Molecular Characterization of the Mouse In(10)17Rk Inversion and Identification of a Novel Muscle-Specific Gene at the Proximal Breakpoint

Kathleen F. Benson and Kiran Chada

Department of Biochemistry, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, Piscataway, New Jersey 08854

Manuscript received July 19, 2001
Accepted for publication October 8, 2001

ABSTRACT

Chromosomal rearrangements provide an important resource for molecular characterization of mutations in the mouse. In(10)17Rk mice contain a paracentric inversion of ~50 Mb on chromosome 10. Homozygous In(10)17Rk mice exhibit a pgmy phenotype, suggesting that the distal inversion breakpoint is within the pgmy locus. The pgmy mutation, originally isolated in 1944, is an autosomal recessive trait causing a dwarf phenotype in homozygous mice and has been mapped to the distal region of chromosome 10. The pgmy phenotype has subsequently been shown to result from disruption of the Hmgi-c gene. To identify the In(10)17Rk distal inversion breakpoint, In(10)17Rk DNA was subjected to RFLP analysis with single copy sequences derived from the wild-type pgmy locus. This analysis localized the In(10)17Rk distal inversion breakpoint to intron 3 of Hmgi-c and further study determined that a fusion transcript between novel 5' sequence and exons 4 and 5 of Hmgi-c is created. We employed 5' RACE to isolate the 5' end of the fusion transcript and this sequence was localized to the proximal end of chromosome 10 between markers Cni-rs2 and Mtap7. Northern blot analysis of individual tissues of wild-type mice determined that the gene at the In(10)17Rk proximal inversion breakpoint is a novel muscle-specific gene and its disruption does not lead to a readily observable phenotype.

The pgmy (pg) locus on mouse chromosome 10 is an important regulator of body size. Pygmy mice originally appeared in 1944 as undersized segregants in a strain of mice selected for small size (MacArthur 1944). Previously in our laboratory, a transgenic inser- tional mutant that was shown to be allelic to the spontaneous mutant pgmy was isolated (Xiang et al. 1990). The founder transgenic mouse harbored two different transgene integration events at the pgmy locus, which were segregated in the progeny, forming two distinct lines, designated pgTgV60B3a (A) and pgTgV60B5a (B). Subsequent characterization of the A, B, and pg mutations determined that sizable deletions had occurred in each of these mutants (Xiang et al. 1990; Zhou et al. 1995; K. F. Benson and K. Chada, unpublished results).

The creation of new mutations at the pgmy locus by transgene insertion resulted in the identification of the Hmgi-c gene at the pgmy locus and a targeted disruption demonstrated that its loss of expression was the cause of the pgmy phenotype (Zhou et al. 1995). Translocations involving human Hmgi-c that create fusion transcripts composed of exons 1 through 3 of Hmgi-c and novel 3' partners have been shown to cause tumors of mesenchymal origin (Ashar et al. 1995), which is consistent with the wild-type expression pattern of murine Hmgi-c (Zhou et al. 1995, 1996). While Hmgi-c is a developmentally regulated gene with highest levels of expression in the embryo (Zhou et al. 1995, 1996), recent studies have demonstrated the importance of Hmgi-c expression for expansion of adipocytes in white adipose tissue of the adult mouse (Anand and Chada 2000).

Another line of mice with a pgmy phenotype arose as a consequence of exposure to the chemical mutagen triethylenemelamine during a large scale mutagenesis experiment conducted at Jackson Laboratory. In(10)17Rk mice contain a paracentric inversion of approximately 50 Mb within mouse chromosome 10 between proximal region A4 and distal region D2 (Roderick 1983). Homozygous In(10)17Rk mice exhibit a pgmy phenotype and the pgmy locus maps to the D2 region of mouse chromosome 10 (Lyon and Searle 1989), suggesting that the distal inversion breakpoint may be within the pgmy locus. Ongoing studies in our laboratory to identify the pgmy gene by positional cloning methods provided an array of probes along the pgmy locus for investigating the precise location of the In(10)17Rk distal inversion breakpoint. This study describes the molecular characterization of both the proximal and distal In(10)17Rk inversion breakpoints and the consequent identification of a novel muscle-specific gene disrupted by the proximal breakpoint.
MATERIALS AND METHODS

**Genotyping:** The genotypes were established for mice in line pg^{6710McG} and the spontaneous mutant pg as previously described (Xiang et al. 1990), and mice containing the In(10)17Rk inversion were detected by a PCR-based restriction fragment length polymorphism as described in Cherath et al. (1995).

**DNA preparation and blot hybridization:** High-molecular-weight genomic DNA was isolated from mouse livers using standard procedures (Ausubel et al. 1988). Genomic DNA (10 µg) of each genotype was digested with various restriction enzymes, electrophoresed in 0.8% agarose gels, and transferred to Duralon (Stratagene, La Jolla, CA). DNA hybridizations were performed as described in Ashar et al. (1994).

**Determination of the 5’ end of the fusion transcript:** We performed 5’ rapid amplification of cDNA ends (RACE) on RNA from homozygous In(10)17Rk embryos using antisense primer 4862, (5’-ATGGATCCCTAATCCTCCTCTGC-3’) from the beginning of exon 5 of HmgI-c, for first-strand cDNA synthesis. For reverse transcription, 20 pmol antisense oligonucleotides were used to prime cDNA synthesis in a 20 µl reaction containing 1 µg of total RNA from 12.5-days post-coitum (dpc) homozygous In(10)17Rk embryos. First-strand cDNAs were A-tailed and then used for subsequent PCR in a 50 µl reaction containing 1X PCR buffer (0.05 M KCl, 0.01 M Tris-HCl, pH 8.3, 0.015 M MgCl2, 0.001% gelatin), 50 pmol sense oligo(T) primer 4999, (5’-GCAATACGACTCACTATAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
sions predominantly result in retention of all genetic material (Bultman et al. 1992; Quaderi et al. 1997; Perry et al. 1998), breakpoints can be identified only by probes that hybridize to restriction fragments bridging the inversion breakpoint. It was reasoned that if the distal inversion breakpoint was within the pygmy locus, it would be identified as a restriction fragment length polymorphism (RFLP) between the wild-type and In(10)17Rk alleles. Single copy sequences were isolated by restriction enzyme analysis of λ-clones derived from a yeast artificial chromosome clone containing the pygmy locus (Zhou et al. 1995) and hybridized to Southern blots of DNA from C57BL/6J, hemizygous In(10)17Rk, and homozygous In(10)17Rk mice digested with EcoRI and BamHI. A large (>30 kb) BamHI genomic fragment migrated faster in In(10)17Rk DNA (data not shown) and was investigated further. This region of the pygmy locus is encompassed by λ-clones Y-408 and 312 (Figure 1a). Further analysis of these clones identified a 2.5 kb SstI/XbaI single copy fragment that, when used as a probe, detected a polymorphism for the restriction enzyme XbaI (Figure 1b). DNA from the mouse strain DBA/2J was also included in the analysis since this is the original strain on which the In(10)17Rk inversion was induced (Roderick 1983). Southern blot analysis with the restriction enzymes XbaI, SstI, and EcoRI and single copy sequences from λ-clones 312 and Y-408 (Figure 1a) localized the In(10)17Rk distal inversion breakpoint to a 6.5-kb EcoRI/SalI fragment present in clone 312. This 6.5-kb fragment was cloned into pBluescript KS+ (Stratagene) and further analyzed by restriction enzyme digestion (Figure 2A).

A comparison between the Southern blot results (including the additional enzyme PvuII) and the restriction enzyme sites within the 6.5-kb sequence determined that the In(10)17Rk distal inversion breakpoint was located within a 0.6-kb PvuII/SstI fragment (Figure 2A). Southern blots were then prepared from DNA of C57BL/6J, hemizygous In(10)17Rk, homozygous In(10)17Rk, and DBA/2J mice digested with the restriction enzymes HindIII, PstI, EcoRI, and XbaI (Figure 2B) as well as the additional enzymes BamHI, HincII, KpnI, and PvuII (data not shown) and hybridized with the 0.6-kb PvuII/SstI sequence. For all enzymes analyzed, the homozygous In(10)17Rk DNA produced a novel hybridization pattern consisting of two unique bands that differed from the C57BL/6J and DBA/2J lanes, while hemizygous In(10)17Rk DNA produced a compound hybridization pattern reflecting the presence of both a wild-type and In(10)17Rk inversion allele (Figure 2B). This places the distal inversion breakpoint within intron 3 of Hmgi-c (Figure 3).

The In(10)17Rk inversion disrupts Hmgi-c expression and creates a fusion transcript: Hmgi-c is encoded by five exons that span ~110 kb (Zhou et al. 1996). To investigate the status of Hmgi-c transcripts within In(10)17Rk mice, Northern blots were prepared from total RNA of 12.5-dpc embryos of the genotypes C57BL/6J, hemizygous In(10)17Rk, and homozygous In(10)17Rk and hybridized with Hmgi-c exonic sequences that were either 5’ (exons 2 and 3) or 3’ (exon 5) to the distal inversion breakpoint (Figure 3). The 5’ probe failed to detect any transcripts in homozygous In(10)17Rk embryos (Figure 4). The identification of a transcript in homozygous In(10)17Rk mice by the exon 5 probe, which is similar in size to Hmgi-c although at greatly reduced levels (Figure 4), was observed only upon overexposure of the Northern blot. On the basis of the size of this transcript (3.7 kb) and the positioning of the In(10)17Rk distal inversion breakpoint within intron 3 of Hmgi-c, we postulated that the RNA species detected was a fusion transcript consisting of exons 4 and 5 of Hmgi-c (3.1 kb) fused to an ~600-bp novel 5’ sequence.

5’ RACE isolates novel sequence from fusion transcript: Identification of a transcript from the In(10)17Rk allele provided an opportunity to isolate the 5’ sequence of this fusion transcript that is not derived from Hmgi-c. Performing 5’ RACE on RNA from homozygous In(10)17Rk embryos using a specific primer from exon 5 of Hmgi-c resulted in the isolation of a 670-bp product. When this product was cloned and sequenced, it demonstrated a novel 586-bp 5’ sequence followed by the 33-bp exon 4 of Hmgi-c spliced to the beginning of exon 5 of Hmgi-c (Figure 5A).

RT-PCR was performed to confirm the origin of the
fusion transcript from the \textit{In(10)17Rk} allele. Using the combination of a primer from exon 5 of \textit{Hmgi-c} and a primer derived from the novel sequence of the 5' RACE product, a predicted product of 216 bp was amplified upon RT-PCR of hemizygous and homozygous \textit{In(10)17Rk} RNA samples (Figure 5B). RT-PCR of wild-type RNA did not produce an amplification product.

**Chromosomal localization of \textit{In(10)17Rk-p}:** A 219-bp sequence was PCR amplified from the novel 5' sequence of the fusion transcript (see materials and methods). Southern blot analysis of the parental mouse strains C57BL/6JEi and SPRET/Ei with the restriction enzyme \textit{HindIII} identified fragments of 3.2 and 5.0 kb, respectively. This polymorphism was utilized for analysis of 94 progeny from the Jackson BSS backcross mapping panel (Rowe et al. 1994). The data place \textit{In(10)17Rk-p} in the proximal end of chromosome 10 between the markers \textit{Cni-\alpha2} and \textit{Mtap7} (Figure 6A).

**Tissue distribution of \textit{In(10)17Rk-p} transcripts:** The 219-bp sequence derived from the \textit{In(10)17Rk-p} fusion transcript was used for Northern blot analysis of wild-type tissues. This identified transcripts of 2.5 and 4.4 kb in adult heart, femur muscle, and diaphragm (Figure 7A). Transcripts were not detected in skin, white adipose tissue, spleen, thymus, or kidney (Figure 7A) or in the additional 10 adult tissues of liver, lung, adrenal, brain, pituitary, stomach, small and large intestine, pancreas, and testes (data not shown). When 11 newborn tissues (brain, heart, lung, liver, kidney, stomach, small and large intestine, spleen, pancreas, and skin) were analyzed, only heart exhibited expression, at levels lower than those of adult (data not shown).

Next the presence of transcripts was assessed in heart and femur muscle of adult hemizygous and homozygous \textit{In(10)17Rk} mice. Compared to C57BL/6J, expression levels are reduced in hemizygous \textit{In(10)17Rk} muscle tissue and completely absent in homozygous mice (Figure 7B). Hemizygous \textit{In(10)17Rk} RNA contains both wild-type and fusion transcripts (lanes 3–4) while only the fusion transcript is detected in homozygous \textit{In(10)17Rk} RNA (lanes 5–6).

**Isolation of the full-length 2.5- and 4.4-kb transcripts:** Gene-specific primers were designed from the novel sequence of the fusion transcript and used for 3' RACE of adult C57BL/6J heart RNA (see materials and methods). This resulted in the isolation of both the

---

**Figure 2**—(A) Detailed restriction enzyme map of a 6.5-kb \textit{EcoRI}/\textit{SalI} fragment derived from \textit{\lambda}-clone 312. The striped box indicates a 0.6-kb single copy sequence that spans the \textit{In(10)17Rk} distal inversion breakpoint. \textit{E}, \textit{EcoRI}; \textit{H}, \textit{HindIII}; \textit{K}, \textit{KpnI}; \textit{P}, \textit{PstI}; \textit{Pv}, \textit{PvuII}; \textit{S}, \textit{SalI}; \textit{Ss}, \textit{SstI}. (B) Southern blot of C57BL/6J (lanes A), hemizygous \textit{In(10)17Rk} (lanes B), homozygous \textit{In(10)17Rk} (lanes C) or DBA/2J (lanes D) adult genomic DNA digested with \textit{HindIII}, \textit{PstI}, \textit{EcoRI} or \textit{XbaI}. Each sample was hybridized with a 0.6-kb \textit{PvuII}/\textit{SstI} sequence derived from a 6.5-kb \textit{EcoRI}/\textit{SalI} fragment from \textit{\lambda}-clone 312. \textit{M}, \textit{\lambda} \textit{HindIII} molecular weight marker (GIBCO-BRL).

**Figure 3**—\textit{\lambda}-clones spanning the murine \textit{Hmgi-c} locus. Gap refers to an 11-kb unclonable region of the locus (K. Przybysz, K. F. Benson and K. Chada, unpublished data). Exons are represented as solid boxes and are not drawn to scale. An arrow indicates the \textit{In(10)17Rk} distal breakpoint within intron 3.
In(10)17Rk Inversion Disrupts Two Genes

Figure 4.—Northern blot of 5 μg total RNA from C57BL/6J (lane 1), hemizygous In(10)17Rk (lane 2), or homozygous In(10)17Rk (lane 3) 12.5-dpc embryos probed with either exons 2 and 3 (left) or exon 5 (right) of murine Hmgi-c. The blots were exposed to film at –70°C for 4 and 48 hr, respectively, and reprobed with an oligonucleotide complementary to murine 28S ribosomal RNA (Barbu and Dautry 1989) to ensure equal loading of samples.

Gene disrupted by In(10)17Rk-p is homologous to a chicken muscle-specific gene: BLAST searches of the NCBI databases with the 2.5- and 4.4-kb transcript sequences returned significant homology to a chicken muscle-specific protein called MDP77 (Uyeda et al. 2000; GenBank accession no. BAA94755). There is 74% identity and 85% similarity over 463 residues beginning with the initiating methionine (Figure 8). This homology is revealed when the mouse sequences are translated in the +1 frame. There are a few gaps in alignment in the amino-terminal end of the proteins with the chicken MDP77 gene encoding a 676-amino-acid protein and the two mouse transcripts encoding an identical 684-amino-acid protein. The carboxy-terminal one-third of the chicken and mouse proteins completely diverges (data not shown). No other similarities were detected at the protein level.

Currently the mouse genomic sequence from the region of chromosome 10 containing the MDP77 gene is not available in the NCBI sequence databases. However, homology searches performed at the nucleotide level returned matches to the chicken MDP77 gene as well as to separate but consecutive regions of a genomic clone derived from human chromosome 6 (clone RP3-522B19, GenBank accession no. AL158850). This was seen only with the regions of the 2.5- and 4.4-kb cDNAs that are 3’ to the sequence present in the fusion transcript. These regions of punctuated homology between the mouse cDNA sequence and human chromosome 6 are presumed to represent individual exons, indicating that the 2.5-kb mouse RNA is encoded by a minimum of 11 exons (data not shown).

The PCR amplification and sequencing of a 3.5-kb fragment from mouse genomic DNA using primers derived from the 5’ (primer 4826) and 3’ (primer 6876) regions of the novel 586-bp sequence of the fusion transcript (Figure 5A) revealed the presence of a single intron (data not shown). This information determines that the fusion transcript is composed of exons 1 and 2 of mouse MDP77 fused to exons 4 and 5 of Hmgi-c and places the In(10)17Rk-p inversion breakpoint within intron 2 of the mouse MDP77 gene. Figure 8 indicates the position of the In(10)17Rk-p inversion breakpoint within the protein separating the first 143 amino acids (encoded by exons 1 and 2) from the remainder of the protein.

DISCUSSION

Allelism studies confirm In(10)17Rk as another pg allele: The pygmy mutation arose spontaneously in 1944 (MacArthur 1944) as an autosomal recessive trait causing a dwarf phenotype and was mapped to mouse chromosome 10.
Figure 6.—(A) Map figure showing the proximal end of Jackson BSS chromosome 10 with loci linked to In(10)17Rk-p. The map is depicted with the centromere toward the top. A 3-cm scale bar is shown to the right. Loci mapping to the same position are listed in alphabetical order. (B) Haplotypes of the region surrounding the In(10)17Rk-p inversion breakpoint on chromosome 10 in 94 progeny from the Jackson BSS cross. Solid boxes represent the C57BL6/J allele; open boxes represent SPRET/Ei alleles. The number of animals used to derive the data is given at the bottom of each column of boxes. The percentage recombination (R) between adjacent loci is given to the right, with the standard error (SE) for each R. Missing typings were inferred from surrounding data when assignment was unambiguous. Raw data from the Jackson Laboratory were obtained from http://www.jax.org/resources/documents/cpdata.

(Lyon and Searle 1989); however, a molecular probe was not available to identify the gene whose disruption gave rise to the phenotype. Remutation of the pygmy locus by transgene integration (Xiang et al. 1990) facilitated cloning of the locus and the subsequent identification of the pygmy gene as Hmgi-c (Zhou et al. 1995). During the molecular analysis of the transgenic insertional mutant we initiated studies of In(10)17Rk mice that contain an ~50-Mb inversion in chromosome 10 and also exhibit a pygmy phenotype (Roderick 1983). The allelism studies confirm the hypothesis that In(10)17Rk represents another allele of pygmy. Both Rk/+ × pg/+ and Rk/+ × A/+ matings produced small progeny in proportions consistent with an autosomal recessive trait and all small mice were shown to be compound heterozygous for the In(10)17Rk inversion and either the pg or A mutations, respectively.

In(10)17Rk-d localized to intron 3 of Hmgi-c: Next a detailed RFLP analysis of the pygmy locus was undertaken to identify the In(10)17Rk distal inversion breakpoint. The identification of an XbaI polymorphism by Southern analysis with a single copy sequence from λ-clone 312 led to the final localization of the In(10)17Rk-d inversion breakpoint to a 0.6-kb PvuII/SstI fragment within intron 3 of Hmgi-c. The relatively minimal loss of genetic material found in the analysis of some inversions is well illustrated by rearrangements occurring at the agouti (a) locus on mouse chromosome 2. Characterization of the distal inversion breakpoint of the Is1Gso mutant detected a deletion of only 29 bp (Bultman et al. 1992), while characterization of the proximal and distal inversion breakpoints in the a18H mouse identified deletions of only 18 and 20 bp, respectively (Perry et al. 1998). Since we found large deletions in the transgenic insertional and spontaneous pygmy mutants (Xiang et al. 1990; Zhou et al. 1995; K. F. Benson and K. Chada, unpublished results), analysis of the In(10)17Rk inversion has proven crucial for narrowing down the location of the pygmy gene.

The localization of the In(10)17Rk distal inversion breakpoint to intron 3 of Hmgi-c provides a molecular explanation for the disruption of Hmgi-c expression in In(10)17Rk mice resulting in the pygmy phenotype. Interestingly, intron 3 of human Hmgi-c is the most common site of chromosomal rearrangement in human tumors (Kazmierczak et al. 1998), suggesting that intron 3 of Hmgi-c may contain some sequence element such as a low copy repeat (Shaffer and Lupski 2000) that confers susceptibility to rearrangement. Evidence that this element could be conserved in intron 3 of mouse Hmgi-c is the relative absence of repetitive sequences in this region (K. F. Benson and K. Chada, unpublished results).
In(10)17Rk inversion disrupts two genes.

Fusion transcript created by inversion facilitates mapping of In(10)17Rk-p. Northern blot analysis of homozygous In(10)17Rk embryonic RNA with exon 5 of Hmgi-c identified a transcript of ~5.7 kb. Since the In(10)17Rk distal inversion breakpoint is within intron 3 of Hmgi-c, resulting in the separation of exons 1–3 from 4 and 5, this transcript was hypothesized to consist of novel 5’ sequence and Hmgi-c exons 4 and 5. A 586-bp novel sequence was identified by RT-PCR and further utilized from homozygous In(10)17Rk embryos (data not shown).

In(10)17Rk-p disrupts a muscle-specific gene: We investigated the wild-type expression pattern of the novel portion of the In(10)17Rk-p fusion transcript. Northern blot analysis of 18 adult and 11 newborn tissues with the 219-bp probe detected transcripts of 2.5- and 4.4-kb chromosome 10. Northern analysis with exons 2 and 3 of Hmgi-c failed in adult heart, femur muscle, and diaphragm and newborn heart. All other tissues were negative for expression. The finding that the disrupted gene is homologous to a chicken muscle-specific gene that encodes a 77-kD protein called MDP77 (Uyeda et al. 2000) supports our conclusion that the In(10)17Rk-p inversion breakpoint creates a reciprocal fusion transcript composed of Hmgi-c exons 1–3 and novel 3’ sequences or results in truncated Hmgi-c consisting of exons 1–3. This is supported by the findings that both Hmgi-c truncations and fusion with novel 3’ partners result in tumors (Ashar et al. 1995; Fedele et al. 1998; Battista et al. 1999), which were not seen in In(10)17Rk identified a transcript of ~3.7 kb. Since the In(10)17Rk distal inversion breakpoint is within intron 3 of Hmgi-c, resulting in the separation of exons 1–3 from 4 and 5, this transcript was hypothesized to consist of novel 5’ sequence and Hmgi-c exons 4 and 5. A 586-bp novel sequence was identified by RT-PCR and further utilized from homozygous In(10)17Rk embryos (data not shown).

In(10)17Rk-p disrupts a muscle-specific gene: We investigated the wild-type expression pattern of the novel portion of the In(10)17Rk-p fusion transcript. Northern blot analysis of 18 adult and 11 newborn tissues with the 219-bp probe detected transcripts of 2.5- and 4.4-kb chromosome 10. Northern analysis with exons 2 and 3 of Hmgi-c failed in adult heart, femur muscle, and diaphragm and newborn heart. All other tissues were negative for expression. The finding that the disrupted gene is homologous to a chicken muscle-specific gene that encodes a 77-kD protein called MDP77 (Uyeda et al. 2000) supports our conclusion that the In(10)17Rk-p inversion breakpoint creates a reciprocal fusion transcript composed of Hmgi-c exons 1–3 and novel 3’ sequences or results in truncated Hmgi-c consisting of exons 1–3. This is supported by the findings that both Hmgi-c truncations and fusion with novel 3’ partners result in tumors (Ashar et al. 1995; Fedele et al. 1998; Battista et al. 1999), which were not seen in In(10)17Rk identified a transcript of ~3.7 kb. Since the In(10)17Rk distal inversion breakpoint is within intron 3 of Hmgi-c, resulting in the separation of exons 1–3 from 4 and 5, this transcript was hypothesized to consist of novel 5’ sequence and Hmgi-c exons 4 and 5. A 586-bp novel sequence was identified by RT-PCR and further utilized for interspecific backcross mapping studies. The placement of In(10)17Rk-p in the proximal end of chromosome 10 between markers Cni-rs2 and Mtap7 determines that the In(10)17Rk inversion encompasses ~60 cM of chromosome 10.

Northern analysis with exons 2 and 3 of Hmgi-c failed to detect transcripts in homozygous In(10)17Rk embryos when levels of Hmgi-c are normally high. Therefore, it is unlikely that the In(10)17Rk inversion creates a reciprocal fusion transcript composed of Hmgi-c exons 1–3 and novel 3’ sequences or results in truncated Hmgi-c consisting of exons 1–3. This is supported by the findings that both Hmgi-c truncations and fusion with novel 3’ partners result in tumors (Ashar et al. 1995; Fedele et al. 1998; Battista et al. 1999), which were not seen in In(10)17Rk identified a transcript of ~3.7 kb. Since the In(10)17Rk distal inversion breakpoint is within intron 3 of Hmgi-c, resulting in the separation of exons 1–3 from 4 and 5, this transcript was hypothesized to consist of novel 5’ sequence and Hmgi-c exons 4 and 5. A 586-bp novel sequence was identified by RT-PCR and further utilized for interspecific backcross mapping studies. The placement of In(10)17Rk-p in the proximal end of chromosome 10 between markers Cni-rs2 and Mtap7 determines that the In(10)17Rk inversion encompasses ~60 cM of chromosome 10.
disrupts a muscle-specific gene. MDP77 was isolated from a cDNA library derived from chicken denervated crus muscle and Northern blotting and in situ hybridization localized expression to cardiac and skeletal muscle (Uyeda et al. 2000). The high level of amino acid conservation and similar expression pattern suggest that the gene residing at the In(10)17Rk-p inversion breakpoint is the murine homolog of MDP77.

Expression of murine MDP77 is completely absent in homozygous In(10)17Rk mice as determined by Northern blot analysis of adult heart and femur muscle. Nevertheless, homozygous In(10)17Rk mice do not exhibit any obvious phenotype in addition to that of pygmy. Based upon translation of the complete mouse cDNA sequence and comparison to chicken MDP77, the majority of the protein is deleted, including a putative leucine zipper (see Figure 8). Therefore it is unlikely that the fusion transcript retains function of the muscle-specific gene. Instead, lack of an additional phenotype may reflect gene redundancy as exemplified by the muscle-specific genes dystrophin and utrophin (Grady et al. 1997).

There is a (CTG)9 (CTC)1 repeat in the 5' end of the mouse cDNA that is 69 bp preceding the translational initiation codon (position –105 to –69) and interestingly, there is a (CTG) repeat at a similar position in the chicken MDP77 cDNA, position –79 to –70 (GenBank accession no. D89999). Recently a number of CUG repeat-containing RNAs have been identified (Lu et al. 1999; Miller et al. 2000; Ladd et al. 2001). These appear to be expressed predominantly in muscle tissue and the CUG repeats are important for the appropriate processing of these RNAs by RNA-binding proteins (Philips et al. 1998; Miller et al. 2000; Ladd et al. 2001). One such example is the DMPK protein kinase gene that contains a CTG repeat in the 3' untranslated region, expansion of which leads to the human illness myotonic dystrophy (Brook et al. 1992; Fu et al. 1992; Mahadevan et al. 1992). Evidence suggests that the disease occurs as a consequence of the deregulation of processing of multiple CUG repeat-containing RNAs (Philips et al. 1998; Lu et al. 1999; Mankodi et al. 2000). This is supported by the recent report by LiQüori et al. (2001), identifying the expansion of a CCTG repeat in the ZNF9 gene as the genetic cause of myotonic dystrophy type 2 (DM2). The underlying pathogenic mechanism common to these two similar diseases is the production of RNAs containing an expanded CUG repeat (DM1) or CUCG repeat (DM2).

One feature of myotonic dystrophy is cardiomyopathy, which has been suggested to reflect the degree of CTG expansion (Finsterer et al. 2001). Based on loci flanking the In(10)17Rk-p inversion breakpoint, the human syntenic region is chromosome 6q22-25 (Mouse Genome Database; http://www.informatics.jax.org). Intriguingly, two forms of autosomal dominant dilated cardiomyopathy have been localized here. These are dilated cardio-

myopathy CMD1F, mapping to 6q23 [Online Mendelian Inheritance in Man (OMIM) entry 602067; MESSINA et al. 1997], and dilated cardiomyopathy CMD1J, mapping to 6q23-24 (OMIM 605362; SCHONBERGER et al. 2000). The detection of homology at the nucleotide level between the mouse MDP77 cDNA sequence and the human chromosome 6-derived genomic clone RP3-522B19 strongly suggests the existence of a human MDP77 gene in this region. Additionally, clone RP3-522B19 has been placed within a larger contiguous sequence of human chromosome 6 (GenBank accession no. NT_025741) and is flanked by some of the same microsatellite markers used to map the CMD1F and CMD1J loci (MESSINA et al. 1997; SCHONBERGER et al. 2000), further emphasizing the suitability of human MDP77 as a candidate gene for these diseases.

Characterization of the proximal and distal In(10)17Rk inversion breakpoints has provided both an explanation for the pygmy phenotype of In(10)17Rk mice and identified a novel muscle-specific gene in mouse. A homolog of this gene has been reported in chicken and a human version is likely to be present in the syntenic region of human 6q22-25. The mapping of two forms of autosomal dominant dilated cardiomyopathy to this chromosomal region makes MDP77 a candidate gene for these diseases.

This work was supported by a grant from the National Institutes of Health (CA77929).

LITERATURE CITED


FEDDE, M., M. T. BERLINGIERI, S. SCALA, L. CIARIOTTI, G. VIGLIETTO et al., 1998 Truncated and chimeric HMGIC genes induce neo...
plastic transformation of NIH3T3 murine fibroblasts. Oncogene 17: 413–418.


Kazmierczak, B., J. Bullerdiek, K. H. Pham, S. Bartnitzke and H. Wiesner, 1998 Intron 3 of HMGIC is the most frequent target of chromosomal aberrations in human tumors and has been conserved basically for at least 30 million years. Cancer Genet. Cytogenet. 103: 175–177.


Communicating editor: N. A. Jenkins