Neurogenetics (1998) 2:34–42
DOI 10.1007/s100489800049


*Original article*

# Refined mapping and characterization of the recessive familial amyotrophic lateral sclerosis locus (*ALS2*) on chromosome 2q33

**Betsy A. Hosler · Peter C. Sapp · Ralph Berger · Gilmore O'Neill · Khemissa Bejaoui Mongi Ben Hamida · Faycal Hentati · Wendy Chin · Diane McKenna-Yasek Jonathan L. Haines · David Patterson · H. Robert Horvitz · Robert H. Brown Jr**

Accepted: July 15, 1998 / Published online: October 28, 1998

## ABSTRACT

**Amyotrophic lateral sclerosis (ALS) is a progressive degenerative neuromuscular disease that shows familial, autosomal dominant inheritance in 10%–15% of cases. Previous genetic analysis of one large family linked a recessive form of familial ALS (FALS-AR type 3) to the chromosome 2q33–35 region. Using additional polymorphic markers, we have narrowed the size of the linked region to approximately 1.7 cM by linkage and haplotype analysis. We have also established a yeast artificial chromosome contig across the locus that covers an approximate physical distance of 3 million bases. Based on this contig, genes and expressed sequences that map near the 2q33 region have been examined to determine whether they are located within this *ALS2* candidate locus. Five identified genes and 34 expressed sequence tags map within the region defined by crossover analysis and merit further consideration as candidate genes for this disease.**

**Key words** Amyotrophic lateral sclerosis · Yeast artificial chromosome contig · Mapping

B.A. Hosler · P.C. Sapp · G. O'Neill · K. Bejaoui · W. Chin · D. McKenna-Yasek · R.H. Brown Jr (✉)
Cecil B. Day Laboratory for Neuromuscular Research, Neurology Service, Massachusetts General Hospital, MGH-East, Building 149, 13th Street, Charlestown, MA 02129, USA

P.C. Sapp · W. Chin · H.R. Horvitz
Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

R. Berger · D. Patterson
Eleanor Roosevelt Institute, Denver, CO 80206, USA

M. Ben Hamida · F. Hentati
Institut National de Neurologie, 1006 Tunis, Tunisia

J.L. Haines
Vanderbilt University Medical Center, Nashville, TN 37232, USA

D. Patterson
University of Colorado Health Sciences Center, Denver, CO 80262, USA

## INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is a progressive, degenerative disease that affects both upper and lower motor neurons. Although most cases of ALS occur in unrelated individuals, both dominant (FALS-AD) and recessive (FALS-AR) inherited forms of the disease have been observed [1, 2]. A subset of FALS-AD cases are associated with mutations in the gene encoding cytosolic Cu,Zn superoxide dismutase (SOD1) [3, 4]. Mutations in the neurofilament heavy chain gene have been identified in rare cases [5], but their relationship to the etiology of ALS is unclear. The genetic causes of the other inherited ALS cases have not been identified.

FALS-AR can be divided into three clinical subtypes [2]. Type 1 involves mild spasticity of all limbs and amyotrophy of the upper limbs. Symptoms (spasticity and peroneal atrophy) are primarily in the lower limbs in type 2. In type 3, prominent spasticity of the facial and limb muscles is accompanied by distal amyotrophy of the upper and lower extremities.

Previous genetic analysis linked FALS-AR to the chromosome 2q33–35 region in one large consanguineous family with type 3 disease [6]. Significant log of the odds (LOD) scores were obtained with markers D2S117, D2S116, D2S72, and D2S155, and crossover analysis delimited an 11-cM candidate region for the responsible gene, designated *ALS2* [6]. If one assumes that the same ancestral chromosome was inherited through the several branches of this family, the locus is further limited to an 8-cM region by a presumptive an-

cestral recombination event. We have now narrowed the region to approximately 1.7-cM using additional polymorphic markers in this family and have established a yeast artificial chromosome (YAC) contig across the region. Based on this contig, genes and expressed sequence tags (ESTs) that map near the 2q33 region have been examined to determine whether they are present within the *ALS2* candidate region and merit further consideration as candidate genes for FALS-AR.

## MATERIALS AND METHODS

### Pedigree analysis

The clinical symptoms of the FALS-AR family have been previously described [2, 6]. The FASTLINK (version 3.0P) set of analysis programs was used to calculate two-point LOD scores for the FALS-AR locus and each polymorphic marker examined [7–12]. Autosomal recessive inheritance with complete penetrance and a disease allele frequency of 0.0001 were assumed [6].

### STS analysis

Polymerase chain reaction (PCR) was performed using Taq polymerase (Perkin Elmer, Foster City, Calif.; Fisher Scientific, Pittsburgh, Pa., USA) following the manufacturer's instructions. Primers for PCR were selected with the help of the program Oligo (version 4.0, National Biosciences, Plymouth, Minn., USA) and obtained from several sources (GIBCO BRL, Grand Island, N.Y.; Integrated DNA Technologies, Coralville, Iowa; Genosys Biotechnologies, The Woodlands, Tex.; Research Genetics, Huntsville, Ala., USA). Human genomic DNA and a YAC (CEPH 747H8) that does not contain DNA from the 2q33 region were used as positive and negative controls, respectively, for all markers tested.

### YAC and PAC analysis

YAC A188E12 was from the library described by Burke et al. [13]. All other YACs were from the Centre d'Étude du Polymorphisme Humain (CEPH) mega YAC libraries [14]. Alu fingerprinting data and initial map and marker information were obtained from databases maintained and made publicly available by Genethon [15, 16] and the Whitehead Institute Center for Genomic Research [17].

Primers AOX F7 (5′-GAG-CAA-TTC-CTC-AGC-AAG-TGC-C-3′), AOX R6 (5′-CCT-TCC-AAT-GAG-TTT-CTG-GCA-GG-3′), and AOX R12 (5′-CCG-AGC-CCA-AAA-GGG-AGA-CTT-C-3′) were derived from the aldehyde oxidase (AOX) cDNA sequence (GenBank entry L11005 [18]). AOX F7 corresponds to bases 1310–1331, while AOX R6 and AOX R12 are complementary to bases 1557–1579 and bases 1628–1649, respectively. Primer AOX F101 (5′-CCA-GTG-CAT-AGG-TGA-TGC-TCA-G-3′) was derived from the approximately 1.5- kilobase (kb) fragment amplified by primers AOX F7 and AOX R6 from YAC A188E12 and is positioned near the 3′ end of the *AOX* intron contained in this fragment. The 242-base pair (bp) fragment amplified by primers AOX F101 and AOX R6 from YAC A188E12 was used to probe a chromosome 2-specific P1-derived artificial chromosome (PAC) library [19] using standard techniques [20]. An *Eco*RI digest of one of the PACs identified in this screening, PAC 5I17, was subcloned into vector pZERO-2.1 (Invitrogen, Carlsbad, Calif., USA) and similarly probed with the YAC end-clone STS GM-21 (Table 1).

YAC and PAC end fragments were amplified using vectorette PCR [21]. Inter-Alu fragments were generated from the indicated YACs (Table 1) by PCR with either primer 47-23 or primers S-J [22] as described previously [23]. DNA sequences were defined using Sequenase following the manufacturer's instructions (United States Biochemical Co., Cleveland, Ohio, USA). Some samples were analyzed on a Model 4200L automated DNA sequencer (LiCor, Lincoln, Neb., USA). The sequences and primers for markers developed in this study can be found in GenBank and dbSTS (Table 1).

The ends of PAC 5I17 and PAC 12J15 overlap by approximately 1.7 kb. GM-16 (200-bp STS marker at the T7 end of PAC 5I17) and GM-18 (190 bp STS marker at the T7 end of PAC 12J15) are both present in the fragment whose sequence is given in GenBank entry G36356. The primers for these two markers are: GM-16F (5′-TGA-GGT-GCT-CTG-TTC-TCT-TTG-3′), GM-16R (5′-TGG-AAG-AAA-GCA-AAA-TGA-ATG-C-3′), GM-18F (5′-CAG-GCA-GGA-AAG-GCA-AAC-CTC-3′), and GM-18R (5′-CTG-ATG-CTG-ATG-CTG-AAC-3′).

### Polymorphism analysis

Marker GM-1, a polymorphic $(CA)_N$ repeat, was tested in 106 control DNA samples. The observed allele sizes (and frequen-

**Table 1** New markers in the FALS-AR (familial amyotrophic lateral sclerosis-autosomal-recessive) region (*YAC* yeast artificial chromosome, *PAC* PI-derived artificial chromosome, *PCR* polymerase chain reaction)

| Marker | GenBank ID | dbSTS ID | Type | Source |
|---|---|---|---|---|
| GM-1[a] | G31889 | 48838 | $(CA)_N$ | A188E12 |
| GM-4 | G31874 | 48823 | Alu 47-23 PCR | 168H9 |
| GM-5 | G31875 | 48824 | Alu 47-23 PCR | 168H9 |
| GM-7 | G31876 | 48825 | Alu S-J PCR | 168H9 |
| GM-8 | G31877 | 48826 | Alu 47-23 PCR | 964G9 |
| GM-10 | G31878 | 48827 | Alu 47-23 PCR | 964G9 |
| GM-13 | G31879 | 48828 | Alu S-J PCR | 925A3 |
| GM-15 | G31880 | 48829 | Alu S-J PCR | 841A8 |
| GM-16 | G36356 | 53347 | PAC end | PAC 5I17 |
| GM-17 | G35367 | 53348 | PAC end | PAC 5I17 |
| GM-18 | G36356 | (see Methods) | PAC end | PAC 12J15 |
| GM-20 | G31881 | 48830 | YAC end | 796C3 |
| GM-21 | G31882 | 48831 | YAC end | 983G8 |
| GM-25 | G31883 | 48832 | Alu S-J PCR | 893C5 |
| GM-26 | G31884 | 48833 | YAC end | 845E4 |
| GM-28 | G31885 | 48834 | YAC end | 865E11 |
| GM-41 | G31886 | 48835 | YAC end | 848B7 |
| GM-42 | G31887 | 48836 | YAC end | 983G8 |
| GM-43 | G31888 | 48837 | YAC end | 168H9 |
| GM-124[a] | G36358 | 53349 | PAC subclone | PAC 5I17 |

[a] These markers are polymorphic

cies) are 239 bp (0.057), 237 bp (0.175), 235 bp (0.160), 233 bp (0.217), 231 bp (0.184), and 225 bp (0.208).

GM-124 was amplified from individuals VI:10 and VI:14, and the PCR products were directly cloned into vector pCR2.1 (Invitrogen). For each individual, DNA was isolated from 8 separate colonies and sequenced. The polymorphism identified in this fragment was examined in the family by single-strand conformation polymorphism analysis (SSCP) by the method described previously [3]. The primers for SSCP, GM-126F (5′-CAA-TTA-GGC-CAC-TAT-TCT-ATG-TG-3′) and GM-126R (5′-CCC-AGG-AAT-GTA-TTT-TGA-TTT-ATG-3′), are located within the GM-124 fragment and amplify a 330-bp fragment.

Two members of the family (affected individual VI:1 and carrier individual V:3) were screened for alterations in two known genes. DNA from a normal human lymphoblast cell line (CGM) was used as a control. Primers placed in intronic regions were used to amplify each of the three exons encoding CTLA-4 [24] and the four exons encoding CD28 [25]. The primer sequences are available from the authors (R.B. and D.P.) upon request.

## RESULTS

### Generation of new markers

To assist our investigation, we generated new markers across the *ALS2* locus. Those markers that proved useful for the work discussed here are listed in Table 1. Two of these markers, GM-1 and GM-124, are polymorphic, and help to define the boundaries of the locus, as described below. Eight STSs were selected from fragments isolated by Alu PCR. Ten others were generated from the ends of the YACs and PACs used in our study.

### Linkage analysis

Previous linkage analysis of one large consanguineous family with type 3 recessive FALS produced positive LOD scores for the markers D2S116, D2S72, and D2S155 [6]. To confirm and extend this work, we tested additional polymorphic markers between D2S115 and D2S157 in this family and determined linkage scores (Table 2). All markers tested from this region showed positive linkage to FALS-AR. The scores also imply the occurrence of crossover events within this family in this extended region (e.g., LOD = $-\infty$ at $\Theta=0.00$ for markers D2S115, D2S1367, and GM-1). These crossovers, which are more evident when haplotype analysis is applied to the pedigree, help define the limits of the disease locus.

**Table 2** Two-point LOD scores for FALS-AR versus chromosome 2q markers

| Locus | Recombination fraction ($\Theta$) | | | | | | |
|---|---|---|---|---|---|---|---|
| Cent | 0.00 | 0.05 | 0.10 | 0.20 | 0.30 | $Z_{max}$ | $\Theta_{max}$ |
| D2S115 | $-\infty$ | 2.14 | 2.25 | 1.82 | 1.13 | 2.26 | 0.087 |
| D2S1367 | $-\infty$ | 3.33 | 3.13 | 2.41 | 1.56 | 3.34 | 0.044 |
| GM-1 | $-\infty$ | 3.92 | 3.99 | 3.18 | 1.99 | 4.15 | 0.078 |
| D2S116 | 5.29 | 4.73 | 4.15 | 2.96 | 1.77 | 5.29 | 0.00 |
| D2S346 | 8.38 | 7.51 | 6.63 | 4.78 | 2.87 | 8.38 | 0.00 |
| D2S72 | 8.14 | 7.28 | 6.39 | 4.55 | 2.65 | 8.14 | 0.00 |
| D2S2189 | 7.98 | 7.12 | 6.24 | 4.41 | 2.56 | 7.98 | 0.00 |
| D2S2237 | 4.60 | 5.27 | 4.85 | 3.63 | 2.22 | 5.34 | 0.030 |
| D2S155 | 4.40 | 3.84 | 3.28 | 2.16 | 1.11 | 4.40 | 0.001 |
| D2S369 | 3.85 | 4.48 | 4.08 | 2.98 | 1.80 | 4.54 | 0.030 |
| qter | | | | | | | |

### Determination of the telomeric limit of the *FALS-AR* locus

Previous analysis [6] identified a crossover that involves marker D2S157 in individual VI:4. With data from additional markers (D2S154 and D2S155), we have now determined that both chromosomes of this individual independently show evidence of crossover events (Fig. 1). Although affected individual VI:4 is homozygous (alleles *3 3*) for marker D2S157 as would be expected for a recessive condition in this inbred family, both of these alleles derive from parental non-disease-associated chromosomes (Fig. 1). Thus, both chromosomes have been involved in crossovers centromeric to marker D2S157. In addition, individual VI:4 is heterozygous (alleles *2 3*) for marker D2S154. Each parent of this individual carries a disease chromosome with a *2* allele at this marker, and a non-disease-associated chromosome with a *3* allele (Fig. 1). Thus, one of the crossover events must have occurred between markers D2S155 and D2S154, although we cannot determine in which parent this crossover occurred. These data establish an explicit telomeric limit for this locus at marker D2S154.

In earlier work, a crossover that involved marker D2S155 could be inferred if the assumption is made that all disease chromosomes in this highly consanguineous family derive from the same ancestral chromosome [6]. When additional markers are examined, two crossovers, both of which involve D2S155, can be inferred. The effect of one of these crossovers is seen in individual V:2 (Fig. 1), but it may have been present in an earlier generation. The second inferred ancestral crossover also involves the more-centromeric marker D2S2237. This haplotype is seen in individuals IV:5 and V:3 (and their descendants, Fig. 1). Unless two independent crossovers that involve marker D2S2237 have occurred, the presence of this haplotype in both individuals implies that the crossover occurred in an earlier generation. Based on the YAC contig we have assembled, the nearest polymorphic markers centromeric to D2S2237 are D2S1384 and D2S2189, which are both homozygous in all affected individuals in this family. Thus, the telomeric end of the *ALS2* region must lie between markers D2S2237 and D2S1384, if the assumption of a single inherited ancestral chromosome is correct. Accordingly, we have designated D2S2237 as the telomeric boundary of this *FALS-AR* locus (Fig. 2).

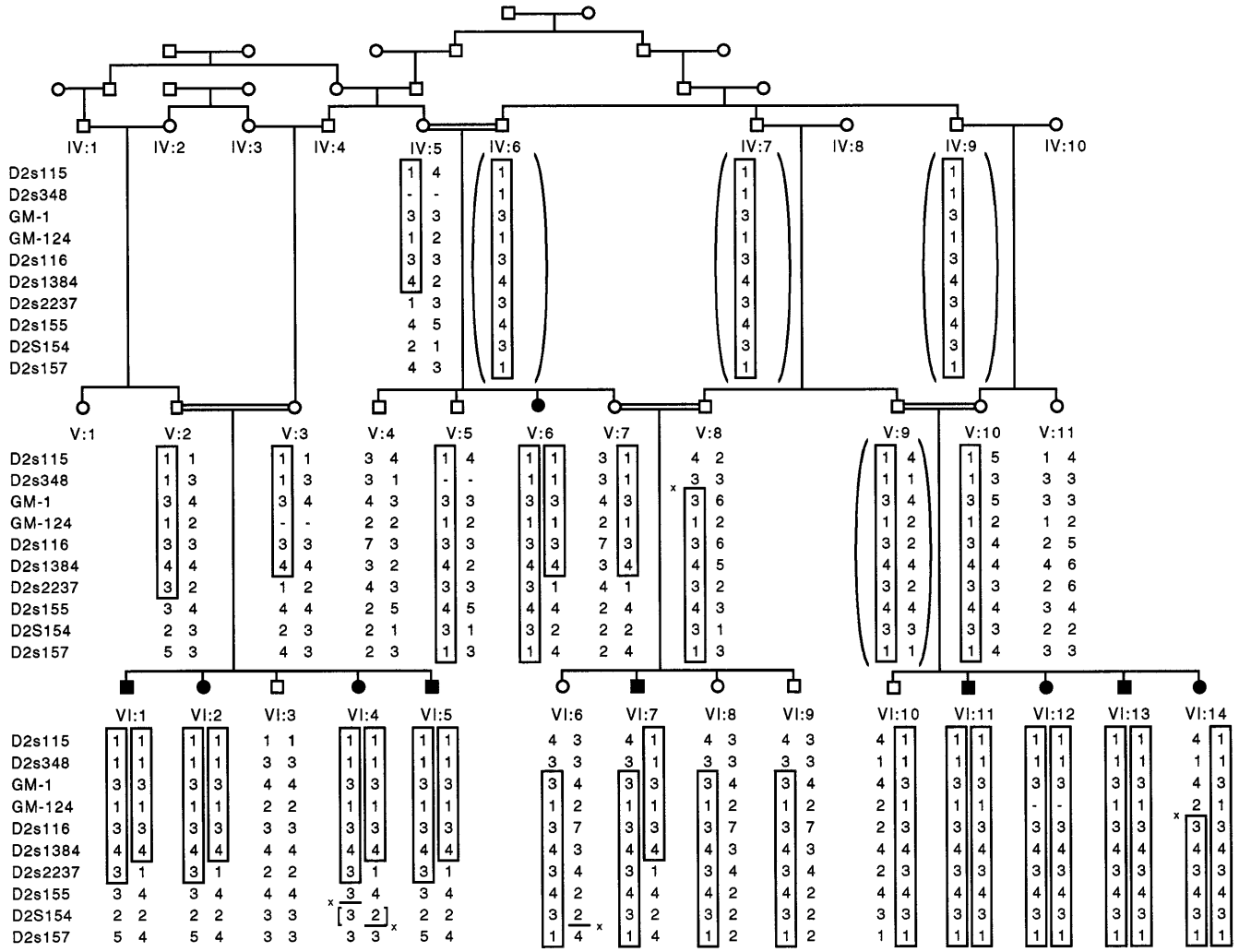


**Fig. 1** Haplotype analysis of the 2q33 region in family F107. Allele data for selected markers in the 2q32-q33 region are shown, with inferred genotypes in *parentheses*. The marker names are listed on the *far left*. *Horizontal lines* within the haplotypes and *x* indicate crossover events. Haplotypes associated with the disease phenotype are *boxed*. For individual VI:4, alleles that cannot be definitively assigned to a specific chromosome are shown in *square brackets*. *Squares* and *circles* represent males and females, respectively. *Filled symbols* represent affected individuals. *Double horizontal lines* in the pedigree indicate marriages between related individuals

## Characterization of the centromeric limit of the *FALS-AR* locus

Previous haplotype analysis [6] of this family indicated in individual VI:14 a crossover event that included marker D2S115, establishing a centromeric limit for this locus. The crossover does not involve D2S116 (LOD score for D2S116 is 5.29 at $\Theta = 0.00$) (Table 2, Fig. 1). The gene that encodes the enzyme AOX [26, 27] maps very close to D2S116, as both are present in the same 280-kb Not I fragment of YAC A188E12 [28]. We initiated a closer study of the genomic region near AOX and D2S116 for two reasons. First, some cases of dominantly inherited ALS are associated with mutations in the gene encoding SOD1, an enzyme involved in reactive oxygen species metabolism [3, 4]. Thus, other enzymes, such as AOX, that potentially generate free radicals are candidates for causing other forms of ALS, including FALS-AR. Also, if the extent of the region involved in the crossover in individual VI:14 could be mapped nearer to D2S116, the size of the candidate region would be substantially reduced and AOX might be excluded as a candidate.

Because of the similarity between the messages for human AOX and mouse xanthine oxidase (XO) [18, 29], the structure of the mouse XO gene [30] was used as a guide for the selection of primers to attempt to amplify introns of the AOX gene. Primers AOX F7 and AOX R6, which span a 270-bp fragment of the AOX cDNA, amplify an approximately 1.5-kb fragment from YAC A188E12, including an intron located after base 1393 of the cDNA sequence for AOX (GenBank entry L11005 [18]). This intron corresponds to intron *13* of the recently reported structure of the AOX gene [31]. Primer AOX F101, positioned near the 3′ end of this intron, and primer AOX R12 amplify an approximately 1.1-kb genomic fragment, including exon *14* and intron

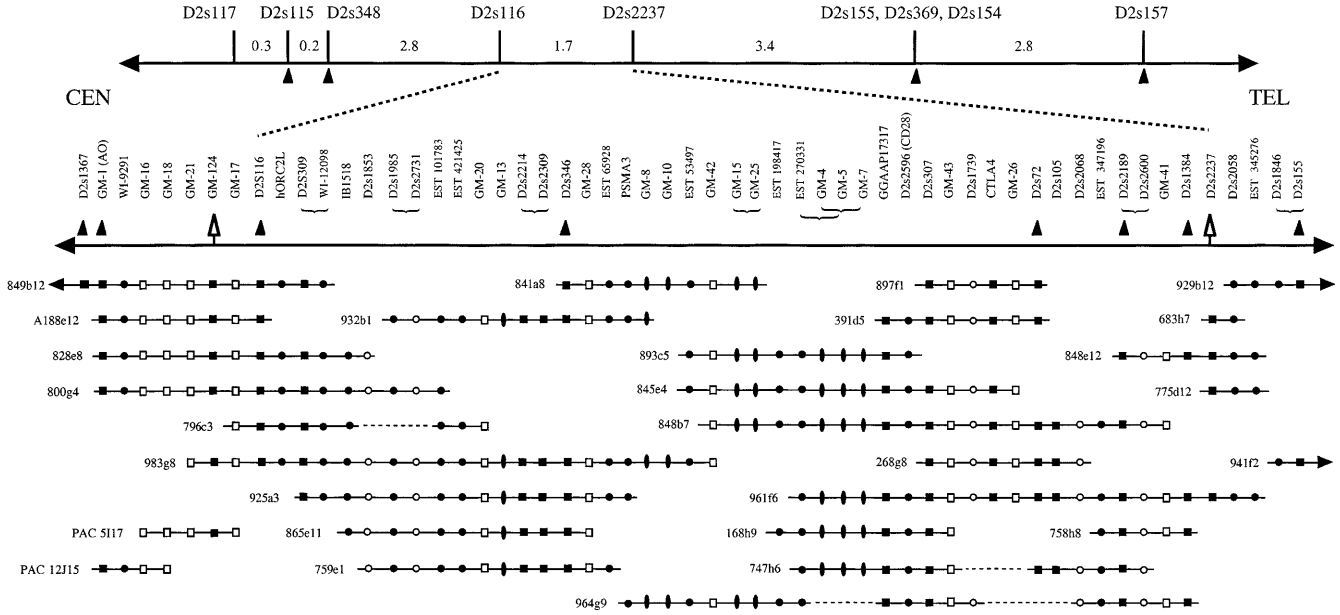


**Fig. 2** Yeast artifical chromosome (YAC) contig across the *ALS2* region. The *line at the top* of the diagram shows the genetic map of the chromosome q233 region, with the centromere (*CEN*) to telomere (*TEL*) orientation indicated. Distances [32, 33] between selected markers are given in cM. In the *lower portion* of the diagram, not drawn to scale, YACs and *PACs* are shown as *horizontal lines* with their names to the left; *dotted lines* indicate potential gaps or deletions. Markers are listed with their presence in individual clones shown by a symbol indicating the type of marker as follows: *solid squares* = polymorphic markers; *open squares* = YAC or PAC end fragment markers; *solid ovals* = Alu polymerase chain reaction fragment markers; *solid circles* = markers representing expressed sequences; *open circles* = other STSs. Single parentheses group markers with positions relative to each other that are undetermined. *Arrows* indicate the limits of the *ALS2* locus. *Arrowheads* indicate other markers shown in Fig. 1 or listed in Table 2 (*PSMA3* proteasome component C3, *hORC2L* human origin recognition complex protein subunit 2)

*14*, which is located after base 1578 of the cDNA. This intron contains a polymorphic $(CA)_N$ repeat (marker GM-1, GenBank Assession number G31889, Table 1). Marker GM-1 is involved in the crossover in individual VI:14 (Fig. 1), thus moving the centromeric boundary of the disease locus from D2S115 to GM-1, within 280 kb of D2S116. This finding essentially limits the ALS2 locus to the 1.7-cM distance between D2S116 and D2S2257 [32].

To refine further the boundary of the centromeric crossover, a chromosome 2-specific PAC library [19] was probed with a 242-bp fragment corresponding to AOX exon *14*. Four PACs were identified and screened for the presence of other markers near D2S116. PACs 10H2 and 14F22 were positive for GM-1. PAC 12J15 contained WI-9291 (derived from the 3′ untranslated region of the AOX cDNA) and GM-1. The hybridization signal for PAC 5I17 was weaker than that for the other PACs, and this PAC does not contain markers WI-9291 or GM-1. However, it is positive for GM-21 (Table 1), an STS marker derived from the end of YAC 983G8 that is near marker D2S116 in our contig that crosses the ALS2 locus. PCR testing of markers derived from the ends of PACs 12J15 and 5I17 confirmed the relative positions of these PACs near the AOX gene and D2S116 (Fig. 2). Thus, while PAC 5I17 does not contain the AOX gene, it does map to this region and may contain sequences homologous to the AOX gene.

PAC 5I17 was digested with *Eco*RI, and the subclones were probed with GM-21. A 5.2-kb fragment that contains GM-21 was identified. Segments from this fragment were amplified from DNA from unaffected individual VI:10 (Fig. 1). This individual is heterozygous for the family disease chromosome and the non-disease chromosome involved in the crossover in individual VI:14. The sequences of the amplified fragments were determined. We identified within one segment (GM-124, located in YAC 983G8 approximately 3.9 kb from the YAC end marked by GM-21), a polymorphism that consisted of two single-base changes separated by 235 bp. As shown in Fig. 1, individual VI:14 was heterozygous for this polymorphism. This result moved the crossover point in individual VI:14 within YAC 983G8, excluding the AOX gene from the *ALS2* locus. Based on minimal size estimates for the AOX gene and PAC 5I17, we estimate that marker GM-124 is within 160 kb of D2S116. Thus, the limit of the centromeric end of the ALS2 region is narrowly defined.

## Assembly of a YAC contig

After analysis of available database information and mapping reported for this region [16, 17, 32–36], 27 YACs were selected for our study based on four criteria: (1) the confirmed presence of markers from the *ALS2* locus, (2) minimal evidence for being chimaeric (indicated by strong hybridization to other chromo-

somes or positive signals for markers from other locations) or containing a deletion (negative for markers within a series), (3) estimated size (YACs ≥ 1,100 kb were selected to ensure a framework, while YACs ≤ 600 kb were expected to provide data for finer resolution), and (4) potential usefulness in closing gaps (probability of overlapping other selected YACs, as indicated by Alu fingerprint data). Total DNA was prepared from the selected YACs and tested by PCR for the presence of markers that map within or near the *ALS2* region. Control human genomic DNA and a YAC (CEPH 747H8) that does not contain DNA from the region of interest were used as positive and negative controls, respectively, for all markers tested.

The initial map assembled using these data contained several ambiguities of order and position and a gap telomeric to marker D2S346. We therefore developed several new STS markers in this locus using inter-Alu PCR and YAC end fragments (Table 1). Some ESTs were also used to refine the map. A marker map consistent with these combined results and that minimized internal YAC deletions was established (Fig. 2). We estimate the total distance between GM-124 and D2S2237 to be 3 million bp.

## Analysis of known genes and ESTs

Many genes have been mapped to the general 2q32-q33 region and could be considered as candidates for the *ALS2* gene. Using the established YAC contig, the presence or absence of these genes within the region defined for *ALS2* can be determined. Some genes were excluded from consideration on the basis of published map data [17, 37, 38]. In addition, we tested markers for ESTs [17, 37–39] and various genes of particular interest to determine whether they could be amplified by PCR from YACs in the FALS-AR contig. The crossover that involves marker GM-124 excludes the AOX gene from consideration. Excluded genes are listed in Table 3. Genes with coding sequences found within the contig include proteasome component C3 (*PSMA3*) [40], the human origin recognition complex protein subunit 2 (*hORC2L*) [41, 42], and Abl interactor 2 (*Abi-2*) [43]. Markers for 34 ESTs are present within the contig and potentially represent candidate genes for *ALS2*.

Two genes, the T-cell membrane glycoprotein CD28 and the cytotoxic T-lymphocyte-associated protein CTLA4, were previously mapped to YACs that are contained in our contig [36]. Two members of the family (affected individual VI:1 and carrier individual V:3) were screened for alterations in the coding regions of these genes. Primers placed in intronic regions were used to amplify each of the four exons encoding CD28 [25] and the three exons encoding CTLA-4 [24], and the sequences of the resulting products were determined. For CD28, no changes were observed in these individuals relative to the published wild-type sequence [25]. Both individuals were homozygous for two changes in the sequence for CTLA-4. In the 5′ untranslated region, we observed two C's (ttggCCatg) at posi-

**Table 3** Exclusion of known genes and expressed sequence tags (ESTs) from the FALS-AR locus

| Gene name | Probe | Method[a] |
|---|---|---|
| *ACADL* | pr: M74096 | B |
| Acetylcholine receptor $\alpha$ | WI-6965 | A, B |
| Acetylcholine receptor $\delta$ | WI-7254 | B |
| Acetylcholine receptor $\gamma$ | pr: X01718 | B |
| cAMP-response element binding protein (CREB) | WI-8988 | A |
| Carbamyl phosphate synthetase | WI-7790 | A |
| Collagen 3a1 | WI-7548 | A |
| Collagen 4a4 | WI-7328 | A |
| Collagen 5a2 | WI-7678 | A |
| Fibronectin 1 | WI-8999 | A |
| Glutamate decarboxylase | UTR-9857 | A |
| Inositol polyphosphate-1 phosphatase | WI-9155 | A, B |
| Insulin-like growth-factor binding protein 5 | WI-7179 | A |
| Integrin a4 subunit | WI-9025 | A |
| Interleukin 8 receptor $\alpha$ | WI-7744 | A |
| Myosin light chain | WI-7607 | A |
| *n*-chimaerin | WI-7611 | A |
| Nucleolin | WI-7767/D2s2594 | A |
| *PAX3A* | WI-7962 | A |
| *PLC-L* | DH82-1,2 | B |
| Titin | WI-9271 | A, B |
| Tyrosine-kinase-type cell surface receptor | pr: L07868 | B |
| Vitronectin receptor $\alpha$ subunit (integrin $\alpha$5) | WI-7853 | A |
| D2S1846 | WI-6330 | B |
| D2S2019 | WI-3694 | B |
| dbEST 02446 | D2s421E | B |
| dbEST 03323 | D2s238E | B |

[a] A = Relative position was determined by inspection of published maps [17, 37, 38]. B = EST or gene marker did not amplify by PCR from YACs in the contig

tion –51 relative to the start of the open reading frame, where the published sequence lists a single C (ttggCatg) [24]. In exon *1*, base 328 of the open reading frame was initially reported as G (GCC→Ala) [24], while we observed an A (ACC→Thr). These two differences were also seen in the sequence of the corresponding region from normal human lymphoblast DNA. We conclude that these two alterations represent either variations in the wild-type sequence or inaccuracies in the originally reported sequence of CTLA-4 and that they are not associated with the disease phenotype caused by the *ALS2* locus.

All ESTs and the 31 map markers within the *ALS2* locus amplify from the genomic DNA of affected individuals. This observation suggests that these individuals do not have extensive DNA deletions across or within this locus.

## DISCUSSION

Our analysis of additional markers in the 2q33 chromosomal region has confirmed the previously reported linkage [6] of this region to type 3 FALS-AR. With these markers the FALS-AR region has been more precisely limited to a genetic distance of about 1.7 cM. Although this localization depends on the assumption of a single ancestral disease chromosome within the family, the overall similarity of the disease haplotype in this region and the high degree of consanguinity in the family support this assumption.

We have established a YAC contig that covers the disease locus with at least threefold depth, allowing us to estimate the physical size as approximately 3 million bp. The relative order of markers in the contig is essentially consistent with previous physical mapping efforts in this region [17, 32, 36]. We have placed marker IB1518 to the centromeric side of adjacent marker D2S1853 (also known as WI-4092), while prior maps [17, 32] show D2S1853 as the more centromeric of these two markers. Similarly, D2S307 maps to the telomeric side of markers GGAA-P17317 and D2S2596 in our contig, while it is placed closer to the centromere by others [17, 32]. Several of the new markers we have identified map between markers D2S346 and D2S307, an area in which few markers had previously been described. Seven of the markers between D2S346 and D2S307 are derived from Alu-PCR fragments and none is a simple sequence repeat, consistent with previous reports that these types of repetative sequences tend to occur in clusters by class [44–46]. This finding confirms the value of using different marker types to close gaps in existing physical maps [44].

By narrowing the defined *ALS2* region and combining information from our YAC contig with other available mapping data, we have been able to eliminate many potential candidate genes from further consideration as a causative gene for this disease in this family. Among the eliminated genes are several that are expressed in skeletal muscle or motor neurons, including titin, nucleolin, myosin light chain, and acetylcholine receptor subunits alpha, delta, and gamma. AOX has also been excluded from the disease region. The mouse genome contains at least two adjacent AOX genes [47, 48], and there may be multiple AOX genes in the human genome as well. The presence of a homologous gene would explain why we detected PAC 5I17 with an AOX exon *14* probe, even though this PAC does not contain the known AOX gene. Any genes that map centromeric to the 2-bp polymorphism (GM-124) within the PAC are also excluded as candidates for the FALS-AR causative gene. No disease-associated sequence alterations were observed in two genes, *CTLA-4* and *CD28*, that are located within the *ALS2* locus.

Traditionally, neurodegenerative diseases have been diagnosed and grouped on the basis of clinical symptoms and criteria, such as patient age at disease onset and the classes and locations of affected nerves and muscles. Knowledge of the genomic alterations causing these diseases is forcing revision of many clinical classifications. There are now examples of nominally distinct diseases being caused by mutations within a single gene, as well as clinically similar conditions associated with changes in multiple genes. The dominant form of familial ALS is among the diseases known to be caused by at least two separate genes [3, 5]. FALS-AR also appears to be heterogeneous. Thus, while ALS in the large family with type 3 FALS-AR that we analyzed here is genetically linked to the chromosome 2q33 region, another subset of FALS-AR families shows genetic linkage to chromosome 15q (M. Ben Hamida and F. Hentati, unpublished data). In the long term, the optimal classification of neuromuscular conditions is likely to be based on the underlying genetic causes.

**Acknowledgements** We are indebted to the patients and their family for their assistance with this project. We thank Généthon for providing the CEPH YAC clones, and Drs. Jeffrey C. Gingrich and Erwin Schurr for providing the chromosome 2-specific PACs. This research has been generously supported by the Muscular Dystrophy Association (R.H.B., D.M.-Y.), the Amyotrophic Lateral Sclerosis Association (D.P., R.H.B.), the Pierre de-Bourgknecht ALS Research Foundation (R.H.B.), the Myrtle May MacLellan ALS Research Foundation (R.H.B.), the CB Day Investment Company, NIH grant 1PO1NS31248-02 (R.H.B., B.A.H.), NRSA Fellowship 1 F32 NS10064-01 (B.A.H.), and the Howard Hughes Medical Institute (H.R.H., P.C.S., W.C.). H.R.H. is an Investigator of the Howard Hughes Medical Institute.

## REFERENCES

1. Mulder DW, Kurland LT, Offord KP, Beard CM (1986) Familial adult motor neuron disease: amyotrophic lateral sclerosis. Neurology 36:511–517
2. Ben Hamida M, Hentati F, Ben Hamida C (1990) Hereditary motor system diseases (chronic juvenile amyotrophic lateral sclerosis). Conditions combining a bilateral pyramidal syndrome with limb and bulbar amyotrophy. Brain 113:347–363

3. Rosen DR, Siddique T, Patterson D, Figlewicz DA, Sapp P, Hentati A, Donaldson D, Goto J, O'Regan JP, Deng H-X, et al (1993) Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. Nature 362:59–62
4. Siddique T, Deng HX (1996) Genetics of amyotrophic lateral sclerosis. Hum Mol Genet 5:1465–1470
5. Figlewicz DA, Krizus A, Martinoli MG, Meininger V, Dib M, Rouleau GA, Julien J-P (1994) Variants of the heavy neurofilament subunit are associated with the development of amyotrophic lateral sclerosis. Hum Mol Genet 3:1757–1761
6. Hentati A, Bejaoui K, Pericak-Vance MA, Hentati F, Speer MC, Hung W-Y, Figlewicz DA, Haines J, Rimmler J, Ben Hamida C, et al (1994) Linkage of recessive familial amyotrophic lateral sclerosis to chromosome 2q33–35. Nat Genet 7:425–428
7. Lathrop GM, Lalouel J-M, Julier C, Ott J (1984) Strategies for multilocus analysis in humans. Proc Natl Acad Sci USA 81:3443–3446
8. Lathrop GM, Lalouel J-M (1984) Easy calculations of LOD scores and genetic risks on small computers. Am J Hum Genet 36:460–465
9. Lathrop GM, Lalouel J-M, White RL (1986) Construction of human genetic linkage maps: likelihood calculations for multilocus analysis. Genet Epidemiol 3:39–52
10. Schaffer AA, Gupta SK, Shriram K, Cottingham RW Jr (1994) Avoiding recomputation in linkage analysis. Hum Hered 44:225–237
11. Schaffer AA (1996) Faster linkage analysis computations for pedigrees with loops or unused alleles. Hum Hered 46:226–235
12. Cottingham RW Jr, Idury RM, Schaffer AA (1993) Faster sequential genetic linkage computations. Am J Hum Genet 53:252–263
13. Burke DT, Carle GF, Olson MV (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science 236:806–812
14. Dausset J, Ougen P, Abderrahim H, Billault A, Sambucy JL, Cohen D, Le Paslier D (1992) The CEPH YAC library. Behring Inst Mitt 91:13–20
15. Bellanné-Chantelot C, Lacroix B, Ougen P, Billault A, Beaufils S, Bertrand S, Georges I, Gilbert F, Gros I, Lucotte G, et al. (1992) Mapping the whole human genome by fingerprinting yeast artificial chromosomes. Cell 70:1059–1068
16. Cohen D, Chumakov I, Weissenbach J (1993) A first-generation physical map of the human genome. Nature 366:698–701
17. Hudson T, Stein L, Gerety S, Ma J, Castle A, Silva J, Slonim D, Baptista R, Kruglyak L, Xu S, et al (1995) An STS-based map of the human genome. Science 270:1945–1954, with supplementary data from the Whitehead Institute/MIT Center for Genome Research, Human Genetic Mapping Project, Data Release 12 (July 1997)
18. Wright RM, Vaitaitis GM, Wilson CM, Repine TB, Terada LS, Repine JE (1993) cDNA cloning, characterization, and tissue-specific expression of human xanthine dehydrogenase/xanthine oxidase. Proc Natl Acad Sci USA 90:10690–10694
19. Gingrich JC, Boehrer DM, Garnes JA, Johnson W, Wong BS, Bergmann A, Eveleth GG, Langlois RG, Carrano AV (1996) Construction and characterization of human chromosome 2-specific cosmid, fosmid, and PAC clone libraries. Genomics 32:65–74
20. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: A laboratory manual (Cold Spring Harbor Press, Cold Spring Harbor, N.Y.
21. Riley J, Butler R, Finniera R, Jenner D, Powell S, Anand R, Smith JC, Markham AF (1990) A novel, rapid method for the isolation of terminal sequences from yeast artificial chromsome (YAC) clones. Nucleic Acids Res 18:2887–2890
22. Munroe DJ, Haas M, Bric E, Whitton T, Aburatani H, Hunter K, Ward D, Housman DE (1994) IRE-bubble PCR: a rapid method for efficient and representative amplification of human genomic DNA sequences from complex sources. Genomics 19:506–514
23. Nelson DL, Ledbetter SA, Corbo L, Victoria MF, Ramírez-Solis R, Webster TD, Ledbetter DH, Caskey CT (1989) *Alu* polymerase chain reaction: a method for rapid isolation of human-specific sequences from complex DNA sources. Proc Natl Acad Sci USA 86:6686–6690
24. Dariavach P, Mattei M, Golstein P, Lefranc M (1988) Human Ig superfamily *CTLA-4* gene: chromosomal localization and identity of protein sequence between murine and human CTLA-4 cytoplasmic domains. Eur J Immunol 18:1901–1905
25. Lee K, Taylor C, Petryniak B, Turka L, June C, Thompson C (1990) The genomic organization of the CD28 gene. Implications for the regulation of CD28 mRNA expression and heterogeneity. J Immunol 145:344–352
26. Badwey JA, Robinson JM, Karnovsky MJ, Karnovsky ML (1981) Superoxide production by an unusual aldehyde oxidase in guinea pig granulocytes. Characterization and cytochemical localization. J Biol Chem 256:3479–3486
27. Shaw S, Jayatilleke E (1990) Ethanol-induced iron mobilization: role of acetaldehyde-aldehyde oxidase generated superoxide. Free Radic Biol Med 9:11–17
28. Berger R, Mezey E, Clancy KP, Harta G, Wright RM, Repine JE, Brown RH, Brownstein M, Patterson D (1995) Analysis of aldehyde oxidase and xanthine dehydrogenase/oxidase as possible candidate genes for autosomal recessive familial amyotrophic lateral sclerosis. Somat Cell Mol Genet 21:121–131
29. Terao M, Cazzaniga G, Ghezzi P, Bianchi M, Falciani F, Perani P, Garattini E (1992) Molecular cloning of a cDNA coding for mouse liver xanthine dehydrogenase. Regulation of its transcript by interferons in vivo. Biochem J 283:863–870
30. Cazzaniga G, Terao M, Lo Schiavo P, Galbiati F, Segalla F, Seldin MF, Garattini E (1994) Chromosomal mapping, isolation, and characterization of the mouse xanthine dehydrogenase gene. Genomics 23:390–402
31. Wright RM, Weigel LK, Varella-Garcia M, Vaitaitis G, Repine JE (1997) Molecular cloning, refined chromosomal, and structural analysis of the human gene encoding aldehyde oxidase (AOX1), a candidate gene for the *ALS2* gene. Redox Rep 3:135–144
32. Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, et al. (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature 380:152–154
33. Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, Marc S, Bernardi G, Lathrop M, Weissenbach J (1994) The 1993–94 Généthon human genetic linkage map. Nat Genet 7:246–339
34. Chumakov IM, Rigault P, Le Gall I, Bellanné-Chantelot C, Billault A, Guillou S, Soularue P, Guasconi G, Poullier E, Gros I, et al (1995) A YAC contig map of the human genome. Nature 377 [Suppl]:175–297
35. Gyapay G, Schmitt K, Fizames C, Jones H, Vega-Czarny N, Spillett D, Muselet D, Prud'Homme JF, Dib C, Auffray C, et al (1996) A radiation hybrid map of the human genome. Hum Mol Genet 5:339–346
36. Buonavista N, Balzano C, Pontarotti P, Le Paslier D, Golstein P (1992) Molecular linkage of the human CTLA4 and CD28 Ig-superfamily genes in yeast artificial chromosomes. Genomics 13:856–861
37. Boguski MS, Schuler GD (1995) ESTablishing a human transcript map. Nat Genet 10:369–371
38. Schuler GD, Boguski MS, Stewart EA, Stein LD, Gyapay G, Rice K, White RE, Rodriguez-Tomé P, Aggarwal A, Bajorek E, et al (1996) A gene map of the human genome. Science 274:540–546

39. Stewart EA, McKusick KB, Aggarwal A, Bajorek E, Brady S, Chu A, Fang N, Hadley D, Harris M, Hussain S, et al (1997) An STS-based radiation hybrid map of the human genome. Genome Res 7:422–433
40. Tamura T, Lee DH, Osaka F, Fujiwara T, Shin S, Chung CH, Tanaka K, Ichihara A (1991) Molecular cloning and sequence analysis of cDNAs for five major subunits of human proteasomes (multi-catalytic proteinase complexes). Biochim Biophys Acta 1089:95–102
41. Gavin KA, Hidaka M, Stillman B (1995) Conserved initiator proteins in eukaryotes. Science 270:1667–1671
42. Takahara K, Bong M, Brevard R, Eddy RL, Haley LL, Sait SJ, Shows TB, Hoffman GG, Greenspan DS (1996) Mouse and human homologues of the yeast origin of replication recognition complex subunit ORC2 and chromosomal localization of the cognate human gene *ORC2L*. Genomics 31:119–122
43. Dai Z, Pendergast AM (1995) Abi-1, a novel SH3-containing protein interacts with the c-Abl tyrosine kinase and modulates c-Abl transforming activity. Genes Dev 9:2569–2582
44. Aburatani H, Stanton VP Jr, Housman DE (1996) High-resolution physical mapping by combined *Alu*-hybridization / PCR screening: construction of a yeast artificial chromosome map covering 31 centimorgans in 3p21-p14. Proc Natl Acad Sci USA 93:4474–4479
45. Weissenbach J, Gyapay G, Dib C, Vignal A, Morissett J, Millasseua P, Vaysseix G, Lathrop M (1992) A second generation linkage map of the human genome. Nature 359:794–801
46. Korenberg JR, Rykowsky MC (1988) Human genome organization: Alu, LINEs, and the molecular structure of metaphase chromosome bands. Cell 53:391–400
47. Holmes RS (1979) Genetics, ontogeny, and testosterone inducibility of aldehyde oxidase isozymes in the mouse: evidence for two genetic loci (*Aox-1* and *Aox-2*) closely linked on chromosome 1. Biochem Genet 17:517–527
48. Holmes RS, Leijten LR, Duley JA (1981) Liver aldehyde oxidase and xanthine oxidase genetics in the mouse. Anim Blood Groups Biochem Genet 12:193–199