lysed in an extraction buffer (TEB) (120 mM piperazino-*N*,*N*′-bis[2-ethanesulfonic acid], 40 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid, 20 mM EGTA, 4 mM $MgCl_2$, 0.25% deoxycholic acid, 1% Triton X-100) plus protease inhibitors (1 mM benzamidine, 1 mM leupeptin, and 1 mM aprotinin) and scraped off the dish with a rubber policeman. The suspension was vigorously vortexed to disrupt the cytoskeletal components and spun at 10,000*g* for 15 min at 4°C to pellet the nuclei. The protein concentrations of the supernatants were determined by the method of M. Bradford [*Anal. Biochem.* **72**, 248 (1976)], and equal amounts of each cell lysate were boiled in denaturing SDS buffer and then separated by electrophoresis on a 5 to 15% SDS-polyacrylamide gel and transferred to Immobilon-P (Millipore, Bedford, MA) by the method of H. Towbin, T. Staehelin, and J. Gordon [*Proc. Natl. Acad. Sci. U.S.A.* **76**, 4350 (1979)]. The membrane was incubated at room temperature for 1 hour either with a monoclonal antibody (2 μg/ml) to human gelsolin [D. J. Kwiatkowski, *J. Biol. Chem.* **263**, 13857 (1988)] or with a 1:500 dilution of rabbit antiserum to rat gelsolin (*17*), and then with 0.5 μC/ml of the appropriate ^{125}I-labeled secondary goat immunoglobulin G (IgG) (New England Nuclear, Boston, MA), with washes of 0.2% Tween PBS in between. After further washes, the membrane was exposed to autoradiograph film for up to 12 hours and the film developed. The autoradiograph bands were scanned with a laser densitometer (LKB Pharmacia, Piscataway, NJ) for quantitation of intensity. We used known amounts of mouse gelsolin or human gelsolin to compare band intensity and estimate the amount of native mouse gelsolin and transfected human gelsolin. Expression of native NIH 3T3 gelsolin was not changed in the CGP cells. On the basis of these calculations and estimates of total cell protein and actin content, we determined that actin accounts for 4% of total cell protein and gelsolin 0.1% of total cell protein in NIH 3T3 cells, giving a molar ratio of gelsolin:actin of 1:84, and that in the cell line C5, for example, this ratio was increased to 1:37.

11. F-actin stress-fiber architecture was visualized by rhodamine phalloidin staining. Cells were plated on cover slips, fixed with 0.1% paraformaldehyde and permeabilized in acetone at −20°C, and incubated at room temperature with rhodamine phalloidin (50 U/ml) (Polysciences). The cells were viewed on a Zeiss Axioplan microscope equipped for epifluorescence. Two-dimensional gel electrophoresis was performed by Kendrick Laboratories, Madison, WI, with 2.0% ampholines, pH 4 to 8.
12. Migration through a membrane in response to a chemoattractant was assayed with the use of a 48-well chamber (Nuclepore, Pleasanton, CA) with a 5-μm polycarbonate filter. Cells were trypsinized, then counted, and 5×10^4 cells per well were loaded in the top wells in media plus or minus serum supplementation. The bottom wells were similarly filled with media with or without calf serum so that the cells were exposed to either a gradient, reverse gradient, or no gradient of serum factors. The chambers were incubated for 2 hours at 37°C; the membranes were removed and stained and then examined with a Zeiss Axiovert microscope with a ×40 objective. The number of cells that had migrated through the membrane was counted for each well.
13. A. Kupfer, D. Louvard, S. J. Singer, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2603 (1982).
14. Elimination of cytoskeletal genes in lower eukaryotes has provided strong evidence for involvement or noninvolvement of the gene products in certain cell functions [D. A. Knecht and W. F. Loomis, *Science* **236**, 1081 (1987); A. De Lozanne and J. A. Spudich, *ibid.*, p. 1086; A. A. Noegel and W. Witke, *Dev. Genet.* **9**, 531 (1988); E. Andre *et al.*, *J. Cell Biol.* **108**, 985 (1989); D. Wessels *et al.*, *Dev. Biol.* **128**, 164 (1988)].
15. J. A. Cooper *et al.*, *J. Cell Biol.* **104**, 491 (1987); A. Huckriede *et al.*, *Cell Motil. Cytoskeleton* **16**, 229 (1990); Z. Franck, M. Footer, A. Bretscher, *J. Cell Biol.* **111**, 2478 (1990).
16. For this transient expression experiment, the cytoplasmic gelsolin cDNA was inserted in the vector CDM8, and this construct was used to transfect NIH 3T3 cells by the DEAE-dextran method [D. J. Kwiatkowski, P.A. Janmey, H. L. Yin, *J. Cell Biol.* **108**, 1717 (1989)]. Cells were prepared for staining 2 days later as described (*11*). After permeabilization, cells were incubated with monoclonal antibody (50 μg/ml) to human gelsolin, and then with fluorescein-conjugated goat anti-mouse IgG (50 μg/ml) (Cappel Labs, Cochranville, PA), with washes in PBS with 0.2% gelatin in between. After rinsing, the cells were incubated in rhodamine phalloidin (50 U/ml), rinsed again, and viewed by epifluorescence.
17. We thank K. Weimer for technical help. We thank C. Chaponnier, University of Geneva, for providing the antibody to rat gelsolin. This research was supported by NIH grants HL19429, AI28465, and HL07680. T.P.S. is an American Cancer Society Clinical Research Professor, and D.J.K. is an Established Investigator of the American Heart Association.

10 October 1990; accepted 10 January 1991

Isolation of Sequences That Span the Fragile X and Identification of a Fragile X–Related CpG Island

D. HEITZ,* F. ROUSSEAU,* D. DEVYS, S. SACCONE, H. ABDERRAHIM, D. LE PASLIER, D. COHEN, A. VINCENT, D. TONIOLO, G. DELLA VALLE, S. JOHNSON, D. SCHLESSINGER, I. OBERLÉ, J. L. MANDEL

Yeast artificial chromosomes (YACs) were obtained from a 550-kilobase region that contains three probes previously mapped as very close to the locus of the fragile X syndrome. These YACs spanned the fragile site in Xq27.3 as shown by fluorescent in situ hybridization. An internal 200-kilobase segment contained four chromosomal breakpoints generated by induction of fragile X expression. A single CpG island was identified in the cloned region between markers DXS463 and DXS465 that appears methylated in mentally retarded fragile X males, but not in nonexpressing male carriers of the mutation nor in normal males. This CpG island may indicate the presence of a gene involved in the clinical phenotype of the syndrome.

THE FRAGILE X MENTAL RETARDATION syndrome is the most frequent cause of inherited mental retardation (with an incidence of one in 1500 newborn males) (*1*). The diagnosis is based on the presence of a fragile site on the X chromosome, at Xq27.3, induced in vitro by culture conditions affecting deoxynucleotide synthesis (*2*). Partial penetrance is observed in males and females and varies in different sibships, even within the same family (*3*). Many hypotheses have been proposed to account for the unique characteristics of the inheritance of this syndrome. In particular, Laird (*4*) suggested that the fragile site is a region of late replication, resulting from a local inability to reactivate a previously inactive X chromosome, during oogenesis. Using pulsed-field gel electrophoresis (PFGE), we have recently obtained data that support an imprinting mechanism, and found restriction sites that appear methylated in males who express the mutation, but not in normal males whether or not they carry the mutation (*5*). We report here the isolation of a DNA region that spans the fragile site and the identification of a CpG island which appears critically involved in the expression of the syndrome.

We have recently cloned two probes, St677 (DXS463) and Do33 (DXS465), that map within a 3-Mb Not I fragment and that flank breakpoints on the X chromosome purported to be at or very near the fragile X site (*5, 6*). The distal probe Do33 detects abnormal PFGE patterns in mentally retarded fragile X patients and lies within 120 kb of a region important for the expression of the syndrome (*5*). We have now isolated four yeast artificial chromosomes (YACs) containing St677, Do33, or both probes by direct colony screening (*7*) or by polymerase chain reaction (PCR) screening of YAC pools (*8*). The two larger clones (141H5 and 209G4) were obtained from the total human library of the Centre d'Etude du Polymorphisme Humain (CEPH), and had been derived from a normal male (*9*), and mapped by PFGE after digestion with rare cutting restriction enzymes. We also analyzed the two smaller clones (XY120 and XY530), which were obtained from a Xq24–q28 library that had been derived from the X chromosome of a fragile X

D. Heitz, F. Rousseau, D. Devys, A. Vincent, I. Oberlé, J. L. Mandel, Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Institut de Chimie Biologique. Faculté de Médecine, 67085 Strasbourg Cédex France.
S. Saccone and G. Della Valle, Dipartimento Genetica e Microbiologia, University of Pavia, Pavia 27100 Italy.
H. Abderrahim, D. Le Paslier, D. Cohen, Centre d'Etude du Polymorphisme Humain, Paris 75010 France.
D. Toniolo, Istituto di Genetica, Biochimica, ed Evolutionistica, CNR, Pavia 27100 Italy.
S. Johnson and D. Schlessinger, Washington University School of Medicine, Center for Genetics in Medicine, St. Louis, MO 63110.

*These two authors contributed equally to the report.

patient (*10*) (Fig. 1).

Two YAC clones (209G4 and XY530) were initially isolated by screening with St677, and both contained also probe 2-34 (DXS477) (*6*). The smaller size of XY530 allowed mapping of the two probes, which were separated by 25 to 45 kb. To further study the 425-kb YAC 209G4, we generated end probes by inverse PCR (*11*). These probes were hybridized to a physical localization panel (*6*), and both mapped to the fragile X region, eliminating the possibility of an artifactual ligation of unrelated sequences during cloning. Moreover, they appeared separated by several chromosomal breakpoints (Fig. 1). Four of these correspond to the cell lines generated by Warren *et al.* (*12*) who selected for breakage of a fragile X chromosome after induction of the fragile X site. These breakpoints are in fact located between 2-34 and the right arm of 209G4, and their clustering within 200 kb indicates that they are indeed specific for the fragile site region.

To verify that 209G4 spans the fragile site, fluorescent in situ hybridization was performed on metaphases obtained from five fragile X patients after induction with fluorodeoxyuridine (Fig. 2). Of the fluorescent signals scored in 120 metaphases, 49% were on the fragile X region, evenly distributed on the fragile site gap itself, on the proximal or on the distal side of the gap (Table 1). The same distribution was observed in all patients. The finding on the same chromosome of well-separated hybridization signals flanking the fragile site was especially striking given the size of the probe (Fig. 2, d and f). A control YAC clone containing the proximal probe 2-71 (DXS476) which maps within the same 3-Mb Not I fragment as 2-34, St677, and Do33 (*13*) showed in most cases hybridization proximal to the fragile site, although a signal was sometimes found over the gap (Fig. 2j).

A larger 520-kb YAC (141H5), isolated later by PCR screening for Do33 (*14*), was also shown to contain St677 and 2-34 and to extend further than 209G4 by 25 kb in the telomeric direction. In fact Do33 is within 10 kb from the right end probe of 209G4 (*15*). The maps of the two YACs are in very good agreement in the overlapping region. The tight clustering of rare restriction sites reveals the presence of a CpG island about midway between probes 2-34 and Do33. Another region with many rare sites is found closer to Do33, but it appears spread over about 20 kb (Eag I and Bss HII sites) and may thus not be a true CpG island (*16*). Preliminary mapping of the two smaller YACs derived from the genome of a fragile X patient showed good correspondence with the map established from the CEPH YACs, given the relative lack of precision in PFGE (Fig. 1).

We have recently reported that Do33 detects, in normal males, 620-kb Bss HII and Sac II fragments and a 120-kb Eag I fragment which all appear missing or present in much lower amounts in mentally retarded fragile X males (*5*). The most likely explanation is that at least one site for each of these enzymes is specifically methylated in patients and thus resistant to digestion. Double digestion of normal genomic DNA indicated that the Bss HII and Sac II sites should be very close. The proximal CpG island between 2-34 and Do33 is thus a

Table 1. Distribution of the hybridization signals of 209G4 and XY300 DNAs on the Xq27.3 fragile site region. Metaphase spreads from four fragile X males and a female were used. Only metaphases expressing the fragile X chromosome and showing hybridization signals on the fragile site region were scored. These signals correspond to 94% of those found on X chromosome and to 52% of the total fluorescent spots detected in the analyzed metaphases. Signals were recorded as proximal, central, or distal, depending on their location on the centric chromatids, over the gap, or on the acentric fragment, respectively. All patients are mentally retarded and their fragile X expression was 20, 5, 7, 14, and 6% for subjects 1, 2, 3, 4, and 5, respectively.

Subject	Probe	Proximal	Central	Distal	Metaphases analyzed
1	YAC 209G4	32	29	37	69
2	YAC 209G4	11	12	12	15
3	YAC 209G4	9	6	7	16
4*	YAC 209G4	16	10	20	20
5	XY 300	16	5	0	15

*Mother of case 3.

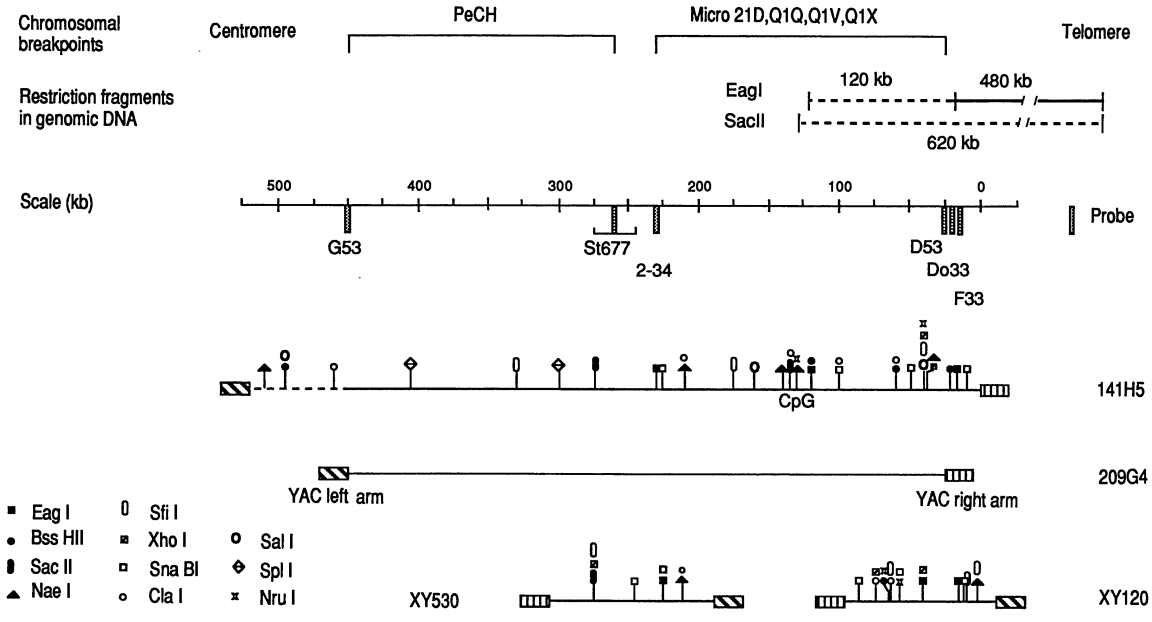

Fig. 1. Map of a 550-kb region spanning the fragile X site cloned in four YACs. The interval containing the breakpoint of cell lines Micro21D, Q1Q, Q1V, and Q1X (*12*) is shown. Restriction fragments detected in PFGE of genomic DNA are represented for Sac II and Eag I digests (fragment present (———) or specifically methylated (– – – –) in fragile X patients). The location of probes used for mapping is indicated under the scale. Probes specific for the right or left arm were also used for mapping with partial digests. G53 and D53 are end probes of 209G4. Two YACs (141H5 and 209G4) contain DNA from a normal male. The two smaller YACs (XY120 and XY530) are derived from a fragile X chromosome. The broken line in 141H5 indicates that this left-end part of the YAC has not been checked to be derived from the fragile X region. Sites with filled symbols are those most specific for CpG islands. "CpG" indicates the cluster methylated in fragile X patients. We did not map all the restriction sites for each tested enzyme. The most telomeric Bss HII site in 141H5 was difficult to digest. Some minor differences were observed between YACs isolated from the normal male and the fragile X patient. For instance, a Sal I site (position 40 kb) was cut in 141H5 and in 209G4 but was not detected in XY120. Fine ordering of rare cutting sites in the most telomeric CpG cluster appear somewhat different in 141H5 and XY120 but we do not know yet if these differences are real.

good candidate for being the site of abnormal methylation since it is the only place in the cloned region where sites for all three enzymes are clustered. Further support was obtained by means of probe F33, which had been isolated by PCR jumping (*17*) from the right arm of 141H5. This probe is adjacent and distal to the most telomeric Eag I site (see Fig. 1). Southern blot analysis showed that this Eag I site is unmethylated in leukocyte DNA from normal or fragile X males and is about 50% methylated in female DNA (*15*). This is the pattern found in CpG islands that are unmethylated on the active X and methylated on the inactive X chromosome (*18*). F33 detects in genomic DNA a 480-kb Eag I fragment present in both normal and affected males but hybridizes to the same differentially methylated 620-kb Sac II fragment as Do33 (Fig. 3). We deduce the position of the 120-kb Eag I fragment detected by Do33 and conclude that it is the Eag I site within the proximal CpG island that is affected by differential methylation (Fig. 1).

We have isolated two YAC clones that hybridize to sequences flanking the fragile site, and that also flank four chromosomal breakpoints generated by induction of fragile X expression (*12*). The clustering of these breakpoints and the similarity in in situ hybridization patterns observed in different fragile X chromosomes suggest that the affected region is always the same and less than 200 kb long. The folding of the chromatin in metaphase chromosomes does not resolve hybridization signals for probes closer than few megabases (*19*). The fact that in some chromosomes, well-separated signals were seen with 209G4 on both sides of the fragile site (Fig 2, d and f), indicates that the corresponding DNA is highly unfolded. The gap observed at the fragile site is usually from 0.2 to 1 μM in length. Given the size of the probe (425 kb), our data could account for an unfolding to the 30-nm chromatin fiber, as previously suggested (*20*). The gap may thus be occupied by 10 to 100 kb of DNA.

The highly specific localization of the Xq27.3 fragile site is mirrored by the highly specific abnormalities in methylation at probably a single CpG island. Three other Eag I sites present in 141H5 do not show abnormal methylation in fragile X patients. The Eag I sites localized in Fig. 1 at position 230 kb (adjacent to 2-34) and 40 kb appear methylated (*5, 21*) whereas the site at 16 kb is unmethylated, in genomic DNA from both normal and fragile X males. The centromeric CpG island, between 2-34 and Do33, appears to have a critical role in expression of the fragile X syndrome and it will be of utmost importance to find if this island signals the presence of an adjacent gene. Four polymorphic probes (St677, 2-34, Do33, and F33) lie within 120 kb of this CpG island and may be used for highly reliable carrier or prenatal diagnosis (*6, 22*).

The fragile site expression, the tendency to site-specific chromosomal breakage in hybrids, as well as the observed differential methylation, must be secondary effects of a primary mutation. The search for the mutations is still open. The detailed comparison of the YAC clones originating from a normal and a fragile X chromosome, as well as the use of probes derived from those YACs to analyze with frequent cutting enzymes large series of patients, should allow the detection of such mutations. It appears unlikely that point mutations could account for the predicted very high mutation frequency (*3, 23*) and for the magnitude of the chromosomal effect.

Note added in proof. Finer mapping has

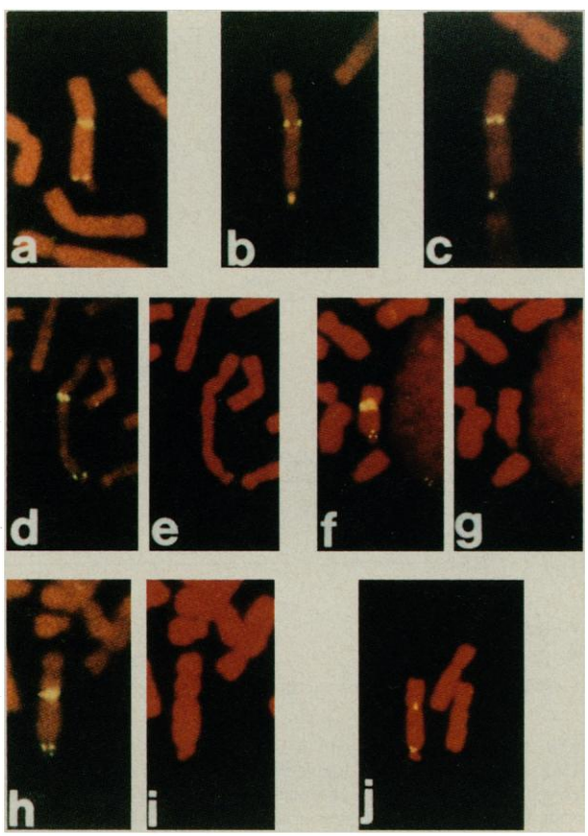

Fig. 2. Representative partial metaphases showing the location of YAC209G4 and XY300 on fragile X chromosomes. In situ hybridization with 209G4 showing signals located **a**, proximal; **b**, central; **c**, distal; **d** and **f**, proximal and distal; **h**, central and distal to the fragile site. Panels **e**, **g**, and **i** are the same partial metaphases shown, respectively in **d**, **f**, and **h** photographed with Leitz N2.1 filter bloc, in order to better visualize the gap. Panel **j**, in situ hybridization with XY300 showing the most distal signal observed with this probe. XY300 (400 kb) was isolated from the Xq24–q28 library (*10*) by screening with probe 2-71 (DXS476). Metaphase chromosome spreads were obtained, by standard cytogenetic methods, from phytohemagglutinin-stimulated lymphocyte cultures of patients with fragile X syndrome. Fragile X was induced by treatment with 5-fluoro-2′ deoxyuridine. YAC DNA, isolated by PFGE, or total yeast DNA were nick-translated with biotin-16-dUTP and hybridized in the presence of 1000-fold excess sonicated human DNA. Plasmid pBX5 (*24*), containing pericentromeric X sequences, was added to the hybridization and detection with fluorescein isothiocyanate–conjugated avidin was as described (*25*). Chromosomes were counterstained with propidium iodide in antifade medium.

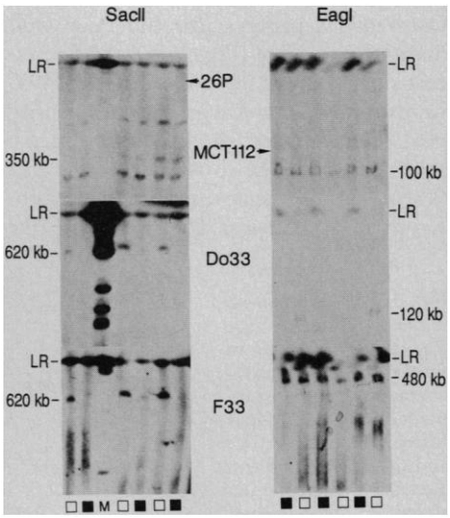

Fig. 3. Hybridization of probe Do33 and F33 to genomic DNA from normal males and fragile X male patients. DNA from lymphoblastoid cell lines (Sac II digests) or from leukocytes (Eag I digests) was prepared in low melting agarose blocks. The Sac II digests were separated by electrophoresis on a CHEF apparatus in 0.25× tris Borate EDTA (TBE) at 130V for 39 hours with a pulse time of 60 s. The Eag I digests were separated in a Waltzer apparatus (*26*) in 0.25 × TBE at 110 V for 15 hours with a pulse time of 30 s. Lambda ladder (Eag I digests) as well as yeast chromosomes (Eag I digests and slot "M" in Sac II digests) were used as size markers. DNA from normal males (open box) and from mentally retarded fragile X males (closed box) belonging to independent families were analyzed with autosomal probes 26P (D9S5) and MCT112 (D9S15) (*27*) in control hybridization. The same membranes were hybridized to probe Do33 (DXS465) and F33. The size of the restriction fragments are indicated as well as the limit of resolution (LR) for each gel.

shown that Bss HII, Eag I, Nae I and, Sac II sites are clustered in less than 2 kb in the proximal CpG-rich region, which is thus a true CpG island.

REFERENCES AND NOTES

1. R. L. Nussbaum and D. H. Ledbetter, *Annu. Rev. Genet.* **20**, 109 (1986).
2. G. R. Sutherland, *Am. J. Med. Genet.* **31**, 125 (1979); ———, E. Baker, A. Fratini, *ibid.* **22**, 433 (1985).
3. S. L. Sherman, P. A. Jacobs, N. E. Morton, U. Froster-Iskenius, P. N. Howard-Peebles, *Hum. Genet.* **69**, 289 (1985).
4. C. D. Laird, *Genetics* **117**, 587 (1987); ———, M. M. Lamb, J. L. Thorne, *Am. J. Hum. Genet.* **46**, 696 (1990).
5. A. Vincent *et al.*, *Nature* **349**, 624 (1991).
6. F. Rousseau *et al.*, *Am J. Hum. Genet.* **48**, 108 (1991).
7. B. H. Brownstein *et al.*, *Science* **244**, 1348 (1989).
8. E. D. Green and M. V. Olson, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 1213 (1990).
9. H. M. Albertsen *et al.*, *ibid.*, p. 4256.
10. F. E. Abidi, M. Wada, R. D. Little, D. Schlessinger, *Genomics* **7**, 363 (1990); M. Wada *et al.*, *Am. J. Hum. Genet.* **46**, 95 (1990).
11. T. Triglia, M. G. Peterson, D. J. Kemp, *Nucleic Acids Res.* **16**, 8186 (1988). Total yeast DNA was digested to completion with Eco RV, and circularized at a concentration of 3 μg/ml with 6 U Weiss of T4 DNA ligase (Biolabs) in a volume of 20 μl. Fifteen nanograms of the ligated DNA was amplified by PCR using inverted primer pairs corresponding to the right arm or the left arm of the pYAC4 vector [D. T. Burke, G. F. Carle, M. V. Olson, *Science* **236**, 806 (1987)]. The PCR was performed in 50 μl with 2.5 U of AmpliTaq DNA polymerase (Perkin-Elmer Cetus) with the following amplification conditions: a 5-min initial denaturation at 94°C, 40 cycles of 1-min denaturation at 94°C, 1 min annealing at 60°C, 4 min elongation at 72°C (extended for 2 s per cycle) and a final elongation step of 7 min, at 72°C. The sequences of the two right arm primers were: NL10: 5′-CTCCCGGGGGCGAGTCGAACGCCCG-3′ and NL21: 5′-GCCCAGTATTCTTAACCCAACTGCACAG-3′. After amplification, the 2.9-kb fragment corresponding to the probe D53 was cloned as a 2.4-kb Eco RI–Pst I fragment in pBluescript KS+. The sequences of the two left arm primers were: NL12: 5′-CAATTAAATACTCTCGGTAGCCAAG-3′ and NL13: 5′-CCGGTCCTTCCAAGATGGTTCAGAGTG-3′. After amplification, a 1.4-kb fragment corresponding to the probe G53 was obtained and used as a LMP purified probe.
12. S. T. Warren *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 3856 (1990).
13. A. Vincent, unpublished results.
14. The PCR was performed in 50 μl with: 50 ng of a DNA pool of 400 YACs, 0.5 μg of each primers OI39 (5′-GTGAAGCAATTTGAGTTGGTGGGAG-3′) and OI41 (5′-CCCAAGTTTAGCTTCCGTAGAAGCC-3′), 2.5 U of AmpliTaq DNA polymerase with the following amplification conditions: a 5-min initial denaturation at 94°C, 35 cycles of 1-min denaturation at 94°C, 1-min annealing at 60°C, 8-min elongation at 72°C (extended for 2 s per cycle), and a final elongation step of 10 min at 72°C. Identification of YAC 141H5 from the positive YAC pool was performed by dot blot hybridization.
15. D. Heitz *et al.*, unpublished data.
16. A. P. Bird, *Trends Genet.* **3**, 342 (1987).
17. R. P. Kandpal, H. Shukla, D. C. Ward, S. M. Weissman, *Nucleic Acids Res.* **18**, 3081 (1990); A. Poustka and H. Lehrach, in *Genetic Engineering: Principles and Methods*, J. Setlow, Ed. (Plenum, New York, 1988), vol. 10, p. 169. Total yeast DNA, prepared in low melting agarose block (about 5 μg per block) was partially digested with Eag I in 1 mM $MgCl_2$ and circularized at a concentration of 200 ng/ml with 30 U Weiss of T4 DNA ligase in a volume of 1 ml. The ligase was inactivated at 68°C and 10 ng of ligated DNA was digested by Eco RI after adjusting the NaCl concentration to 100 mM. Afterwards, the enzyme was inactivated at 68°C for 20 min and DNA precipitated with NaCl, glycogen and ethanol. The digested DNA was then recircularized at a concentration of 1 ng/μl with 1.5 U Weiss of ligase in a total volume of 50 μl. Ligation product (5 ng) was amplified in 50 μl as described (*11*) in presence of right-arm-specific inverse-PCR primers NL9 and OR19 (NL9: 5′-CGAGTCGAACGCCCGATCTCAAG-3′ and OR19: 5′-GGCCATTATCGCCGGCATGG-3′). The 1.1-kb fragment corresponding to probe F33 was cloned as a 1-kb Eco RI–Eag I fragment in pBluescript KS+.
18. S. F. Wolf *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2806 (1984); D. Toniolo *et al.*, *EMBO J.* **7**, 401 (1988).
19. P. Lichter *et al.*, *Science* **247**, 64 (1990); J. Shaw, P. Lichter, A. J. Driesel, R. Williamson, S. Chamberlain, *Cytogenet. Cell Genet.* **53**, 221 (1990).
20. C. J. Harrison, E. M. Jack, T. D. Allen, R. Harris, *J. Med. Genet.* **20**, 280 (1983).
21. E. Maestrini *et al.*, *Genomics* **8**, 664 (1990).
22. F. Rousseau, unpublished results.
23. J. A. Sved and C. D. Laird, *Am. J. Hum. Genet.* **46**, 443 (1990).
24. H. F. Willard *et al.*, *Nucleic Acids Res.* **11**, 2017 (1983).
25. V. Montanaro *et al.*, *Am. J. Hum. Genet.*, in press.
26. E. M. Southern *et al.*, *Nucleic Acids Res.* **15**, 5925 (1987).
27. R. Fujita *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 1796 (1990).

28. We wish to thank C. Kretz for her expert technical assistance, S. T. Warren for the generous gift of DNA from hybrid cell lines, and R. Tupler and C. Barletta for help in the in situ hybridization experiments. We are also grateful to G. R. Sutherland for communicating his data about XTY26 previously to publication. Supported by grants from Ministère de la Recherche et de la Technologie (87C.0397) and Caisse Nationale d'Assurances Maladie to J.L.M. from Association Française contre les Myopathies to D.C., from the Associazione Italiana per la Ricerca sul Cancro and Progetti Finalizzati CNR to G.D.V. and D.T. F.R. was supported by a postdoctoral fellowship from the Medical Research Council of Canada.

8 January 1991; accepted 8 February 1991

Subtle Cerebellar Phenotype in Mice Homozygous for a Targeted Deletion of the *En-2* Homeobox

A. L. JOYNER, K. HERRUP, B. A. AUERBACH, C. A. DAVIS, J. ROSSANT

The two mouse genes, *En-1* and *En-2*, that are homologs of the *Drosophila* segmentation gene *engrailed*, show overlapping spatially restricted patterns of expression in the neural tube during embryogenesis, suggestive of a role in regional specification. Mice homozygous for a targeted mutation that deletes the homeobox were viable and showed no obvious defects in embryonic development. This may be due to functional redundancy of *En-2* and the related *En-1* gene product during embryogenesis. Consistent with this hypothesis, the mutant mice showed abnormal foliation in the adult cerebellum, where *En-2*, and not *En-1*, is normally expressed.

TWO HIGHLY CONSERVED HOMEOBOX-containing genes, *En-1* and *En-2*, have been identified in mouse (*1*) and other vertebrates (*2–5*), based on their sequence similarity to the *Drosophila* genes, *engrailed (en)* and *invected (inv)* (*6, 7*). All *en*-like genes contain four conserved protein domains including the homeobox (Fig. 1A), suggesting that they have similar biochemical functions. The *Drosophila en* gene has been shown by mutational and mosaic analysis to be required for embryonic segmentation (*8*) and later development of the nervous system (*9*). The *En* and *inv* genes are coexpressed in a pattern consistent with these roles (*7, 10*).

Comparative expression studies have shown that in vertebrates, from zebrafish to mice, the *En* genes are expressed in a highly specific pattern in the developing embryo (*3–5, 11–14*), indicating conservation of function. During and after neural tube closure, *En-1* and *En-2* are expressed in a band of cells spanning the junction between mid- and hindbrain, apparently defining a specific spatial domain in the developing nervous system. *En-1*, unlike *En-2*, is also expressed in specific domains within the spinal cord, somites, and limbs, beginning at 9.5 days of development. In the mouse, expression of both *En* genes in the brain later becomes progressively limited to different specific groups of neurons (*11, 12*). In the adult, *En-1* and *En-2* continue to be coordinately expressed in a number of motor nuclei in the pons region and in cells within the substantia nigra. However, *En-2* alone is expressed in the granule cell layer of the cerebellum.

The two phases of *en* expression in neural development have been suggested to reflect roles in both compartmentalization of the developing neural tube and later specification of particular neuronal populations (*11*). Proof of such roles requires mutational analysis in animals. The advent of targeted mutagenesis by homologous recombination in mouse embryonic stem (ES) cells has

A. L. Joyner, C. A. Davis, J. Rossant, Division of Molecular and Developmental Biology, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario, Canada M5G 1X5, and Department of Molecular and Medical Genetics, University of Toronto, Canada.
K. Herrup, Shriver Center for Mental Retardation Inc., 200 Trapelo Road, Waltham, MA 02254.
B. A. Auerbach, Division of Molecular and Developmental Biology, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario, Canada M5G 1X5.