
</tr>
<tr>
<td></td>
<td>D17Mit16</td>
<td>4</td>
<td>123</td>
<td>3.4</td>
<td>173</td>
<td>3.4</td>
<td>103</td>
<td>0.489</td>
<td>0.007*</td>
</tr>
<tr>
<td></td>
<td>D17Mit231</td>
<td>4</td>
<td>109</td>
<td>3.4</td>
<td>205</td>
<td>3.4</td>
<td>85</td>
<td>0.495</td>
<td>0.679</td>
</tr>
<tr>
<td></td>
<td>D17Mit11</td>
<td>3.7</td>
<td>138</td>
<td>3.5</td>
<td>188</td>
<td>3.5</td>
<td>73</td>
<td>0.018</td>
<td>0.747</td>
</tr>
<tr>
<td></td>
<td>D17Mit10</td>
<td>3.5</td>
<td>160</td>
<td>3.6</td>
<td>159</td>
<td>3.7</td>
<td>80</td>
<td>0.254</td>
<td>0.092</td>
</tr>
<tr>
<td><strong>Pas3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D19Mit42</td>
<td>3.2</td>
<td>39</td>
<td>3.4</td>
<td>199</td>
<td>3.8</td>
<td>161</td>
<td>0.687</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D19Mit86</td>
<td>3.2</td>
<td>46</td>
<td>3.4</td>
<td>212</td>
<td>3.8</td>
<td>141</td>
<td>0.274</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D19Mit152</td>
<td>3.4</td>
<td>22</td>
<td>3.4</td>
<td>182</td>
<td>3.8</td>
<td>195</td>
<td>0.975</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D19Mit19</td>
<td>3.6</td>
<td>44</td>
<td>3.7</td>
<td>176</td>
<td>3.5</td>
<td>179</td>
<td>0.795</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D19Mit69</td>
<td>3.3</td>
<td>50</td>
<td>3.6</td>
<td>182</td>
<td>3.6</td>
<td>167</td>
<td>0.797</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D19Mit53</td>
<td>3.3</td>
<td>54</td>
<td>3.7</td>
<td>149</td>
<td>3.5</td>
<td>196</td>
<td>0.500</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D19Mit10</td>
<td>3.3</td>
<td>56</td>
<td>3.8</td>
<td>174</td>
<td>3.4</td>
<td>169</td>
<td>0.192</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D19Mit37</td>
<td>3.3</td>
<td>63</td>
<td>3.6</td>
<td>210</td>
<td>3.6</td>
<td>126</td>
<td>0.093</td>
<td></td>
</tr>
</tbody>
</table>

*Indicates the marker is significant in the analysis.
Three nucleotide polymorphisms were identified in the skipped exon by sequencing; i.e., in codons 258, 266, and 269 (Fig. 6b). Codon 258 encodes for glycine in A/J mice and glutamate in C57BL/6J. Codons 266 and 269 carry silent polymorphisms. These three polymorphisms correlated closely with the transcript splicing pattern of each strain (Table 2), suggesting that they might be associated with mRNA alternative splicing.

DISCUSSION

There are at least three lines of evidence to support the significance of identifying the Pas1 gene(s). Firstly, Pas1 is a major susceptibility locus based on the genetic linkage studies in several crosses as reviewed in the previous section (8–12). This locus accounts for ~50% of the observed phenotypic variance indicating that Pas1 plays a major role in inherited predisposition to lung tumor development in mice (8–12). Secondly, there is a significant association between a Kras2/RsaI polymorphism (located within the Pas1 locus), and the risk and prognosis of lung adenocarcinoma in both Italian and Japanese populations in case-control studies (24, 25). These results suggest that the Pas1 gene may also play an important role in the susceptibility to human lung cancer. And thirdly, the recently completed mouse genome sequences will make positional cloning effort much easier than ever before (26). Accordingly, we reports here that the Pas1 locus on Chr6, which was predicted high degree of resolution for the Pas1 locus on Chr6, which now makes positional cloning a practical possibility.

In the Celera CDS 3.6 genome database, 27 known or unknown genes have been identified in the refined Pas1 candidate region. Kras2 has long been considered a major candidate for Pas1 (reviewed in Refs. 27, 28). We performed a lung tumor bioassay in heterozygous Kras2-deficient mice recently to evaluate the effect of presence of the wt Kras2 allele on lung tumorigenesis (29). Mice with a heterozygous Kras2 deficiency had an increased susceptibility to the chemical induction of lung tumors when compared with wt mice (29). Treatment of mice heterozygous for Kras2-deficiency produced four times as many tumors per lung as Kras2 wt mice regardless of the remaining allele being wt A/J Kras2 allele or wt C57BL/6J Kras2 allele (29). This result suggests that Kras2 is not likely to be the major candidate for Pas1 QTL in terms of its effect on lung tumor multiplicity, because the Pas1 locus has been reported to be responsible for >20-fold difference in lung tumor multiplicity between A/J and C57BL/6J mice (8–12).

Among all of the genes located in the refined Pas1 candidate region, our study identified the Lrmp and Ak016641 genes as strong candidates for the Pas1 QTL. In the Celera mouse genome, these two genes are located in a 276-kb physical region. Previous studies suggest that Lrmp may play a role in lymphoid development. Specifically, Lrmp may be involved in developmentally regulated intracellular trafficking of antigen receptors of lymphocytes (23). Lrmp protein can efficiently deliver a COOH-terminal antigenic peptide to MHC class I molecules in a transporter associated with antigen processing-independent manner (30). In the present study, we found eight nucleotide changes in the coding region of the Lrmp gene, and five of which result in amino acid variations between A/J and C57BL/6J strains of mice. The strain distribution pattern demonstrates that these polymorphisms cosegregate with mouse lung tumor susceptibility. Four of the five polymorphisms in codons 31, 56, 58, and 438 are localized in the cytosolic domain of the protein. Codons 31, 56, and 438 exhibit nonconservative changes, which may lead to conformational changes in the protein. Codon 537 is localized in the luminal domain near the COOH terminus, which is cleaved during post-translation processing (31). Whether the codon 537 substitution affects this processing is unclear. Lrmp is expressed in lung tissues of all of the inbred mouse strains examined. Interestingly, Lrmp expression increased in lung tumors relative to normal lungs (data not shown). The human 12p region containing SOX5, KRA2, and LRMP is amplified frequently in testicular germ cell tumors (32). It would be interesting if the same were true for mouse lung tumors.

The Pas1c1 (Ak016641) gene encodes a 413-amino acid interme-

**Fig. 3. Physical positions of six Pas1 candidate genes. The refined Pas1 candidate region is flanked by the D6Osu6 and D6Mit294 markers. Six putative Pas1 candidate genes (Eca39, Lrmp, Ak015530, Ak016641, mHoj-1, and Krag) were investigated in the present study. The position of K-ras relative to these Pas1 candidates is indicated. Their physical chromosomal positions were derived from Celera genomic database CDS 3.6.**
diate filament tail domain-containing protein. Although there is no available functional information, our finding that the transcript-splicing pattern cosegregates with mouse *Pas1* allele status provides genetic evidence for *Pas1c1* (**Ak016641**) as a candidate gene for *Pas1*. The nucleotide polymorphisms in the spliced exon (codons 258, 266, and 269) may play a role in determination of the splice pattern. A mechanism for exon skipping caused by nonsense or missense mutations in BRCA and other genes may also apply to the *Pas1c1* (**Ak016641**) gene (33).

The *Eca39* gene was isolated by a subtraction/coexpression strategy with Myc-induced tumors in transgenic mice, and demonstrated that *Eca39* is a direct genetic target for Myc regulation (34). The **Ak015530** gene encodes for an 86 amino acid small protein with unknown function, and it is unlikely a candidate *Pas1* gene, because

---

**Table 2** Nucleotide polymorphisms in the Lrmp, **Ak015530**, and **Ak016641** genes that alter amino acids

<table>
<thead>
<tr>
<th>Inbred strains</th>
<th><em>Pas1</em> alleletype</th>
<th>Lrmp</th>
<th><strong>Ak015530</strong></th>
<th><strong>Ak016641</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>129/SvJ</td>
<td><em>Pas1/s</em></td>
<td>GAC</td>
<td>GGC</td>
<td>TTC</td>
</tr>
<tr>
<td>BALB/cJ</td>
<td><em>Pas1/s</em></td>
<td>GAC</td>
<td>GCC</td>
<td>TTC</td>
</tr>
<tr>
<td>SWR/J</td>
<td><em>Pas1/s</em></td>
<td>GGC</td>
<td>TTC</td>
<td>AGG</td>
</tr>
<tr>
<td>CBA/J</td>
<td><em>Pas1/s</em></td>
<td>GGC</td>
<td>TTC</td>
<td>AGG</td>
</tr>
<tr>
<td>SM/J</td>
<td><em>Pas1/s</em></td>
<td>GAC</td>
<td>GCC</td>
<td>TTC</td>
</tr>
<tr>
<td>C57BL/6J</td>
<td><em>Pas1/r</em></td>
<td>GGC</td>
<td>TTC</td>
<td>GGG</td>
</tr>
<tr>
<td>SII/J</td>
<td><em>Pas1/r</em></td>
<td>GGC</td>
<td>TTC</td>
<td>GGG</td>
</tr>
<tr>
<td>AKR/J</td>
<td><em>Pas1/r</em></td>
<td>GGC</td>
<td>TTC</td>
<td>GGG</td>
</tr>
<tr>
<td>C3H/HeJ</td>
<td><em>Pas1/r</em></td>
<td>GGC</td>
<td>TTC</td>
<td>GGG</td>
</tr>
<tr>
<td>DBA/2J</td>
<td><em>Pas1/r</em></td>
<td>GGC</td>
<td>TTC</td>
<td>GGG</td>
</tr>
<tr>
<td>Mus. Spretus</td>
<td><em>Pas1/r</em></td>
<td>GGC</td>
<td>TTC</td>
<td>GGG</td>
</tr>
</tbody>
</table>
the amino acid polymorphism detected does not cosegregate with mouse lung tumor susceptibility (Table 2). We derived the mouse mHoj-1 gene from its human homologue HOJ-1 in the present study. Although the function of the mHoj-1 gene remains to be determined, searching the National Center for Biotechnology Information Conserved Domain Database for the protein sequence of mHoj-1 led to a prediction of a Ras association (RalGDS/AF-6) domain in its NH₂ terminus. This suggests that mHoj-1 may function as a RasGTP effector. The human KRAG gene is related to autosomal recessive limb-girdle muscular dystrophies 2C through 2F diseases (35), and is coamplified with KRA52 and ITPR2 genes in certain tumors (36). Our results show the above-mentioned four genes (Eca39, RIKEN Ak015530, mHoj-1, and Krag) are expressed in mouse lung tissue but to similar extents in susceptible and resistant strains of mice. Moreover, there was no amino acid polymorphism in the coding regions except for Ak015530. These results suggest that these 4 genes are less likely to be candidates for Pas1.

Results from the present study have significant implications for identifying candidate genes for the Pas1 locus, of which the human homologue may predispose some individuals to lung cancer. Our results establish the candidacy of Lrmp and Pas1c1 (Ak016641) genes for the Pas1 locus, and these genes may be responsible for differences in lung tumor multiplicity between susceptible and resistant strains.

Additional functional evidence is required for precise identification of the Pas1 gene(s). However, it is important to note that whereas AIL mapping can considerably improve QTL resolution, our studies indicate that, despite care in randomizing mating and maintaining effective population size, QTL alleles may be lost in the development of AIL.

ACKNOWLEDGMENTS

We thank Gary Stoner for critical reading of this manuscript and Elizabeth Wiley for secretarial assistance, and Bob King of the International Livestock Research Institute, Nairobi, Kenya, for excellent stewardship of the AIL lines used in this study.

REFERENCES


Fine Mapping and Identification of Candidate Pulmonary Adenoma Susceptibility 1 Genes Using Advanced Intercross Lines

Min Wang, William J. Lemon, Gongjie Liu, et al.


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/63/12/3317

Cited articles
This article cites 33 articles, 9 of which you can access for free at:
http://cancerres.aacrjournals.org/content/63/12/3317.full#ref-list-1

Citing articles
This article has been cited by 6 HighWire-hosted articles. Access the articles at:
http://cancerres.aacrjournals.org/content/63/12/3317.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link:
http://cancerres.aacrjournals.org/content/63/12/3317.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.