Volume 15 Number 22 1987 Nucleic Acids Research

A 230kb cosmid walk in the Duchenne muscular dystrophy gene: detection of a conserved sequence and of a possible deletion prone region

R.Heilig, C.Lemaire and J.-L.Mandel*

Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Faculté de Médecine, 11 rue Humann, 67085 Strasbourg Cedex, France

Received October 1, 1987; Accepted October 27, 1987 Accession no. Y00494

SUMMARY

A 230 kb genomic region from the Duchenne muscular dystrophy gene has been cloned in a cosmid walk, using an improved vector and by screening the same unamplified library for all steps. The region cloned surrounds the translocation breakpoint characterized by Worton et al (1) and Ray et al (2), and overlaps by 70 kb the Pert region cloned by Monaco et al (3). We have identified a region of strong sequence conservation in mammals and chicken, and comparison of the homologous sequences in chicken and man has indicated the presence of two putative protein coding exons. Comparison with the sequence recently published by Koenig et al (4) shows that only one is present in the Duchenne cDNA, and this raises the question of the functional significance of the other conserved sequence. Single copy probes and whole cosmids generated during this work have been used to analyse the corresponding region in Duchenne patients. Of five independant patients shown to be deleted for a probe 30 kb in 3' of the translocation breakpoint, three have the 5' endpoint of the deletion within a region of less than 20 kb, 100 kb away from the probe used to ascertain the deletion. This might suggest the presence of a region where deletions occur preferentially.

INTRODUCTION

Duchenne muscular dystrophy (and the milder Becker form) affect about one male in 4000. Despite intensive research the biochemical function of the gene product has remained completely elusive. The elegant work of the groups of Kunkel (3, 4, 5, 6) and Worton (1, 2, 7) has recently led to the isolation of genomic and cDNA sequences corresponding to part of the DMD gene. The mapping of incomplete cDNA clones (3) and the results of restriction mapping experiments using pulse field gel electrophoresis (8, 9) suggested that the DMD gene in truly enormous and might cover as much as 1000 to 2000 kb, and codes for a 16 kb mRNA. This has been very recently confirmed by Koenig et al (4) who cloned the entire cDNA. This complexity will render very difficult the study of its genomic organisation and of the mutations that cause the disease, especially since it can be expected that each Duchenne family will

carry an independant mutation. In an effort to characterize a region of the DMD gene different from that analysed by Monaco et al. (3, 5) we have engaged in a cosmid walk around the translocation breakpoint caracterized by Ray et al (2). We have cloned 230 kb of contiguous genomic DNA, partially characterized a sequence conserved in man and chicken, and defined end points of three independant Duchenne deletions, located within a 20 kb region of the gene.

MATERIALS AND METHODS

Cloning of junction fragment from the X-21 translocation.

A fragment containing both a rDNA sequence from chromosome 21 and an X linked sequence was cloned from the A 2-4 somatic cell hybrid line that contains the translocated chromosome 21qter-p11 : Xp212-qter as sole human contribution on a mouse background (1). This cell line was provided to us by Dr. R. Worton. The A2-4 DNA was first digested with BglII and fragments in the 3,5 kb range were isolated after electrophoresis in low melting point agarose. The resulting DNA was further digested with BamHI and HindIII, and the 2 kb region was similarly isolated. This allowed us to separate the 2 kb BamHI - HindIII junction fragment from human and mouse rDNA fragments ($\sim$ 7,5 kb). The enriched fraction was cloned in a λ 1149 insertion vector (10) modified to clone HindIII - BamHI fragments. 200 000 recombinants were screened with a 1.6kb BamHI - EcoRI probe corresponding to part of the Human 28 S rDNA (11). Two identical clones were obtained which fit with the map established by Ray et al.(2). The low yield of positive clones confirm the observation by Ray et al. (2) that the junction fragment is extremly underepresented in genomic libraries.

Cosmid cloning.

DNA from a 49 XXXXY human lymphoblastoid cell line was prepared by standard methods. Mecanical shearing was minimised and ethanol precipitation avoided in order to obtain DNA of more than 200 kb in length. This DNA was partially digested with MboI for various length of time and was fractionated by sedimentation on a 5-30 % NaCl gradient (in Tris-HCl 20 mM, pH 7.8, EDTA 5 mM) for 16 h at 36 000 rpm, 20°C in a SW41 Beckmann rotor. Fractions containing fragments larger than 40 kb were used for cosmid cloning. The Sp cos cosmid vector was first cleaved at the PvuII site located between the cos sites and dephosphorylated with calf intestine alkaline phosphatase. After phenol-chloroform extraction and ethanol precipitation, the vector was further cut with BamHI. A 10 fold molar excess of vector was ligated to

fractionated human DNA with T4 DNA ligase and packaging was performed as described in Maniatis et al (12). Packaging extracts were made as described by Scalenghe et al (13).

The library was titrated on the DH1 strain and 1.10^6 cosmid clones were plated directly on Pall Biodyne membrane placed over agar plates (with 100 µg/ml Ampicillin). Replicates were made on the same membrane, covered by an additional nylon membrane and stored at -80°C between two sheets of 3MM Whatman paper impregnated with freezing medium (freezing medium is L broth containing 40mM potassium phosphate (pH 7.2) 7mM $(NH4)_2SO_4$, 1.5mM Na Citrate, 0.4mM Mg SO_4 5 % glycerol and 100 mg/l ampicillin.

The master filter was used for hybridization screening (ten to twelve successive hybridizations were performed). Some replicas were used up to 5 times for clone isolation without detectable loss in efficiency (see text).

Other methods

Genomics blots were performed on Diazobenzyloxymethyl (DBM) paper as described by Oberlé et al (14). Hybridization with whole cosmids was as described by Litt and White (15) with slight modifications. Subcloning of fragments was in general performed using a 2.9 kb vector (Spl) derived from the Sp cos vector (R. Heilig unpublished). For mapping of cosmid clones, individual cosmids were digested with SalI or with SmaI (enzymes that cut rarely human DNA) and were then partially digested with EcoRI, BamHI or BglII. The digests were analysed after electrophoresis on 0.5 % agarose gels by Southern blotting and hybridization to the cosmid vector. Complete digests were analysed after end labelling of fragments using the filling in reaction of DNA polymerase (Klenow fragment). The Maxam Gilbert method (16) was used for DNA sequencing.

RESULTS

The Cosmid Vector

A 6.2 kb cosmid vector, Sp cos, (Fig. 1A) was constructed by modification of the pcos4 vector (Lehrach et al, unpublished). It contains the pBR322 origin of replication (ORI), the β lactamase gene (Amp^r), a 115bp polylinker region containing 15 unique restriction sites, flanked on one side by the Sp6 specific RNA polymerase promoter, and three "cos" sites. On each side of the cloning site (BamHI in the present case) are unique sites for enzymes that cut relatively frequently vertebrate genomic DNA and that can be used for rapid subcloning of end fragments of the genomic insert (Fig. 1B). The DNA from a recombinant cosmid clone is digested by an enzyme corresponding to one

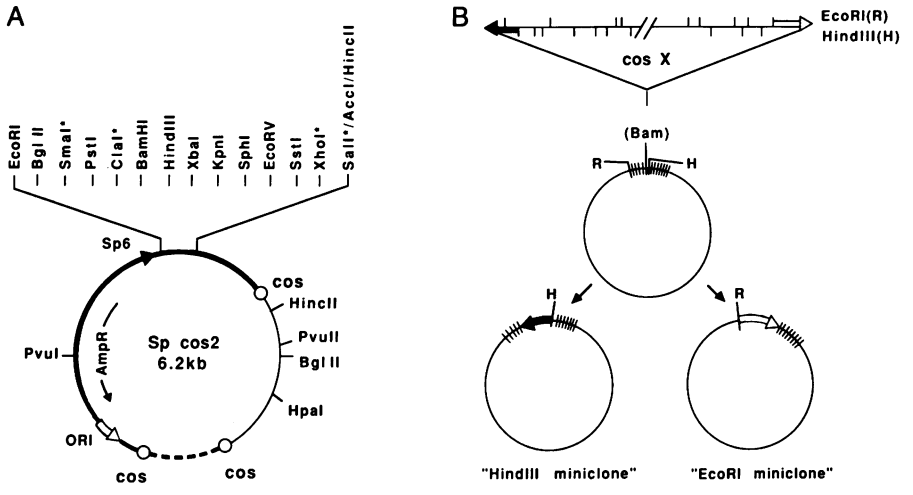

Fig. 1 - Map of the Sp cos 2 vector. A) The heavy line corresponds to the part of the vector (3.8kb) that is conserved in all recombinant cosmids. The region indicated by the thin line is always deleted while that defined by the dashed line can be conserved in some recombinants. Sites which cut rarely in the human genome are marked with an asterisk. All the sites in the polylinker are unique after packaging. Ori is the origin of replication derived from PBR322. Sp6 is the promoter for Sp6 RNA polymerase. B) Scheme for generating "miniclones" corresponding to end fragments of the genomic insert. In the example shown HindIII and EcoRI are chosen to generate clones corresponding to the left and right end, by digestion of the recombinant cosmid and religation at low DNA concentration (≤ 0.4 µg/ml).

of the linker sites, which results in one end fragment linked to the cosmid vector. Ligation of the digest at low DNA concentration favors intramolecular ligation and the selective cloning of the relevant end fragment. By choosing sites on either sides of the BamHI site it is possible to rapidly derive a series of subclones ("miniclones") for both ends of the genomic insert. This facilitates the finding of segments devoid of repetitive sequences to be used as probes for subsequent walking steps. The subcloned fragment can easily be separated from the vector sequences by double digestion of the "miniclone" with the enzyme used to generate it and with an enzyme on the opposite side of the linker. Sites for enzymes that cut only infrequently in vertebrate genomes ("CpG" sites marked by an asterisk in Fig 1A) are useful for restriction mapping of the insert (see Material and Methods).

Strategy for the cosmid walk.

A cosmid library was prepared using DNA from a 49 XXXXY lymphoblastoid cell line partially digested with MboI and size selected by sedimentation on

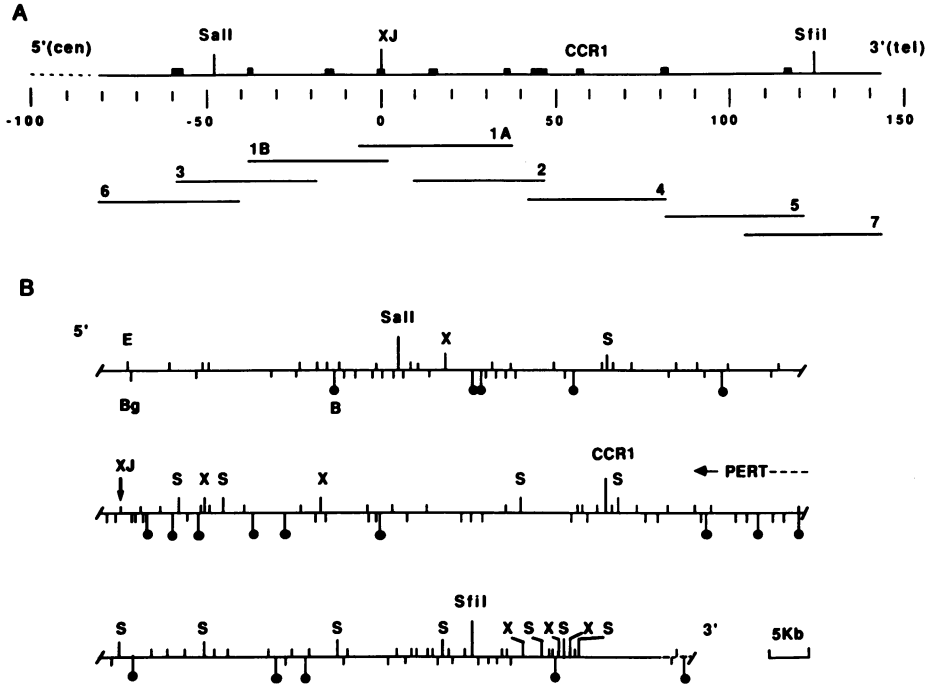

Fig. 2 - Map of the cloned region. A) General map showing the initial start point XJ and the region covered by the best cosmid from each step (labeled 1 to 7). Location of unique sequence probes are indicated by the heavy line, most of them correspond to the extremities of cosmids and are named in the text as C (centromeric) or T (telomeric) followed by the cosmid designation. The position of SalI and SfiI sites is indicated. B) Detailed restriction map of the 230 kb region with EcoRI (above line) BglII (under line) and BamHI (under line, marked with a dot). Additional rare CpG sites are indicated (S = SmaI, X = XhoI, as well as the 5' end of the Pert region).

a 5-30 % NaCl gradient. After in vitro packaging, the unamplified library (1.10^6 clones) was plated at high density on nylon membranes (25,000 clones per 150 cm^2 plate). One replica was made (for each plate) using the same membrane, which was subsequently stored at -80°C using a sandwich method derived from Hanahan and Meselson (17) (see Material and Methods). The original filters were used for all the successive screening tests. For isolation of positive clones, we tried to avoid total thawing of the membranes on which the library was stored. These were placed on dry ice and rapidly punched. Pure clones were in general obtained after two additional screening steps. When several positive clones were obtained for a given step, their pattern of digestion with BglII and with EcoRI were compared between

themselves and to the preceding cosmid clone. This allowed in general to determine the clone that showed the least overlap with the sequences isolated previously. In two cases, (step 2 to 4, and 4 to 5, see fig. 2) the region of overlap was so small that it could only be analysed by digestion with more frequently cutting enzymes (DdeI, HinfI, AluI) followed by hybridization with the screening probe. The cosmid clone that appeared the most interesting at a given step was subjected to "end fragment subcloning". Several restriction sites were chosen on each side of the polylinker and this generated several "miniclones" which were analysed for the presence of repetitive sequences and for their position with respect to the direction of walk. If a unique sequence fragment could not be directly isolated in this way further analysis of end fragment clones was necessary. The "end probes" generated were tested for X linkage and localisation with respect to the X-21 translocation breakpoint, by Southern blot hybridization to appropriate genomic DNAs. In two cases the end probes were found not to be X-linked and alternate cosmid clones had to be analysed. Unique X-linked "end probes" were used to screen the original unamplified library, and it was possible to avoid backward step since clones already positive at the preceding steps were easily identified.

The initial probe for the cosmid walk was a 0.8kb NsiI-HindIII fragment located on the centromeric side from the X-21 translocation. It was cloned, using an enrichment procedure, from the A2-4 mouse-human hybrid line constructed by Worton et al. (1) (see Material and Methods). Seven successive walking steps allowed us to clone a 230 kb region (Fig.2). The 8 corresponding cosmids contain inserts ranging from 37 to 44 kb. A restriction enzyme map was established for EcoRI, BamHI and BglII, and for the "rare cutting" enzyme ("CpG" enzymes") SmaI and XhoI. Single SalI and SfiI sites were found in the centromeric (5') and telomeric (3') regions respectively, separated by 170 kb (Fig.2). This fits approximately with the size (∿ 200kb) of the SalI-SfiI fragment detected by pulse field gel electrophoresis, using the XJ.1.1 probe (18). Since the SfiI site is present within the Pert region (8), this establishes that the genomic region that we have cloned overlaps at its telomeric end with the centromeric end of the Pert region. In fact comparison of the EcoRI restriction map established by Monaco et al (3) to the one presented in figure 2 shows that the two cloned regions overlap by about 70kb.

Search for conserved sequences.

In order to identify potential protein coding sequences, we searched for sequences conserved during vertebrate evolution. A region of about 120 kb

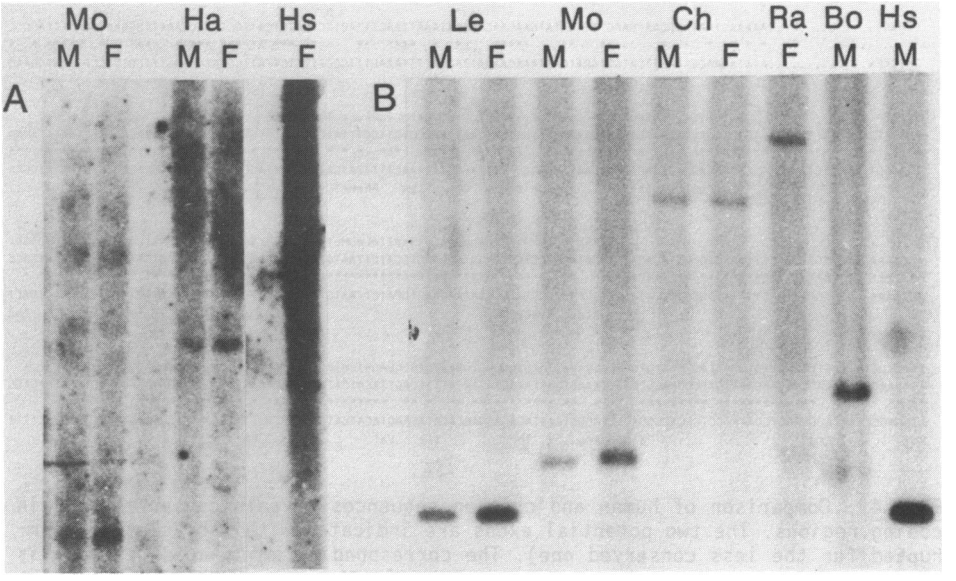

Fig. 3 - Detection of sequence conservation. A) Hybridization of a "zoo" blot to a subcloned 4.9 kb PstI fragment derived from step 4. Mo: Mouse DNA (12 µg) digested with TaqI. Ha: Hamster DNA (12 µg) digested with EcoRI. M: male, F: female. HS: human DNA (3 µg) digested with EcoRI. B) Hybridization to the 1.0 kb HpaI-BstNI fragment (termed CCR1) Le: lemur. Ch: chicken. Ra: Rat. Bo: Bovine. All digests are with TaqI. Hybridizations were performed in 40 % formamide at 37°C as previously described (Oberlé et al 1986).

around the X-21 translocation breakpoint was systematically subcloned in a plasmid vector. We chose, for the subcloning step, restriction enzymes that cut the various cosmids in a limited number (≤ 8) of fragments. The subclones were directly used as probe without eliminating those that contain human repetitive sequences (ie the vast majority of the subclones). We reasoned that repetitive sequences are in general species specific (under appropriate hybridization conditions) and that by testing a sufficient number of mammalian species, it would be possible to increase the probability of detecting cross hybridization with a unique human sequence without interference from repetitive sequences. Since X linkage of protein coding genes is thought to be conserved throughout mammalian evolution (19) we tested male and female DNAs from lemurs, mouse and hamster, so that gene dosage could indicate whether a hybridizing fragment is X linked or not. Chicken DNA was also included in the screening blots. Subclones were tested and a single one corresponding to a 4.9 kb PstI fragment located at 56 kb

Fig. 4 - Comparison of human and chicken sequences reveals conserved protein coding regions. The two potential exons are indicated within brackets (interrupted for the less conserved one). The corresponding amino acid sequence is given in full for the human sequence, and only the amino acids that are not conserved are indicated for the chicken sequence.

from the X-21 translocation breakpoint, on the telomeric side, was found to hybridize to unique X linked fragments in mouse and hamster. This probe detected a smear in human DNA due to presence of repetitive sequence (Fig. 3A). Subcloning allowed to restrict the cross hybridizing sequence to a 1kb HpaI-BstNI subclone that was devoid of human repetitive sequences (laboratory acronym CCR1). Sequence conservation was also detected with this probe in lemurs, rat, bovine and chicken DNA (Fig. 3B). The latter observation is especially significant in view of the large evolutionary distance between human and chicken. Differential dosage in males and females indicated X linkage in mouse, hamster and lemurs. It should be noted that two other clones derived from cosmid 1 A detected a set of conserved fragments in rodent DNA. However these were not X linked and the multiplicity of fragments suggests that they correspond to a pseudogene family.

Comparison to the homologous region in chicken.

We used the 1kb HpaI-BstNI (CCR1) fragment to screen a chicken genomic library (20). Seven bacteriophage clones were obtained that showed overlapping restriction maps. A 1.1kb PstI fragment that contained the cross hybridizing region was subcloned and sequenced. Comparison to the sequence of the corresponding human region showed better than 70 % homology over 385bp (from Nt 63 to 447, in Fig.4) with very few gaps. In particular a long

conserved open reading frame of 60 codons was found between conserved putative acceptor and donor sequences (positions 257 and 438). Amino acid sequence homology was 73 % corresponding to a nucleotide sequence homology of 75 %. Termination codons were found in chicken and man in all the other possible frames. We conclude that this region corresponds to a conserved protein coding exon. In the upstream region the sequence conservation is however much greater than would be expected in the absence of selective pressure. A second putative exon sequence is found in man from Nt122 to 187. This 66bp region is 65 % conserved between man and chicken and a putative donor signal is found in a stretch of 16 completely conserved nucleotides. A single open reading frame is found in man and chicken up to a potential donor site in man. The reading frame would be in phase with that of the following exon, described above. However in the chicken we do not find at the homologous position a characteristic donor sequence, since it does not contain the characteristic GT dinucleotide, which is replaced by GC. It should be noted that GC has been found previously in functional donor splice sites (21), in particular in the chicken and duck α D globin gene (22, 23). Since the protein coding sequence is also less conserved (48 % homology) it is not certain that the sequence between Nt122 and 187 corresponds to a functional exon, especially in the chicken. The very recent work of Koenig et al (4) shows that the region from NT257 to 438 indeed corresponds to exon sequences present in Duchenne cDNA (derived from skeletal muscle), but the conserved upstream sequences are not found in this cDNA (see Discussion).

Analysis of deletion breakpoints in DMD patients.

DNAS from 100 unrelated DMD patients were screened for deletion in collaboration with Dr. C. Junien (Paris) using a probe located 37kb 3' from the X-21 translocation, and derived from the telomeric end of cosmid 1 A (probe T 1A). Five patients were found to be deleted for this probe, two of them were also deleted for probe Pert 87-15, and four for probe Pert 87-1. The DNAs from these patients were further tested with the most 5' cosmid from the present genomic walk (cos 6). The whole cosmid was used as hybridization probe, in the presence of a vast excess of sonicated human DNA, to compete for repetitive sequences present in the cosmid (15). All the genomic fragments detected in this way (a to e in Fig.5B) are from the Duchenne region since they are all absent from two of the patients, and are all present in the somatic hybrid line A2-4 (lane 11) that contains the X p21-qter region as sole human contribution. Three unrelated patients (lanes 2, 5 and 10) were deleted for some but not all of the bands detected by the cosmid probe and a fragment of abnormal size was present. The abnormal

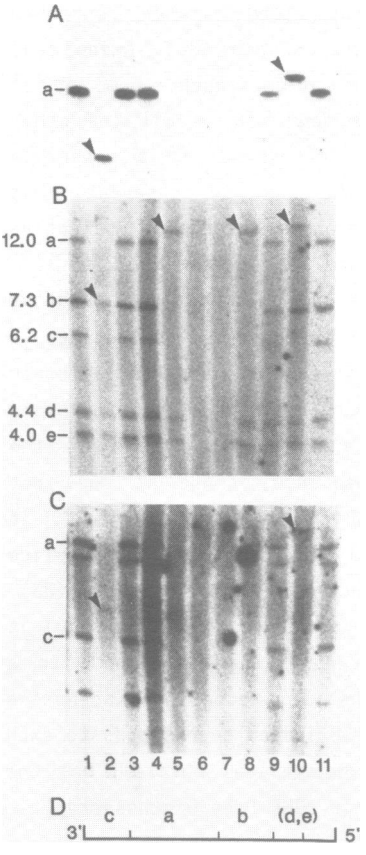

Fig. 5 - Analysis of deletions in patients. DNA (10 µg) from DMD patients (lanes 1 to 10) or from the A2-4 hybrid cell line (lane 11) was digested with BclI, electrophoresed in a 0.8 % agarose gel and transfered to DBM paper. Patients 5 and 8 are first cousins. The blot was analysed successively with various probes from the Pert region and from the region described in Fig 2. A) Probe C cos 3, a 2 kb EcoRI fragment devoid of repetitive sequences derived from the 5' extremity of the step 3 cosmid. B) Step 6 cosmid was used as probe after competition with an excess of human DNA (see Material and Methods). C) Step 3 cosmid was used as in panel B. Normal fragments detected by step 6 cosmid are labelled a to e and their size is indicated in kb. Junction fragments are indicated by an arrow. A map of BclI fragments derived from this analysis is presented in D (BclI does not cut the cosmids because of DNA methylation in bacteria, and thus the map cannot be obtained directly).

fragment was different in each urelated patient, while two affected cousins (lanes 5 and 8) shared the same pattern. Hybridization of the same blot with the cosmid 3 (Fig.5C), which overlaps on the 3' side cosmid 6 (see Fig.2, Fig.5) and with the unique sequence C cos3 (Fig. 5A) located in the region of overlap, allowed a more precise mapping of the deletion end points.

In two patients (lanes 2 and 10) the deletion breakpoint occurs within the 12kb fragment a detected by probe C cos3 (panel A). In the two cousins (lanes 5 and 8), the deletion breakpoint is more centromeric than in the 2 other cases since no signal is found with cosmid 3 (panel C) and it can be concluded that the breakpoint occurs in the 7.3kb fragment b, which is replaced in these patients by a larger fragment (panel B). The order of fragments a to e can be derived from these results (Fig.5D) and this shows that fragments a and b, where the three independant deletion end points are found, are contiguous.

DISCUSSION

We have isolated a 230 kb region around the X-21 translocation breakpoint originally described by Worton et al (1). The cosmid walking method that we have used was rendered efficient by the constant rescreening of the original master filters, thus avoiding backward steps, and by the use of a vector that allows easy isolation of end fragments for further screening. The cloned region must be completely contained within the huge Duchenne gene as indicated by the pulse field gel electrophoresis studies of Van Ommen et al (8) and Kenwrick et al (9). Our placing of SalI and SfiI sites 170 kb apart agrees quite well with the size predicted by Burmeister et al (18). Comparison of restriction maps shows that the region we have isolated overlaps on its 3' side, by about 70 kb, the Pert region. In order to detect potential coding sequences we have searched for sequences conserved in evolution (24), an approach that has proven powerful for detecting exons in the Duchenne gene (3, 25) and for the identification of gene sequences corresponding to the retinoblastoma gene (26) and to a candidate for the cystic fibrosis gene (27). To avoid the tedious and inefficient cloning of single copy sequences, we have used, to search for cross hybridizing sequences, all fragments from a 120 kb region, whether they contained or not human repetitive sequences. By screening for homology in various species, it is in general possible to find at least one where a particular human probe is non repetitive. This proved successful for detecting a region that contains one and possibly two exons as shown by comparison with the homologous chicken sequence. One problem in using cross hybridization as a detection method for expressed genes is the frequent presence of well conserved dispersed families, that in most analysed cases correspond to pseudogene sequences (of the retrotranscript type). We encountered two such sequences in step 1A. However at least for X linked genes, the conservation of X linkage in mammalian species is a efficient way to discriminate between genes and pseudogenes (24). The conserved sequence present in the CCR1 probe (as well as that corresponding to the Pert 87-25 fragment) has recently allowed us to map the Duchenne homolog in mouse, close to the G6PD and coagulation factor VIII loci, and far from the ornithine transcarbamylase locus, thus showing a very different arrangement than in man (28).

While this manuscript was completed, Koenig et al (4) reported the nucleotide sequence of the 5' third of Duchenne cDNA. Their sequence includes exactly the proposed exon in the most conserved region (257-438) but does not include the other conserved sequence and putative exon. Because of the large

evolutionary distance between man and chichen, the homology found upstream from this exon must correspond to a functionnally significant region. Another conserved sequence, (between mouse an man) with putative splice signals was similarly found by Monaco et al (3) near Pert 87-4, but was not present in the muscle cDNA. It would be very interesting to determine whether these two regions correspond to exons subjet to alternative splicing, to exons from an independant gene (or genes) or to conserved regulatory regions.

Finally, although we have analysed only a very limited set of patients with deletions in the region cloned (5 independant patients, from 100 initially screened), the results might suggest that deletion end points occur preferentially within a 20kb region included in cosmid 6 (sum of the size of fragments a and b). The 5 patients were ascertained for deletion with a probe located about 100 kb 3' from this region, where 3 independants end points are located. On a much larger sample (33 end points from 29 independant patients) Monaco et al (29) did not observe significant clustering of end points, with the possible exception of the region around Pert 87-22 (6 in a 25kb region). The finding of 8 endpoints in the 20kb region between Pert 87-1 and 87.8 is probably biased since these probes were the first widely available for such analysis and thus 2000 patients were initially screened for deletion in this small region. It is interesting to note that of the patients analysed by Monaco et al (29) and that were ascertained as being deleted in the Pert region, junction fragments were detected by the J-MD probe for 3 patients. The J-MD breakpoint itself appears located in the region of overlap of cos 3 and cos 6, as shown by comparison of the map in fig. 2 and the maps of this region presented by Monaco et al. (29) and Worton et al. (30). This strengthens our suggestion of a significant clustering of deletion breakpoints in the region defined by cosmid 6 (and probe J-MD). The cloning of a substantial portion of genomic DNA corresponding to the Duchenne gene should be useful to analyse the normal structure of this part of the gene, and its rearrangements in patients.

ACKNOWLEDGEMENTS

We wish to thank Ms. C. Kretz for efficient technical help, Dr R. Worton for providing the A2-4 line, Dr Bachellerie (Toulouse) for rDNA probes, Dr H. Lehrach for the p cos 4 vector and Dr M. Haenlin for help in constructing the Sp cos vector. We gratefully acknowledge the collaboration of Dr C. Junien (Paris) in the screening for deletions in patients. We thank J.P. Moisan for his participation in the initial stages of this work. This work was supported by grants from Caisse Nationale d'Assurance Maladie, Mutuelle Générale de l'Education Nationale, and the Association des Myopathes de France to Dr J.L. Mandel.

*To whom correspondence should be addressed

REFERENCES

1. Worton, R.G., Duff, C., Sylvester, J.E., Schmickel, R.D. and Willard, H.F. (1984) Science, 224, 1447-1449.
2. Ray, P.N., Belfall, B., Duff, C., Logan, C., Kean, V., Thompson, M.W., Sylvester, J.E., Gorski, J.L., Schmickel, R.D. and Worton, R.G. (1985) Nature, 318, 672-675.
3. Monaco, A.P., Neve, R., Colletti-Feener, C., Bertelson, C.J., Kurnit, D.M. and Kunkel, L.M. (1986) Nature, 323, 646-650.
4. Koenig, M., Hoffman, E.P., Bertelson, C.J., Monaco, A.P., Feener, C. and Kunkel, L.M. (1987) Cell 50, 509-517.
5. Monaco, A.P., Bertelson, C.J., Middlesworth, W., Colletti, C.A., Aldridge, J., Fischbeck, K.H., Bartlett, R., Pericak-Vance, M.A., Roses, A.D. and Kunkel, L.M. (1985) Nature, 316, 842-845.
6. Kunkel, L.M. et al (1986) Nature, 322, 73-77.
7. Burghes, A.H.M., Logan, C., Hu, X., Belfall, B., Worton, R.G. and Ray, P.N. (1987) Nature, 328, 434-437.
8. Van Ommen, G.J.B., Verkerk, J.M.H., Hofker, M.H., Monaco, A.P., Kunkel, L.M., Ray, P., Worton, R., Wieringa, B., Bakker, E., Pearson, P.L. (1986) Cell, 47, 499-504.
9. Kenwrick, S., Patterson, M., Speer, A., Fischbeck, K. and Davies, K. (1987) Cell, 48, 351-357.
10. Murray, N.E. (1983) in the Bacteriophage Lambda, vol.2 (Hendrix, R. et al. eds) Cold Spring Harbour Laboratory, New York.
11. Erickson, J.M., Rushford, C.L., Dorney, D.J., Wilson, G.N. and Schmickel, R.D. (1981) Gene, 16, 1-9.
12. Maniatis, T., Fritsch, E.T. and Sambrook, J. (1982) in Molecular CLoning. Cold Spring Harbor Laboratory.
13. Scalenghe, F., Turco, E., Edström, J.E., Pirrota, V. and Melli, M. (1981). Chromosoma 82, 205-216.
14. Oberlé, I., Camerino, G., Kloepfer, C., Moisan, J.P., Grzeschik, K.H., Hellkuhl, B., Hors-Cayla, M.C., Van Cong, N., Weil, D. and Mandel, J.L. (1986) Human Genetics, 72, 43.
15. Litt, M. and White, R.L. (1985) Proc. Natl. Acad. Sci. USA, 82, 206-6210.
16. Maxam, A. and Gilbert, W. (1980) in Methods in Enzymology (eds. L. Grossman and K. Moldave) Vol. 65, Academic press, New York, pp. 499-560.
17. Hanahan, D. and Meselson, M. (1980) Gene, 10, 63-67.
18. Burmeister, M. and Lehrach H. (1986) Nature, 324, 582-585.
19. Ohno, S. (1969) Annu. Rev. Genet., 3, 495.
20. Dodgson, J.B., Strommer, J. and Engel, J.D. (1979). Cell, 17, 879-887.
21. Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S. and Sharp, P.A. (1986) Ann. Rev. Biochem. 55, 1119-150.
22. Dodgson, J.B. and Engel, J.D. (1983) J. Biol. Chem., 258, 4623-4629.
23. Erbil, C. and Niessing, J. (1983) EMBO Journal, 2, 1339-1343.
24. Mandel, J.L., Arveiler, B., Camerino, G., Hanauer, H., Heilig, R., Koenig, M. and Oberlé, I. (1986) in Cold Spring Harbor Symp. Quant. Biol., 51, 195-203.
25. Smith, T.J., Wilson, L., Kenwrick, S.J., Morrest, S.M., Speer, A., Coutelle, C. and Davies, K.E. (1987) Nucleic Acids Research, 15, 2167-2174.
26. Friend, S.H., Bernards, R., Rogelj, S., Weinberg, R.A., Rapaport, J.M., Albert, D.M. and Dryja, T.P. (1986) Nature, 323, 643-646.
27. Estivill, X., Farrall, M., Scambler, P.J., Bell, G.M., Hawley, K.M.F., Lench, N.J., Bates, G.P., Kruyer, H.C., Frederick, P.A., Stanier, P.,

Watson, E.K., Williamson, R. and Wainwright, B.J. (1987) Nature, 326, 840-845.
28. Heilig, R., Lemaire, C., Mandel, J.L., Dandolo, L., Amar, L. and Avner, P. (1987) Nature, 328, 168-170.
29. Monaco, A.P., Bertelson, C.J., Colletti-Feener, C. and Kunkel, L.M. (1987) Human Genetics, 75, 221-227.
30. Worton, R.G., Burghes, A.H.M., Logan, C., Belfall, B., Duff, C., Thompson, M.W. and Ray, P.N. (1987) Hum. Gene Mapping 9, abstract