The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/24067

Please be advised that this information was generated on 2017-08-31 and may be subject to change.
Identification of the first gene (FRG1) from the FSHD region on human chromosome 4q35

Judith C. T. van Deutekom1,2, Richard J. L. F. Lemmers1, Prabhjit K. Grewal3, Michel van Geel1,2,*, Silke Romberg1,2, Hans G. Dauwerse1, Tracy J. Wright3,†, George W. Padberg4, Marten H. Hofker1, Jane E. Hewitt3 and Rune R. Frants1,**

1MGC-Department of Human Genetics, and 2Department of Neurology, Leiden University, Leiden, The Netherlands, 3School of Biological Sciences, University of Manchester, Manchester, UK and 4Department of Neurology, University of Nijmegen, Nijmegen, The Netherlands

Received January 29, 1996; Revised and Accepted February 23, 1996 GenBank accession nos L76159, L76173, L76174

Facioscapulohumeral muscular dystrophy (FSHD) is an autosomal dominant, neuromuscular disorder characterized by progressive weakness of muscles in the face, shoulder and upper arm. Deletion of integral copies of a 3.3 kb repeated unit from the subtelomeric region on chromosome 4q35 has been shown to be associated with FSHD. These repeated units which are apparently not transcribed, map very close to the 4q telomere and belong to a 3.3 kb repeat family dispersed over heterochromatic regions of the genome. Hence, position effect variegation (PEV), inducing allele-specific transcriptional repression of a gene located more centromeric, has been postulated as the underlying genetic mechanism of FSHD. This hypothesis has directed the search for the FSHD gene to the region centromeric to the repeated units. A CpG island was identified and found to be associated with the 5' untranslated region of a novel human gene, FRG1 (FSHD Region Gene 1). This evolutionarily conserved gene is located about 100 kb proximal to the repeated units and belongs to a multigene family with FRG1 related sequences on multiple chromosomes. The mature chromosome 4 FRG1 transcript is 1042 bp in length and contains nine exons which encode a putative protein of 258 amino acid residues. Transcription of FRG1 was detected in several human tissues including placenta, lymphocytes, brain and muscle. To investigate a possible PEV mechanism, allele-specific FRG1 steady-state transcript levels were determined using RNA-based single-strand conformation polymorphism (SSCP) analysis. A polymorphic fragment contained within the first exon of FRG1 was amplified from reverse transcribed RNA from lymphocytes and muscle biopsies of patients and controls. No evidence for PEV mediated repression of allelic transcription was obtained in these tissues. However, detection of PEV in FSHD patients may require analysis of more specific cell types at particular developmental stages.

INTRODUCTION

Facioscapulohumeral muscular dystrophy (FSHD) is a clinically heterogeneous, autosomal dominant myopathy with an estimated prevalence of 1 in 20 000 (1,2). Initially, the facial and shoulder-girdle muscles are affected but during progression of the disease the upper-arm, abdominal, pelvic-girdle and foot-extensor muscles can also become involved (1,3).

The FSHD locus has been assigned to the distal long arm of chromosome 4 (4q35) by linkage analysis (4,5), with a most likely location distal to D4S139 (6–8). More recently, rearrangements within a polymorphic EcoRI fragment were detected in FSHD patients using probe p13E-11 (D4F104S1) (9). These rearrangements were found to be due to deletion of an integral number of 3.3 kb tandemly repeated units (D4Z4) (9,10). Although these 3.3 kb repeated units contain a double homeobox motif (11,12), no transcripts derived from this locus could be identified (11–14). In addition, fluorescence in situ hybridization (FISH) experiments showed that the tandem array of 3.3 kb repeats maps very close to the 4q telomere (13,15,16), suggesting a probable location of the FSHD gene proximal to the repeated units. Recent data suggested that the repeated units are members of a 3.3 kb repeat family dispersed over heterochromatic regions.

*To whom correspondence should be addressed
†Present address: Department of Human Genetics, Roswell Park Cancer Institute, Buffalo, NY, USA
‡Present address: Life Sciences Division LS-3, Los Alamos National Laboratory, Los Alamos, NM, USA
of the genome (14). This is supported by FISH results demonstrating hybridization of the 3.3 kb repeated unit to the short arms of all acrocentric chromosomes as well as to chromosomes 1q12, 4q35, 10qter, and 10cen (13,16). Weak signals were also detected on a large heterochromatic region on chromosome 9q12, and chromosome Y. Therefore, considering that the 3.3 kb repeated unit may be heterochromatin-associated, deletion of a number of these repeated units may lead to position effect variegation (PEV) by alteration of the chromosomal structure. Heterochromatinization or extensive chromatin packaging (17) can transcriptionally silence a gene by (i) inducing a later replication time point which will put the gene at a disadvantage in the competition for transcription factors (18), or (ii) generating a more tightly condensed structure of the chromatin which itself contributes to its transcriptionally silent state (18–20). A classic example of heterochromatinization in mammals involves transcriptional silencing of genes by X chromosome inactivation (21–23). Furthermore, since telomeres have a functional location in the nucleus (24), the FSHD associated deletions may also have implications for the nuclear organization. PEV leading to reduced transcription levels of one allele of a gene has been observed in Drosophila, yeast and mammals (17,25), and may affect genes over a region of several hundred kb (26). When taken into account that the deletion prone chromosomal distribution of 3.3 kb repeated units may also involve several kb in FSXD patients (27), the FSHD-associated deletions of 3.3 kb repeated units may also involve several hundred kb. Therefore, the FSHD candidate gene region subjected to several gene identification methods was chosen to extend about 150 kb proximal to the 3.3 kb repeats.

Here, we report the isolation and characterization of a novel gene cloned from the FSHD region (FRGI). The PEV hypothesis was tested by analysis of allele-specific transcription of this gene in lymphocytes and muscle tissue of patients and controls.

**RESULTS**

**Isolation and sequence analysis of FRGI cDNAs**

Previous mapping studies (28,29) showed the presence of a CpG island located within a 6.5 kb EcoRI fragment of cosmid c171 which was isolated from a cosmid library made from YAC 25C2E (DS81093) (28). A total of 6.2 kb of contiguous sequence from this 6.5 kb EcoRI fragment was determined (GenBank accession number L76174) by sequencing adjacent PstI subclones located within this fragment. The CpG island spans a region of 1 kb showing a G+C content of 63% and a CpG:GpC ratio of 0.8. A 1.2 kb PstI subclone (p10B) encompassing the CpG island (Fig. 1), was used as a probe to screen a human skeletal muscle cDNA library. A 640 bp cDNA clone (10B7, Fig. 1) was isolated and, after sequencing, shown to contain the 3' end of a cDNA. Clone 10B7 hybridized to a 1.1 kb transcript in total RNA from lymphocytes and fetal muscle, and poly-(A)+ RNA from adult muscle, fetal brain and placenta (Fig. 2). Subsequently, two cDNA clones were obtained by screening a total infant brain cDNA library (BS3; 3.5 kb insert), and a fetal brain cDNA library (10B1; 1.3 kb insert), with probe 10B7 (Fig. 1). These clones were sequenced and their unexpectedly large cDNA inserts were shown to be due to splicing aberrations and cloning artefacts (see below). Additional cDNA fragments were obtained by amplification of different cDNA libraries using one of the cDNA library vector-specific primers flanking the cloning site, in combination with 10B7-based primers. Two cDNA fragments, PLC2FB22 (700 bp) and ML8RT3 (780 bp) were amplified from the placenta and muscle cDNA library, respectively. By sequence analysis these fragments were found to overlap, spanning 913 bp (Fig. 1). To determine the transcription start site of the cDNA, two primers at the 5' end of the cDNA (1P1 and 2r) were designed to perform 5' RACE (rapid amplification of cDNA ends (30)) from human brain cDNA. The obtained RACE products were sequenced and found to start at, or just distal of, position −191 upstream from the putative translation initiation site (Fig. 3b). Subsequently, primers (1f and 9r) were chosen at positions −134 and 801 to perform RT-PCR on RNA isolated from a chromosome 4 only cell hybrid, muscle tissue, and lymphocytes. In all cases, an identical 935 bp product was obtained which is in accordance with the total transcript length of about 1.1 kb detected by Northern blot analysis. Further sequence analysis confirmed the identity of these RT-PCR products. The complete cDNA sequence of this novel human gene, named FRGI (FSHD Region Gene 1), was compiled (GenBank accession number L76159); spanning 1042 bp and containing an open reading frame of 258 amino acid residues (Fig. 3b). The putative translation start site is the first in frame ATG codon located 191 bp downstream of the 5' end of the transcript, which conforms to the Kozak consensus sequence (31). The 3' untranslated region is 74 bp and ends with a poly-(A) tail 25 bp downstream of a polyadenylation signal at position 821 (Fig. 3b). Nucleotide sequence database searches (BLAST-N) identified several highly homologous human EST clones: a human HL60 EST clone (accession number D19665), and from the WashU-Merck EST Project; accession numbers T41211, T51111, T51019; all being 96-99% homologous to the 3' end of the cDNA, and R77057; 95% homologous to the internal part of the cDNA. Furthermore, a mouse EST clone (GEG-123, accession number X71639) showed high homology (86%) to the 3' end of the FRGI transcript. Southern blot analysis also indicated that FRGI is evolutionarily conserved (Fig. 4a). No significant homology to known genes was detected in either the nucleotide sequence databases or the protein databases (BLAST-P). Protein motif searches gave no indication of a possible function for the deduced FRGI protein.

**Genomic organization of FRGI**

Southern blot analysis revealed that the FRGI cDNA is distributed over three EcoRI fragments contained within cosmid c171 (Fig. 1, Fig. 4b). These EcoRI fragments were subcloned and partially sequenced using FRGI cDNA-based primers. The 1042 bp cDNA sequence was found to be completely identical to the genomic sequence and is composed of nine exons (Fig. 1). The sequences of the intron-exon boundaries were determined and shown to be in accordance with splice site consensus sequences (Table 1). The promoter region of FRGI (Fig. 3a, GenBank accession number L76173) coincides with the identified CpG island (see above). No TATA or CAAT box motifs were found, but three putative Sp1 binding sites (32) are present at positions −743, −403, and −310, which may function as control elements of transcription (33). Absence of a TATA box in genes with a CpG island has been reported for many housekeeping genes and appears to be related to the regulation of widely expressed genes (34).

**Chromosomal distribution of FRGI**

The 935 bp chromosome 4 FRGI RT-PCR product hybridized to eight different human genomic EcoRI fragments with varying
intensities (Fig. 4). Only three of these fragments are derived from the chromosome 4q35 region (Fig. 1, Fig. 4b), suggesting the presence of multiple FRG1-related sequences in the genome. To identify the chromosomal origin of the different members of this novel multigene family, the three EcoRI fragments of cosmid cT171 containing the FRG1 exons were independently used as probes for fluorescent in situ hybridization (FISH) to metaphase spreads. For each fragment an identical hybridization pattern was detected. Hybridization signals were found at chromosome 4q35, at three sites in the pericentric region of chromosome 9, at the short arm of all acrocentric chromosomes (13, 14, 15, 21, and 22) and at the centromeric region of chromosome 20 (Fig. 5). Subsequent PCR analyses of DNA from a monochromosomal somatic cell hybrid mapping panel using several primer combinations, confirmed the chromosomal dispersion of FRG1-related sequences. With most primer combinations, appropriate PCR products could be amplified from cell hybrids containing one of the human chromosomes 4, 8, 9, 12, 13, 14, 15, 20, 21, and 22. To illustrate this, amplification of a fragment containing part of intron 4 and exon 5 using primers 5f and 5r, is shown in Figure 6a. The appropriate 320 bp product was detected in somatic cell hybrids containing chromosomes 4, 8, 9, 12, 20, 21 or 22. However, this product could not be amplified from cell hybrids containing chromosomes 13, 14, or 15, which may be an indication for sequence differences or the presence of processed FRG1-like pseudogenes. Amplification of the 320 bp fragment from cell hybrids containing chromosomes 8 or 12 does not correspond to the negative FISH results, suggesting that these chromosomes may harbor incomplete FRG1-related sequences not detectable by FISH.

Unlike most FRG1 primers, primer 1f derived from the 5' untranslated region of FRG1, enabled chromosome 4 specific amplification of FRG1 fragments, when used in several primer combinations. For example, using primers 1f and 1r, a 180 bp fragment in exon 1 could only be amplified from DNA of the single chromosome 4 cell hybrid (Fig. 6b). In addition, RT-PCR analysis of total RNA isolated from somatic cell hybrids specific
individuals without significant deviations in the amount of transcript (data not shown). However, possible transcription of FRG1-related sequences may mask any effects on the transcription of the chromosome 4q35 gene. Therefore, detection of alteration of the FRG1 transcript level in FSHD patients compared to unaffected individuals requires the 4q35-focus to be analyzed specifically. As was indicated by the above described genomic and cDNA PCR analyses of the monochromosomal cell hybrids (Fig. 5c), 4q35-specific amplification of FRG1 fragments is possible using primer If. Additional evidence for the specificity of this primer was obtained by SSCP-analysis of a Dutch FSHD family (see below).

Since FSHD is caused by a mutation on only one allele, the complete FRG1 cDNA was screened for polymorphisms to enable analysis of allele-specific transcripts. Primary, reverse transcribed lymphocyte RNA of a number of unaffected individuals was amplified using primers If and 9r to ensure 4q35-specific products and to avoid amplification of potential DNA contaminations. These 935 bp products were used in secondary, nested PCRs. Three overlapping cDNA fragments were obtained using primer combinations 1f/9r, 2f/5r, and 5f/9r (Fig. 1), which were analyzed for sequence variation by single-strand conformation polymorphism analysis (SSCP) (35,36). Subsequent sequence analysis was performed to characterize the variant fragments. So far, five polymorphic sites were observed: three in the 5' untranslated region at positions −82 (C/T), −62 (C/T), and −10 (G/T), one in exon 1 at position 18 (TAT/TAC; Tyr18Tyr), and one in exon 5 at position 321 (ATT/ATC; Ile321Ile). These positions are indicated in Figure 3b. The polymorphisms at positions −82 and −62, which can be analyzed by amplifying cDNA or DNA using primer combination 1f/If, were used to demonstrate co-segregation of the polymorphisms with the chromosome 4q35 haplotype in DNA from part of the large Dutch FSHD family 1 (4). Four alleles were observed which segregate for chromosome 4, 8, 9, 10, 12, 14, 15, 20, 21, or 22, using primer If in combination with primer 9r, revealed the 935 bp cDNA fragment in the chromosome 4 cell hybrid only (Fig. 6c).

**Patient analysis**

Northern blot analysis of total RNA isolated from lymphocytes of four sporadic FSHD patients with a *de novo* deletion and their unaffected parents, revealed a 1.1 kb FRG1 transcript in all

---

Figure 3. (a) Sequence of the GC-rich promoter region of FRG1 (accession number L76173). Three putative Sp1 binding sites are indicated in bold, italic letters. (b) Sequence of the FRG1 cDNA (accession number L76159). The putative transcription start site is at position −191. The predicted amino acid sequence encoded by the FRG1 cDNA is shown below the DNA sequence, assuming that translation begins at the first in-frame methionine of the ORF. Arrowheads (>) indicate the first base pair of the next exon following the exon boundary. The detected polymorphisms in exons 1 and 5 are depicted in bold, italic, and underlined. The asterisk indicates the termination codon.

Figure 4. Southern blots containing (a) EcoRI digested DNA of human, monkey, sheep and rat, and (b) EcoRI and HindIII digested DNA of total human DNA, YAC 25C2E and cosmid cT171, both hybridized with the 935 bp FRG1 RT-PCR product (see text). The sizes of the hybridizing chromosome 4-specific EcoRI fragments are indicated (see Fig. 1).
in the family in a Mendelian fashion and show complete co-segregation with the FSHD haplotype, as determined by the loci D4S163 and D4S139 (Fig. 7).

Amplification of the 4q35-specific polymorphic fragment in exon 1 was used to screen reverse transcribed lymphocyte RNA from the four sporadic patients mentioned above and their unaffected parents. After primary amplification with primers 1F and 9R and subsequent nested amplification using primers 1F and 1R², electrophoresis on a SSCP gel showed equal levels of allele-specific FRG1 transcripts in patients vs controls (Fig. 8a). Subsequently, we considered the possibility that a position effect affecting the transcription level of FRG1 may only be detectable in muscle tissue. Therefore, RNA was isolated from open muscle biopsies of three familial FSHD patients, one sporadic patient with a de novo mutation, and one unrelated unaffected individual; all shown to be heterozygous for the exon 1 polymorphisms (further details of the biopsies, and of the associated p13E-11 fragment sizes are described in Materials and Methods). The exon 1 polymorphic fragment was amplified as described from reverse transcribed lymphocyte and muscle RNA of the four patients as well as from lymphocyte and muscle control cDNA. However, SSCP analysis again showed equal transcript levels for both alleles of FRG1 in muscle tissue of patients, as was also detected in the corresponding patient lymphocytes and in the control lymphocyte and muscle samples (Fig. 8b). To exclude the possibility that the SSCP-based screening method would not be sufficiently sensitive for detection of small alterations in transcript amount, we have tested the method by mixing two homozygous controls in various template ratios. A 30% reduction of transcript level could easily be detected using this SSCP based method (data not shown).

**DISCUSSION**

FSHD has been shown to be associated with deletions within an EcoRI fragment detectable using probe p13E-11 (D4F104S1) (9). These deletions involve integral numbers of 3.3 kb repeated units (9,10) which are located adjacent to the 4q telomere (13,15,16). Since no transcribed sequences were identified within these repeated units (11–14), the FSHD gene search has been focused on the region more centromeric. Deletion of a number of the repeated units may induce position effect variegation (PEV), reducing allelic transcript levels of proximally located genes. Recently, PEV has also been proposed as a potential genetic cause for aniridia (37).

The FSHD candidate region was scrutinized for genes using several gene identification methods. However, the region appeared to be comprised of different kinds of repetitive elements and regions homologous to several loci dispersed over the genome (38). This property has seriously complicated the search for candidate genes, in particular using hybridization-based gene identification methods such as cDNA selection (38). For this reason, other gene identification methods were applied such as exon trapping and screening of cDNA libraries with specifically selected genomic subclones. Several pseudogenes were identified, each originating from a different chromosome (Hewitt, Van

---

**Figure 5.** Fluorescent in situ hybridization of the 9.5 kb EcoRI fragment of c171 to metaphase spreads of normal lymphocytes. Signals were observed at both chromosomes 4q35 and by cross-hybridization, at chromosomes 9, 13, 14, 15, 20, 21, and 22. Only chromosomes which could unambiguously be identified, are numbered in the figure.

**Figure 6.** PCR analysis of FRG1-related sequences. (a) Exon 5 was amplified from DNA of a human monoclonal mapping panel using intron primer 5P in combination with primer 5R. The appropriate 320 bp product was detected in cell hybrids containing human chromosomes 4, 8, 9, 12, 20, 21, or 22. The smaller product detected in the chromosome 3 cell hybrid is probably an aspecific product, since all other primer combinations revealed negative results for this chromosome. Total human DNA (Hu) and DNA derived from Chinese hamster lung cells V79 (V) and mouse A9 (A) cells, representing the cell line backgrounds was also analyzed. M: molecular weight marker. (b) Amplification of part of exon 1 (180 bp) from DNA of the mapping panel using primer combination 1F–1R², was successful only for the chromosome 4 somatic cell hybrid, demonstrating the 4q35 specificity of primer 1F. (c) RT-PCR analysis of RNA isolated from the chromosome 4, 8, 9, 10, 12, 14, 15, 20, 21, and 22 cell hybrids: using primers 1F and 9R; the 935 bp cDNA fragment could be amplified from the chromosome 4 cell hybrid only.
Panel Material

Materials and Methods

The PHO family genes are part of the PHO regulon, which is responsible for the pH-dependent expression of the PHO genes. In yeast cells, the PHO genes are inducible by high pH, and their expression is controlled by the PHO regulon. The PHO regulon consists of several genes, including PHO1, PHO4, PHO5, and others, which are involved in the regulation of pH-dependent gene expression.

The PHO regulon is activated by the high pH, which leads to the activation of the PHO4 gene. The PHO4 gene encodes a transcription factor that binds to the promoter regions of the PHO genes, thereby activating their expression. The PHO4 gene is activated by the PHO2 gene, which encodes another transcription factor. The PHO2 gene is activated by the PHO5 gene, which encodes a protein that is involved in the perception of pH changes.

The regulation of the PHO regulon is complex and involves several mechanisms, including transcriptional and post-transcriptional regulation. The PHO regulon is also implicated in several cellular processes, including the synthesis of amino acids and the utilization of carbon sources.

References


biopsy taken from patient 10403. Lymphocytes were also isolated from these four patients.

DNA of part of the large Dutch FSHD family 1 (4) was used to study co-segregation of the exon 1 polymorphisms with the chromosome 4 haplotype.

Somatic cell hybrid panel

DNA from a human monochromosomal somatic cell hybrid panel was available to determine the chromosomal origin of FRG1-related sequences (UK HGMP Resource Centre). RNA was isolated from somatic cell hybrids specific for chromosome 4, 8, 9, 10, 12, 14, 15, 20, 21, and 22 (41–44).

cDNA isolation

A human skeletal muscle cDNA library (ΔZAPII, Stratagene) was screened with a α-32P labelled genomic probe p10B (Fig. 1), which is a 1.2 kb genomic PstI subclone of cosmid c171 (28). Phase clone 10B7 was identified and the insert (640 bp) was recovered by amplification using the T7 and T3 vector primers (T7: 5′-AATAC-GACTCCTAGAT-3′, T3: 5′-ATTAACCTCTACAAAG-3′). The PCR product was cloned into the pCR™I vector using the TA Cloning Kit (Invitrogen) and subsequently sequenced. A human fetal brain cDNA library (HL1065b, ClonTech) and a total infant brain cDNA library (Soares, LLNL Library id 52) were screened using α-32P labelled 10B7 as probe. Positive phage clones 10B1 (fetal brain cDNA) and BS3 (infant brain cDNA) were isolated and the EcoRI inserts (1.3 kb and 3.5 kb, respectively) were subsequently subcloned in pBluescriptIIISK+ (Stratagene) and further characterized. For each phage library 5×10⁵ plaque forming units were screened. Hybridizations were carried out for 16 h at 65°C as described (45). In addition, by amplifying cDNA libraries with cDNA-vector-specific primers in combination with 10B7-specific primers, two cDNA fragments were obtained: (i) PLC/2F22 was amplified from the human placenta cDNA library (HL1008, ClonTech) using 10B7-specific primer 2F (Fig. 1, Table 1) and ΔZ11-specific primer B22 (5′-GGTGGCGACGACTCGTGG-AG-3′); and (ii) ML8RT3 was amplified from the skeletal muscle cDNA library using 10B7-specific primer 8RT (Fig. 1, Table 1) and ΔZAPII-specific primer T3.

To isolate the 5′ end of the cDNA, the 5′-RACE-Ready Human Brain cDNA kit (ClonTech) was used in combination with one of the primers 1R and 2R (Fig. 1, Table 1). The PCR fragments generated were cloned in the TA cloning vector (Invitrogen) and subsequently sequenced.

Sequence analysis

All sequence reactions were performed according to the dyeoxy method (46), using either the Sequenase Kit (Pharmacia) or Sequenase V2.0 (Amersham). Genomic and cDNA sequences were compared and/or characterized using computer programs BLAST-N (Basic Local Alignment Search Tool) (47), GDE V2.2 (Genetic Data Environment, S. Smith, University of Illinois), and the GCG package of programmes [Genetics Computer Group (1994)]. The putative protein sequence was analyzed using BLAST-P.
**Table 1. Sizes, splice junctions, and primers of FRG1 exons**

<table>
<thead>
<tr>
<th>Exon no.</th>
<th>Nucleotide position</th>
<th>Exon size (bp)</th>
<th>3' splice site</th>
<th>5' splice site</th>
<th>Forward primer</th>
<th>Reverse primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-191*-62</td>
<td>253</td>
<td>GAG/gtagc</td>
<td>5'-TCTACAGAGACGTAGGCTGTC-3'</td>
<td>1f: 5'-TCTACAGAGACGTAGGCTGTC-3'</td>
<td>1r: 5'-CCGGTTTCTGGAGCGATGT-3'</td>
</tr>
<tr>
<td>2</td>
<td>63-133</td>
<td>71</td>
<td>ttgag/TAAG</td>
<td>5'-TTTCCCTTTGTCTTCTTCTGG-3'</td>
<td>2f: 5'-TTTCCCTTTGTCTTCTTCTGG-3'</td>
<td>2r: 5'-CAACAATCTCAGGGTGGTTCTC-3'</td>
</tr>
<tr>
<td>3</td>
<td>134-259</td>
<td>126</td>
<td>AAG/GAGAG</td>
<td>5'-CCGCTGTGAGAATGGTATGT-3'</td>
<td>3f: 5'-CCGCTGTGAGAATGGTATGT-3'</td>
<td>3r: 5'-GCCATCTCAGGGTGGTTCTC-3'</td>
</tr>
<tr>
<td>4</td>
<td>260-317</td>
<td>58</td>
<td>CAG/gtgag</td>
<td>5'-GCCAATTTGAAATGGGTTAG-3'</td>
<td>4f: 5'-GCCAATTTGAAATGGGTTAG-3'</td>
<td>4r: 5'-GAAGAAGATGGGTTAGGCTGTC-3'</td>
</tr>
<tr>
<td>5</td>
<td>318-432</td>
<td>115</td>
<td>CAG/gtgag</td>
<td>5'-CAAGATTAGATCCTGTGCTG-3'</td>
<td>5f: 5'-CAAGATTAGATCCTGTGCTG-3'</td>
<td>5r: 5'-GCCATTGAAATGGATAAGGG-3'</td>
</tr>
<tr>
<td>6</td>
<td>433-537</td>
<td>105</td>
<td>CAG/gtgag</td>
<td>5'-TTTCCCTTTGTCTTCTTCTGGG-3'</td>
<td>6f: 5'-TTTCCCTTTGTCTTCTTCTGGG-3'</td>
<td>6r: 5'-AAATCCATCTTTCGAGGCC-3'</td>
</tr>
<tr>
<td>7</td>
<td>538-629</td>
<td>92</td>
<td>CAG/gtgag</td>
<td>5'-GCCATTGAAATGGATAAGGG-3'</td>
<td>7f: 5'-GCCATTGAAATGGATAAGGG-3'</td>
<td>7r: 5'-AAATCCATCTTTCGAGGCC-3'</td>
</tr>
<tr>
<td>8</td>
<td>630-740</td>
<td>111</td>
<td>CAG/gtgag</td>
<td>5'-CAAGATTAGATCCTGTGCTG-3'</td>
<td>8f: 5'-CAAGATTAGATCCTGTGCTG-3'</td>
<td>8r: 5'-AAATCCATCTTTCGAGGCC-3'</td>
</tr>
<tr>
<td>9</td>
<td>741-851</td>
<td>111</td>
<td>CAG/gtgag</td>
<td>5'-GCCATTGAAATGGATAAGGG-3'</td>
<td>9f: 5'-GCCATTGAAATGGATAAGGG-3'</td>
<td>9r: 5'-CAAGATTAGATCCTGTGCTG-3'</td>
</tr>
</tbody>
</table>

*Prime transcriptional start site

**Southern blot analysis**

The 935 bp FRG1 RT-PCR product, which was amplified from a chromosome 4 only somatic cell hybrid using primers 1f and 9r, was hybridized to Southern blots containing: (i) EcoRI digests of total human (2 μg/lane), monkey, sheep, and rat (8 μg/lane) (‘Zoo blot’, Fig. 4a); and (ii) EcoRI and HindIII digests of total human lymphocytes DNA (7 μg/lane), YAC 25C2E (D4S1093) (1.5 μg/lane), and cosmid cT171 (28) (1.5 ng/lane) (Fig. 4b). Digestions with specified restriction enzymes were performed according to the manufacturer’s instructions (Pharmacia). Hybridizations were performed for 16 h at 65 °C for the Southern blot and at 57 °C for the ‘Zoo’ blot, as described (48). The blots were washed at 65 °C and to a stringency of 0.3 x SSC/0.1% SDS for the Southern blot and of 1x SSC/0.1% SDS for the ‘Zoo’ blot and, followed by autoradiography for 16-36 h at -70 °C using Konica AX film with an intensifying screen.

**RNA isolation and Northern blot analysis**

Total RNA was isolated from various human tissues and somatic cell hybrids as previously described by Chomczynski and Sacchi (49). Poly-(A)+ RNA was isolated using the PolyATtract mRNA Isolation Systems kit (Promega) or purchased from ClonTech. cDNA was synthesized from 5 μg of total RNA isolated from human muscle, human lymphocytes, somatic cell hybrids specific for the particular chromosomes, Chinese hamster lung cells (V79) and mouse A9 cells, using an oligo-dT primer and the AMV Reverse Transcriptase kit from Promega. The primers used for (RT-) PCR analyses are indicated in Figure 1 and Table 1. All primers were provided by Isogen Bioscience BV, Maarssen. The Netherlands. A total of 35 cycles (94 °C for 40 s, 60 °C for 60 s and 72 °C for 90 s) were run in a 50 μl volume containing 2.5 ng of cDNA, 1.0 U Taq-DNA-polymerase (Perkin-Elmer), 200 μM of each dNTP, 100 ng of each primer and 5X PCR standard-buffer [300 mM Tris-HCl, 75 mM (NH4)2SO4, 17.5 mM MgCl2, pH 8.5].

**Fluorescent in situ hybridization**

Metaphase chromosome spreads from normal human lymphocytes were hybridized with the 3 kb, 6.5 kb, and 9.5 kb EcoRI fragments of cosmid cT171. Probes were labelled by nick translation with biotin 11-dUTP (Sigma) (50). Hybridization, washing and staining were performed as previously described (8).

**Allele-specific analysis of FRG1 steady-state transcript levels** (RNA-based SSCP)

cDNA of lymphocytes and muscle tissue was primarily amplified using primers 1f and 9r (Fig. 1, Table 1). For the SSCP analysis (35,36), PCR products were diluted 1:50 of which subsequently 1 μl was used for secondary amplifications with primer combinations 1f/1r, 1f/2r, 2f/6r or 5f/9r (Fig. 1, Table 1) and with incorporation of α-32P labelled dCTP. Products were diluted 1:25 in a formamide loading buffer (1:1 mix of 95% formamide/20 mM EDTA and 0.1% SDS/10 mM EDTA) and denatured at 100 °C for 10 min. Electrophoresis was performed in non-denaturing 0.5X MDE polyacrylamide gels (J.T. Baker) and 0.6x TBE buffer (1X: 0.09 M Tris, 0.09 M Boric acid, 0.002 M EDTA, adjusted to pH 8.0) using a BRL-S2 vertical gel apparatus (Life Technologies, Inc.). Gels of 40 cm were run at 25 Watt for 3-4 h and subsequently dried at 80 °C under vacuum, and exposed at -70 °C for 12-24 h to Konica AX film with an intensifying screen.

**ACKNOWLEDGEMENTS**

The authors thank R. Lyle and C. Wijmenga for comments on the manuscript. This study was funded by The Prinses Beatrix Fonds, The Muscular Dystrophy Group of Great Britain, The Muscular
Dystrophy Association (USA), and the Association Française contre les Myopathies.

REFERENCES


