3 RESULTS

3.1 Clinical and molecular studies in patients with disease-associated X;autosome translocations: the search for candidate MRX genes

As outlined earlier, disease-associated balanced translocations provide a practical starting point in the search for disease genes. This is especially true for translocations that involve the X chromosome, due to the fact that X-chromosomal genes are generally present as single functional copies. The initial aim of this study was to screen several patients with MR and associated balanced X;autosome translocations in order to identify one or more candidate genes for MRX. In the first section, the clinical picture and karyotype of each of three such patients is described. In each case, the presence of a balanced translocation was confirmed using fluorescence *in situ* hybridisation (FISH). In order to select the patients for which a causal relationship between disorder and translocation was highly probable, we used a combination of molecular and *in silico* analyses to further investigate the breakpoint region identified in each patient. More precisely, we searched for specific candidate X-chromosomal genes for which expression was affected by the breakpoint, either due to direct disruption of the coding sequence, or indirectly, by disruption of a critical regulatory element in the neighbouring sequence. In two of the three patients investigated, no such gene was found; however, preliminary analysis of Patient 3 led to the identification of an obvious candidate MRX gene, which was studied further (see section 3.2).

3.1.1 Patient 1 with translocation t(X;7)(q22;p22)

Patient 1 is a female diagnosed with mild mental retardation and ovarian dysgenesis, carrying a balanced translocation t(X;7)(q22;p22), depicted in the ideogram in Figure 5A. Examination of metaphase chromosomes from the patient cell line, and subsequent FISH-mapping of X-chromosomal YAC, PAC, and BAC clones both proximal and distal to the breakpoint confirmed the presence of this translocation. Sequential mapping of genomic clones led to the identification of a breakpoint-spanning PAC clone RP4-596C15 (Accession # AL031387), which hybridised to both the normal X chromosome and the derivative chromosomes (Figure 5B).

Fully-annotated sequence of the X-chromosomal breakpoint-spanning clone enabled thorough *in silico* analysis of the region, depicted in Figure 6A, which contains two genes,

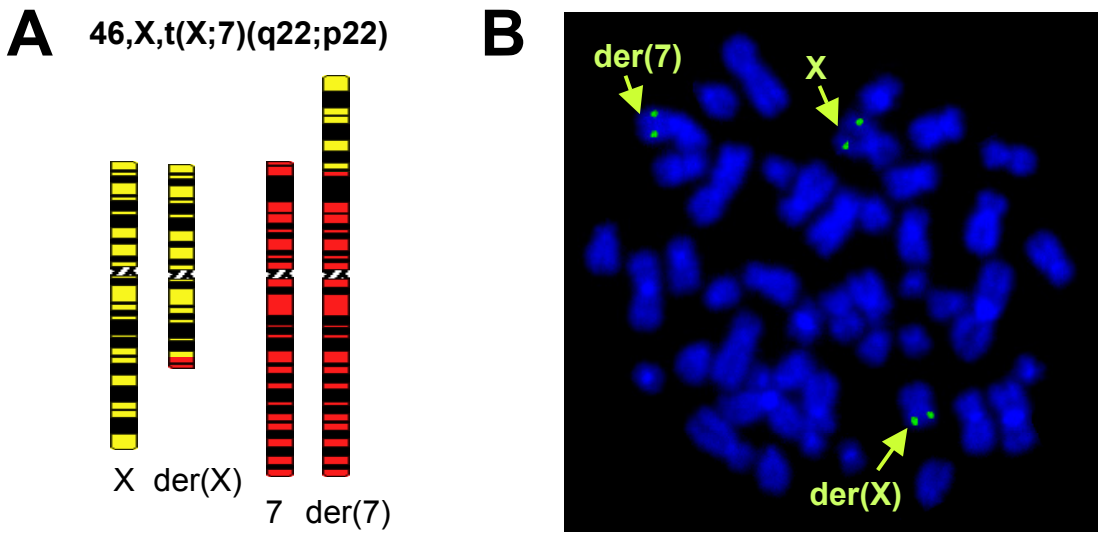
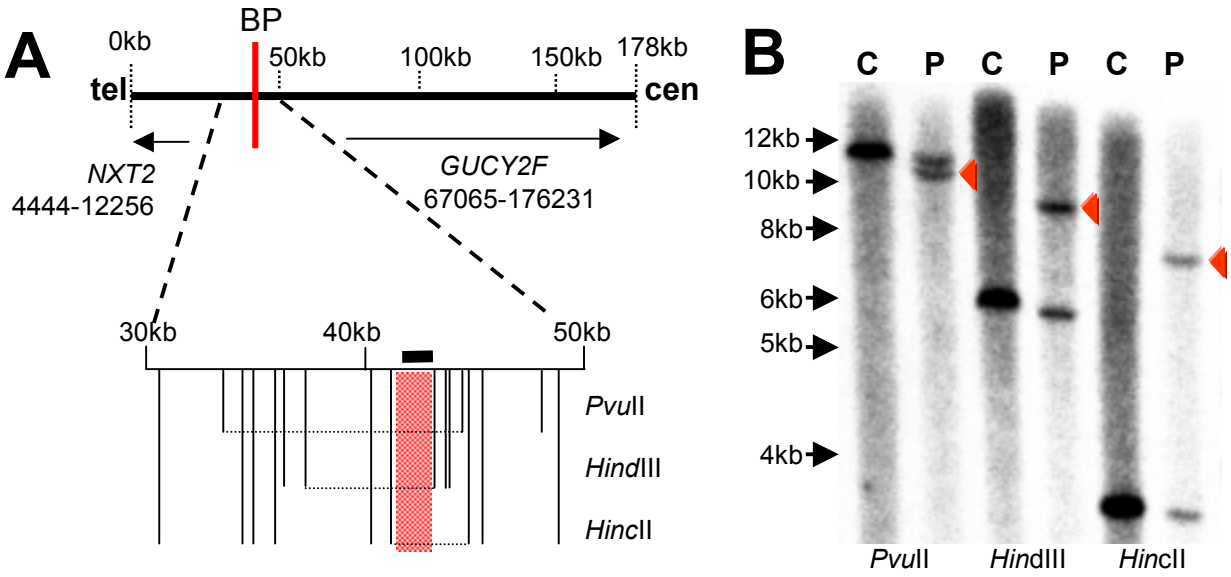


Figure 5: Patient 1 molecular cytogenetic studies
A: Ideogram of Patient 1 metaphase chromosomes. Metaphase spreads were used for fluorescence *in situ* hybridisation (FISH). **B:** FISH result for X chromosome overlapping PAC clone RP4-596C15 with signals on the normal X, der(X), and der(7) as indicated.

the retinal guanylyl cyclase *GUCY2F* and the nuclear transport factor (NTF) 2-like export factor 2 (*NXT2)*. In order to determine if either of these genes was affected by the translocation, further fine-mapping of the breakpoint was necessary. Given that there is no known visual disability in the patient, and that the expression of *GUCY2F* is predominantly retinal (Yang et al. 1996), the genomic region that this gene occupies was excluded in subsequent analyses, and kilobases (kb) 1-60 of clone RP4-596C15 were considered first. PCR products from within this region were hybridised successively to Southern blots with patient and control genomic DNA. As shown in Figure 6B, using the probe corresponding to nucleotides 41632-43225 of clone RP4-596C15, in both patient and control *Pvu*II, *Hin*dIII, and *Hin*cII digests, fragments of the expected sizes, 11.1, 5.9, and 3.6 kb, respectively, were observed. In the patient, this probe also hybridised to aberrant patient-specific fragments with approximate sizes of 10, 8.5, and 7.0 kb, respectively, indicating that the breakpoint on the patient's disrupted X chromosome lies within the sequence common to these three expected fragments (see Figure 6A). The breakpoint could thereby be localised to within the 2500 base pairs between nucleotides 41326 and 43827 of clone RP4-596C15. The *Hin*cII site marks the telomeric boundary of this region, while the *Hin*dIII site marks the centromeric boundary.

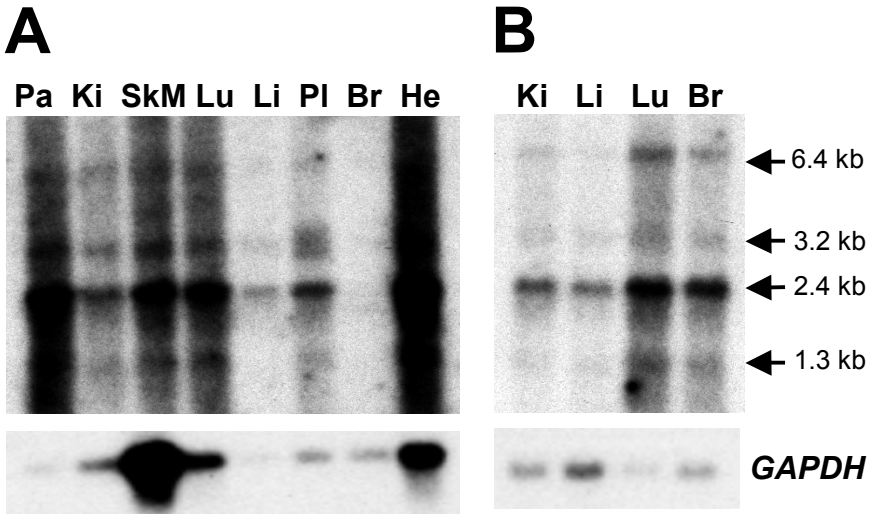
Figure 6: Patient 1 X chromosome breakpoint localisation
A: Schematic diagram of Patient 1 overlapping PAC clone RP4-596C15 indicating the location and orientation of the two genes in the region, with respect to the centromeric (cen) and telomeric (tel) ends of the clone. Restriction map of the breakpoint (BP) region is shown below (expanded insert). Horizontal dotted lines in **A** indicate expected control (C) restriction fragments that hybridise to the probe (black bar corresponding to nucleotides 41632-43225 of clone RP4-596C15) used for Southern blot hybridisation. **B:** Southern blot hybridisation using the probe described in **A**. Fragments of expected sizes 11186, 5897, and 3560 nucleotides, for *Pvu*II, *Hin*dIII, and *Hin*cII, respectively, corresponding to clone RP4-596C15 nucleotides 33695-44881(*Pvu*II), 37930-43827(*Hin*dIII), and 41326-44886 (*Hin*cII), are present in both control (C) and patient (P) DNA. Aberrant restriction fragments (red arrows) observed exclusively in patient (P) DNA, indicate the breakpoint region (highlighted in red in **A**).



Neither *GUCY2F* nor *NXT2* is disrupted by the breakpoint (see Figure 6A). Very little is known about *NXT2*; therefore, we investigated its expression pattern. Using an RT-PCR product corresponding to sequence from the three terminal exons of the *NXT2* gene as a probe for northern blot hybridisation, we detected transcripts in multiple tissues including foetal and adult brain (Figure 7). In addition to the predominant transcript observed at approximately 2.4 kb, we observed additional transcripts exhibiting the same pattern of expression, with approximate sizes of 6.5, 3.2, and 1.3 kb, suggesting the presence of either multiple *NXT2* splice variants or closely related family members. In light of its expression in the brain, *NXT2* remained a candidate MRX gene despite the fact that its published sequence is not directly disrupted by the breakpoint. However, RT-PCR analysis (data not shown) on

RNA from patient and control cell lines provided no evidence for altered *NXT2* expression levels as a result of the chromosome abnormality.

Figure 7: *NXT2* northern blot hybridisations
Northern blots probed with an *NXT2* cDNA probe corresponding to nucleotides 294-690 of Accession # NM_018698. Adult tissues (**A**) include (L-R) pancreas, kidney, skeletal muscle, lung, liver, placenta, brain, and heart. Foetal tissues (**B**) include (L-R) kidney, liver, lung, and brain. *GAPDH* served as a control for RNA loading.



3.1.2 Patient 2 with translocation t(X;15)(q13;p11)

Patient 2 is also a female with mental disability carrying a balanced X;autosome translocation. In this case, the mental retardation is more severe, and the patient has some dysmorphic features. After a normal pregnancy and spontaneous delivery at term, the patient had normal birth parameters: mass of 3600 grams (50^{th} percentile), length of 51 centimetres (75^{th} percentile), and head circumference of 34.5 centimetres (50^{th} percentile). There was no neonatal muscular hypotonia, and feeding was normal. At the age of four months, she was obese, and at 12 months, muscular hypotonia was observed. The patient was sitting freely at 12 months, crawling at 15 months, and standing at 14 months. She began walking at 20 months, and her first words were at 18 months. At the age of 4 years and 9 months, a detailed clinical analysis led to a questionable diagnosis of Prader-Willi syndrome. The patient weighed 29 kilograms (>95^{th} percentile), and was 109 centimetres tall (75^{th} percentile). Her head circumference was 52.5 centimetres (within the normal range). She presented with psychomotor retardation, hyperphagia, and obesity. Routine laboratory work indicated that she did not have diabetes. The proband had some distinct facial features including

hypotelorism and convergent strabismus (right to left). She had a high arched palate and flat occiput. External female genitalia were normal. At the age of 8 years, the patient no longer exhibited hyperphagia, obesity, or hypotonia, and the tentative diagnosis of Prader-Willi syndrome was excluded. However, the severe mental retardation and global developmental delay persisted. She was referred for cytogenetic testing, and a balanced translocation t(X;15)(q13;p11) (depicted in Figure 8A) was detected. The parents were not karyotyped.

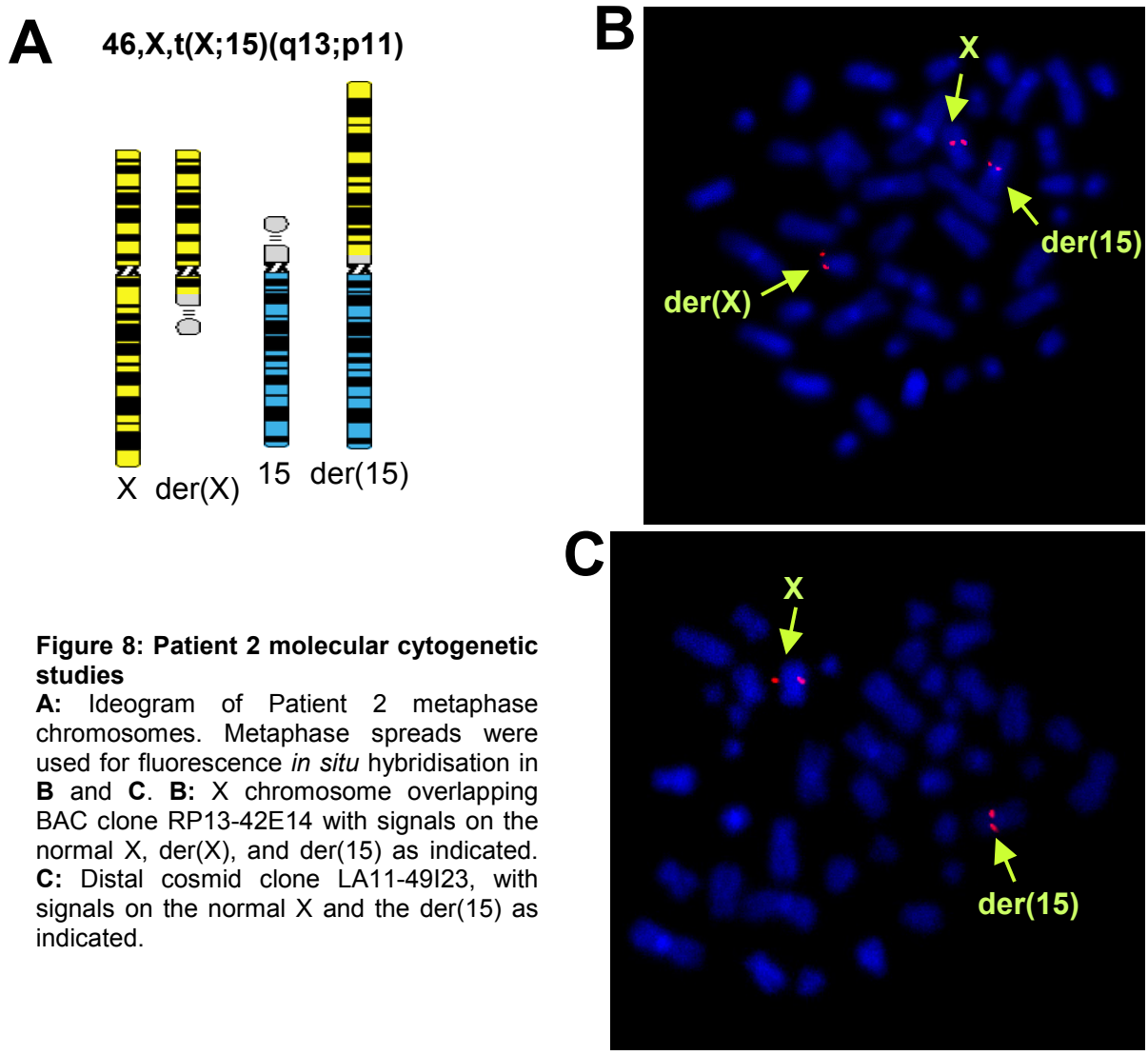


Figure 8: Patient 2 molecular cytogenetic studies
A: Ideogram of Patient 2 metaphase chromosomes. Metaphase spreads were used for fluorescence *in situ* hybridisation in **B** and **C**. **B:** X chromosome overlapping BAC clone RP13-42E14 with signals on the normal X, der(X), and der(15) as indicated. **C:** Distal cosmid clone LA11-49I23, with signals on the normal X and the der(15) as indicated.

Sequential hybridisation of Xq13-mapped genomic clones to patient metaphase chromosomes led to the identification of a breakpoint-spanning BAC clone, RP13-42E14

(Accession # AL359545), which hybridised to the normal X chromosome and to the two hybrid chromosomes (Figure 8B). Fully annotated sequence in the region of the X chromosome breakpoint-spanning clone was not available. However, based on available (fragmented) sequence of clone RP13-42E14, it was clear that the ATP-binding cassette transporter *ABCB7* lay in the vicinity of the breakpoint (for schematic diagram see Figure 9A). *ABCB7* is a ubiquitously expressed gene that has been implicated in X-linked

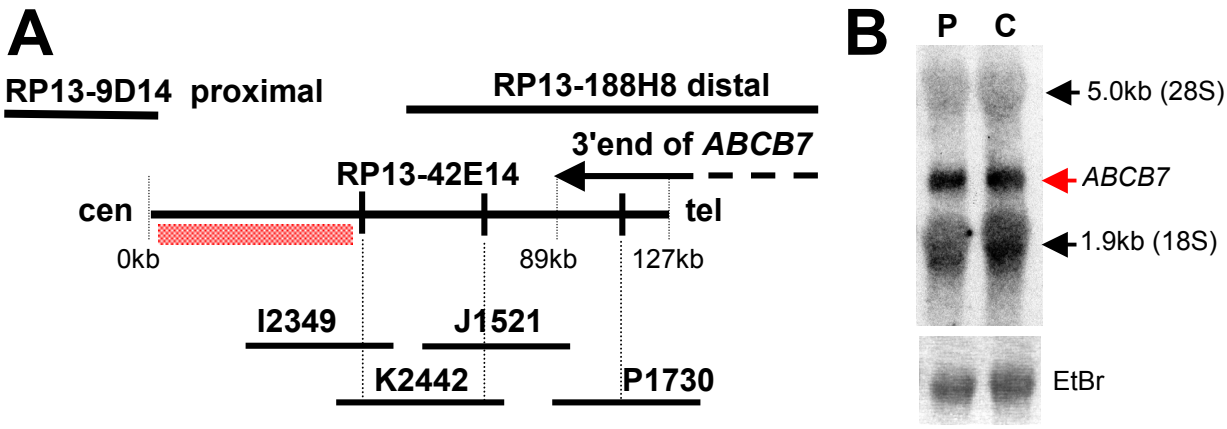


Figure 9: Analysis of *ABCB7* transcripts in Patient 2
A: Schematic diagram of Patient 2 overlapping clone RP13-42E14, with centromeric (cen) and telomeric (tel) ends indicated. The location and orientation of the *ABCB7* gene with respect to probes (black bars) used for cosmid library screening. Selected clones (indicated below), containing probe sequences as shown by dotted lines, were used for FISH, thereby localising the breakpoint to the centromeric end of clone RP13-42E14. The breakpoint region is highlighted in red. **B:** Northern blot hybridisation with RNA from patient (P) and control (C) lymphoblastoid cell lines using a cDNA probe corresponding to nucleotides 62-652 of the *ABCB7* mRNA sequence (Accession # NM_004299). Ethidium bromide (EtBr) staining of 18S ribosomal RNA served as a control for RNA loading. The expected 2.3 kb *ABCB7* transcript is indicated with a red arrow; black arrows indicate non-specific hybridisation of probe to ribosomal RNAs.

sideroblastic anaemia with ataxia (Allikmets et al. 1999; Bekri et al. 2000), which is sometimes associated with MR (Bekri et al. 2000), and is clearly a good candidate gene for XLMR. In order to determine if this gene is affected by the breakpoint in the translocation patient, two strategies were employed. First, available sequence from the distal end of clone RP13-42E14 was used to design primers for amplification of DNA probes for screening an X chromosome-specific cosmid library (Los Alamos, available from RZPD). Positive clones (depicted in Figure 9A, lower part) were FISH mapped (Figure 8C), and all lay distal to the breakpoint, suggesting that the breakpoint lies on the centromeric side of the overlapping BAC clone, beyond the 3' end of the *ABCB7* gene. Secondly, the 2.4 kb *ABCB7* transcript was analysed by northern blot hybridisation, using RNA from lymphoblastoid cell lines

(Figure 9B). No difference between patient and control samples was observed. Although it is possible that the chromosome abnormality contributes to the phenotype in this patient, for example, by exerting effects on distant genes, our data indicate that the phenotype is not a result of obvious alterations in the local *ABCB7* transcript.

3.1.3 Patient 3 with translocation t(X;7)(p11.3;q11.21)

Patient 3 is a young girl presenting with severe developmental delay and carrying a *de novo* balanced chromosome rearrangement t(X;7)(p11.3;q11.21) (depicted in Figure 10A). She is the daughter of healthy unrelated parents, and there is no family history of mental disability or related illness. She has two healthy siblings. After normal pregnancy and birth, she had an Apgar score of 10/10, birth weight of 3130 grams (25^{th} percentile), length of 51 centimeters (50^{th} percentile), and head circumference of 33.5 centimeters (25^{th} percentile). At one week she had metatarsus adductus and at 3 months an umbilical hernia was observed.

At the age of two years, the patient did not yet walk or speak, and she exhibited clearly delayed psychomotor development. At the age of three years and nine months, she was diagnosed with speech delay, within the framework of general developmental delay. Additionally, occasional convulsions predominantly localised to the shoulders and upper back were reported, leading to the diagnosis of myoclonic epilepsy. However, grand mal seizures had never been observed. At five years and six months, her physical development was in the normal range, and a thorough assessment uncovered no dysmorphic features. Her skin was normal, suggesting no genetic mosaicism. She continued, however, to exhibit severe mental and psychomotor delay. She could speak but was unable to create sentences of more than four words. On clinical examination at the age of six years and four months, her height was 118 centimeters (50^{th} percentile), her weight was 25 kilograms (90^{th} percentile), and she was friendly and cooperative. She had some difficulty standing on one leg and walking along a line, suggesting minor ataxia. Routine laboratory tests, including thyroid hormone and ammonium levels, were normal. Screening for inborn errors of metabolism gave no indication of amino acidopathies or organic acidurias. Magnetic resonance imaging (MRI) did not indicate any brain abnormalities, but electroencephalograhy (EEG) exhibited a spike-wave pattern. The patient continued to suffer from myoclonic seizures and was treated with valproate and ethosuximide. One year later EEG was normal.

This patient had been screened for abnormalities in both fragile sites FRAXA and FRAXE, and disease-causing triplet repeat expansions could be ruled out. Molecular characterisation of the X chromosome breakpoint region by fluorescence *in situ* hybridisation of genomic clones led to the identification of two clones, RP13-83I24 (Accession # AL590283) and RP13-479F17 (Accession # AL590223), with partially overlapping

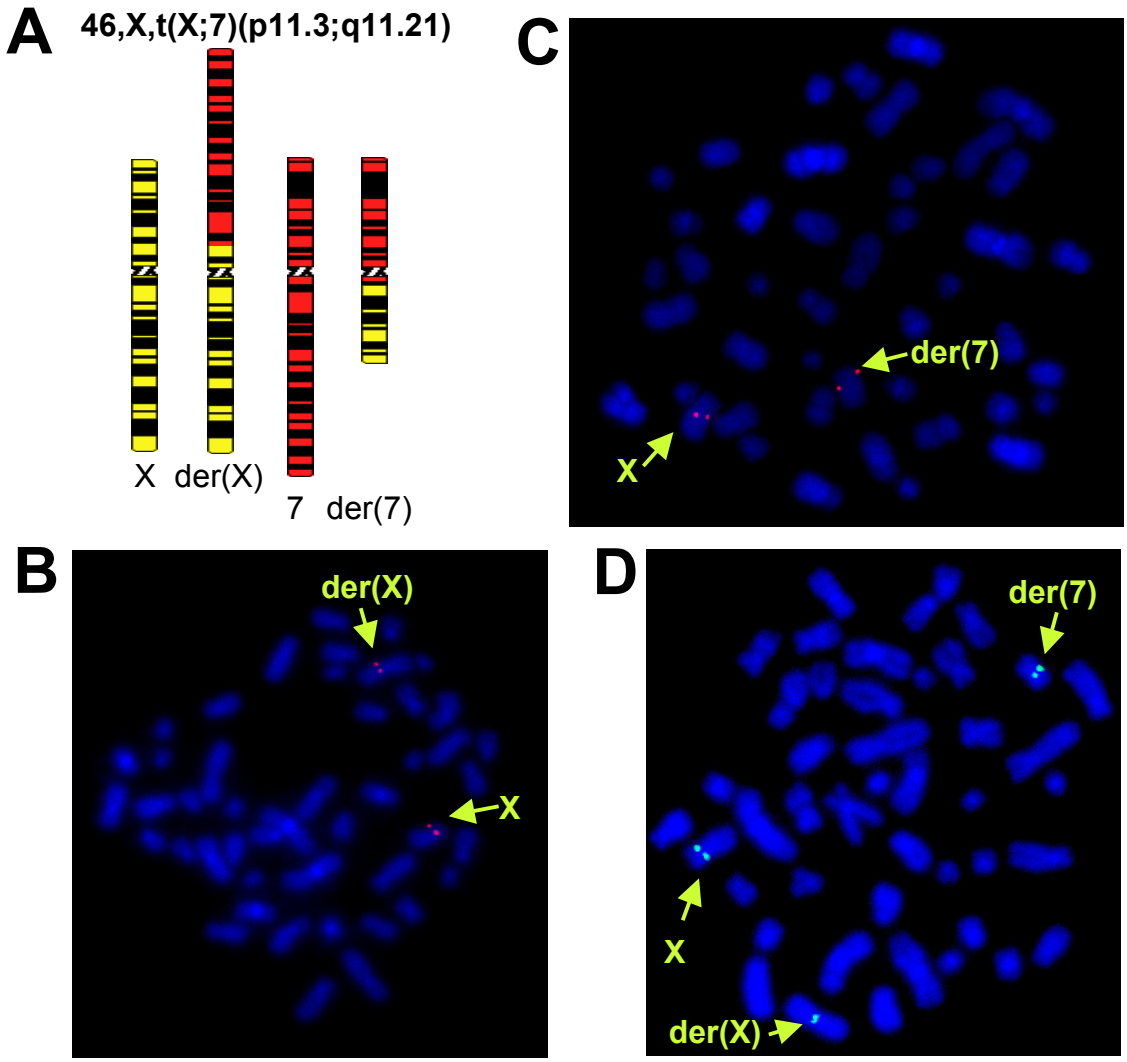
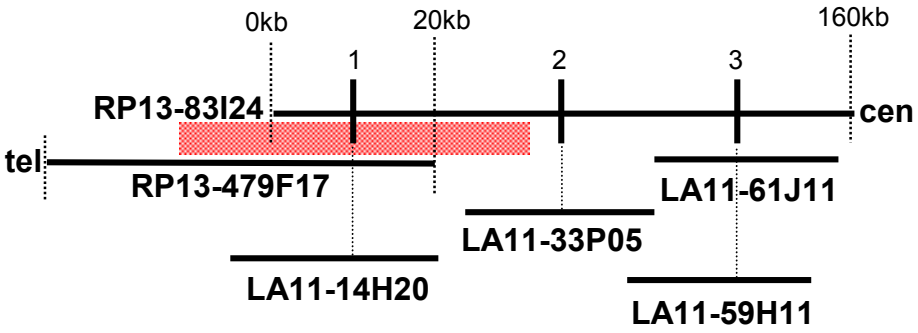


Figure 10: Patient 3 molecular cytogenetic studies
A: Ideogram of Patient 3 metaphase chromosomes. Metaphase spreads were used for fluorescence *in situ* hybridisations shown in **B**, **C**, and **D**. **B:** X chromosome proximal BAC clone RP13-83I24, which gives signals on the normal X and the der(X) as indicated. **C:** X chromosome distal BAC clone RP13-479F17, which gives signals on the normal X and the der(7), as indicated. **D:** Cosmid clone LA11-14H20, which gives a signal on the normal X and split signals on the der(X) and the der(7) as indicated.

sequence, that gave proximal and distal hybridisation signals, respectively (Figures 10B and C), suggesting that the breakpoint might lie in the region of sequence overlap. As in the case of Patient 2, hybridisation of PCR products to a chromosome X-specific cosmid library (Los Alamos) led to the identification of several cosmid clones (depicted in Figure 11), one of which (LA11-14H20) hybridised to both derivative chromosomes (Figure 10D), confirming that this region is indeed disrupted in the patient. Fully-annotated sequence indicated the presence of the 37 kb-spanning zinc finger 41 gene (*ZNF41*) in the breakpoint region (see schematic diagram in Figure 12A).

Figure 11: Patient 3 X chromosome breakpoint region and cosmid library screening results
Schematic diagram of Patient 3 candidate breakpoint region (red shaded bar) indicating the location of the probes (black bars) used for cosmid library screening, with respect to proximal and distal BAC clones RP13-83I24 and RP13-479F17. Centromeric (cen) and telomeric (tel) ends of these clones are indicated. Probes 1,2, and 3 correspond respectively to nucleotides 18500-19465, 61390-62375, and 129191-130130 of clone RP13-83I24 (GI # 14031123, now updated to Accession # AL590283). Selected cosmid clones, containing probe sequences as indicated by dotted lines, were used for FISH.



Zinc finger genes are involved in transcriptional regulation, and various members of this superfamily play an established role in both disease and development (Ladomery and Dellaire 2002a). Moreover, *ZNF41* is subject to X-inactivation (Rosati et al. 1999). As this patient exhibits skewed X-inactivation of the normal X chromosome (personal communication), any alterations in the *ZNF41* copy on the translocated chromosome would likely have functional consequences. Analysis of *ZNF41* at the RNA level confirmed a striking difference between patient and control (Figure 12B): by RT-PCR, full-length *ZNF41* transcripts were absent in the patient cell line but readily detected in a control cell line, suggesting that there is no functional ZNF41 in the patient. This result, in conjunction with

the severe phenotype, made Patient 3 an obvious candidate for further studies, which are discussed in detail in section 3.2.

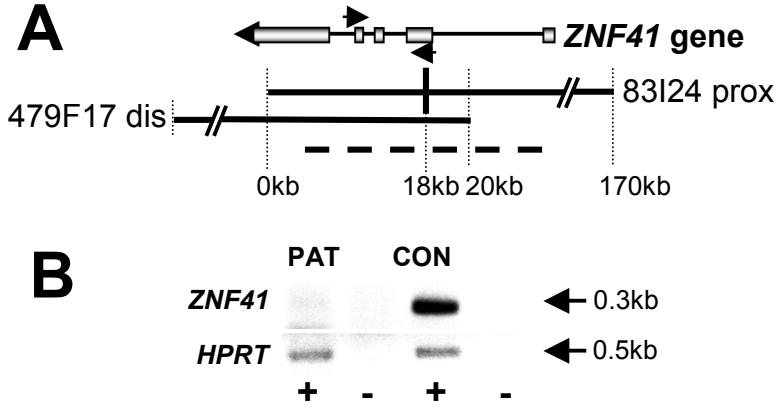


Figure 12: Analysis of *ZNF41* transcripts in Patient 3
A: Schematic diagram indicating the region of sequence overlap between Patient 3 X chromosome proximal (prox) and distal (dis) BAC clones (RP13-83I24 and RP13-479F17, respectively). Numbers below correspond to sequence of BAC RP13-83I24). Angled lines indicate that both clones extend beyond the region depicted. The horizontal dotted line indicates the approximate location of the overlapping cosmid clone, based on the presence of confirmed cosmid sequence, at approximately kb 18 of BAC clone RP13-83I24 (indicated by the black bar). Above these clones, the respective location of the *ZNF41* mRNA sequence is shown. Arrows indicate location of primers used for RT-PCR amplification of the *ZNF41* mRNA (**B**) from patient (PAT) and control (CON) lymphoblastoid cell lines. + and – indicate with reverse transcriptase or mock reaction, respectively.

3.2 Further investigations into the cause of MR in Patient 3

Preliminary studies on three patients with MR and associated balanced translocations led to the identification of a single promising candidate XLMR gene, *ZNF41*, which is the focus of section 3.2.

3.2.1 The X-chromosomal *ZNF41* gene is disrupted near the 5' end

In order to localise the chromosome X breakpoint in Patient 3 relative to the coding sequence of *ZNF41*, the region was mapped by restriction endonuclease digestion and Southern blot hybridisation using probes from within the region of the overlapping cosmid. Aberrant restriction fragments in patient DNA were detected in several independent digests (Figure 13) using a probe corresponding to nucleotides 9529-10112 of clone RP13-83I24. In particular, the aberrant fragment of approximately 1 kb in the patient *Eco*RI digest enabled us to locate the breakpoint precisely. It lies within intron 2 of *ZNF41*, thereby truncating the predicted protein (Accession # NP009061) just after amino acid 24 (see Figure 13A). Cloning of the breakpoint fragment observed in the *Eco*RI digest (Figure 13B) and subsequent sequence analysis (Figure 14A) indicated the precise location of the breakpoints on each

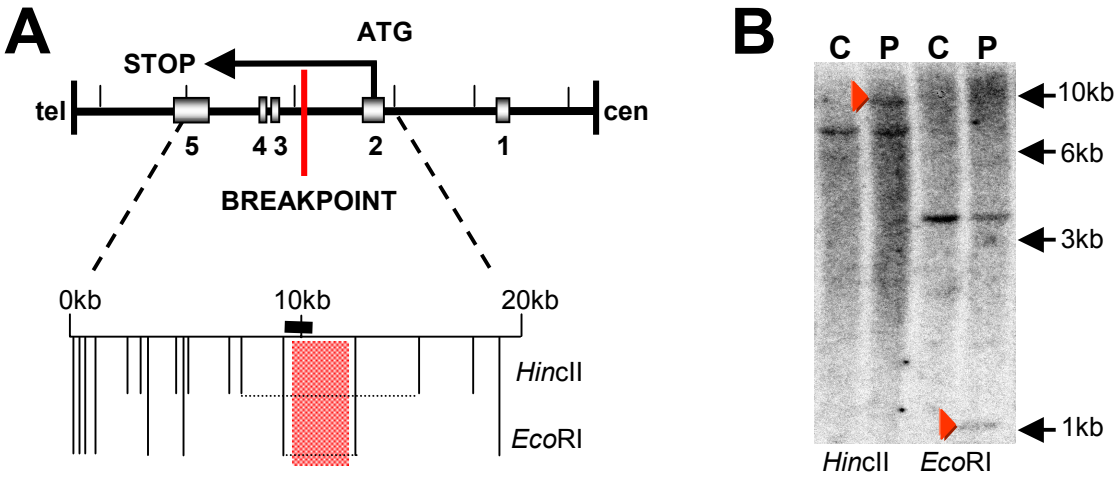


Figure 13: Patient 3 X chromosome breakpoint localisation
A: Schematic diagram depicting the breakpoint location (highlighted with red shading), restriction sites, and location of the probe used for Southern blot hybridisation in **B**, relative to the *ZNF41* gene (exonic sequence indicated by shaded boxes). Orientation relative to centromere (cen) and telomere (tel) is indicated. **B:** Southern blot of patient (P) and control (C) genomic DNA probed with sequence corresponding to nucleotides 9529-10112 of clone RP11-83I24. Aberrant fragments specific to patient DNA (red arrows) in *Hinc*II and *Eco*RI digests correspond to sizes of approximately 10 kb and 1 kb, respectively.

chromosome. Both are disrupted within LINE1 repetitive elements (L1ME on the X, and L1 on 7), and the three bases, AAG, indicated in large type in Figure 14A, are present on both chromosomes; it is therefore not clear whether these three bases come from the X chromosome or from chromosome 7, or from a combination of both. The chromosome 7 breakpoint lies within BAC clone RP11-196D18 (Accession # AC069285), which was confirmed, by FISH, to span the breakpoint (Figure 14B). The chromosome 7 breakpoint does not disrupt any known genes, and based on both UCSC "Golden Path" (April 2003 freeze) and NIX analysis, there are no ESTs bridging the breakpoint region. The closest documented spliced ESTs or mRNA sequences (Accession #s AW661872 and BC041841) map approximately 96 kb and 132 kb to either side of the breakpoint. Further studies, therefore, focussed on the X chromosome where an obvious candidate had been identified.

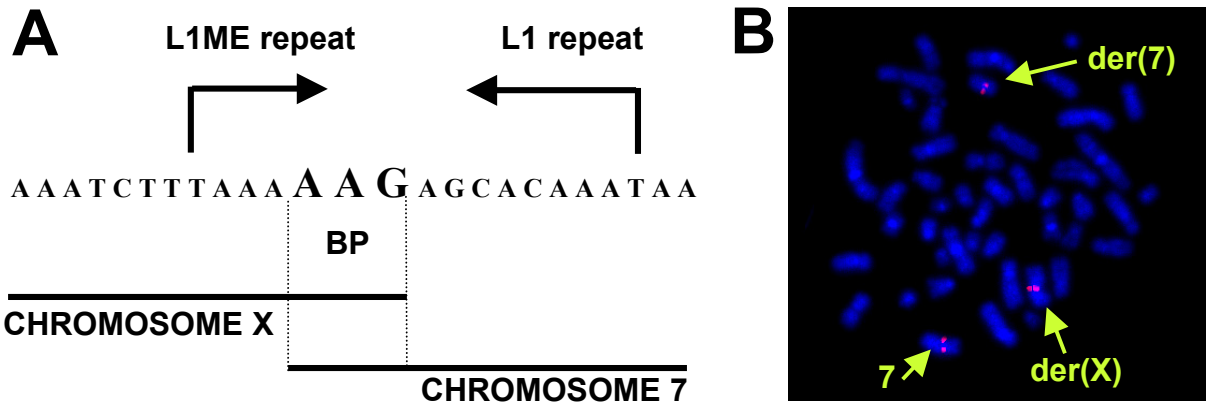


Figure 14: Patient 3 breakpoint cloning and chromosome 7 breakpoint confirmation by FISH
A: Sequence of the breakpoint fragment, indicating (in large type) the three nucleotides common to both chromosome X and chromosome 7, the neighbouring X chromosome sequence (corresponding to BAC clone RP11-83I24), and the neighbouring chromosome 7 sequence (corresponding to BAC clone RP11-196D18). **B:** Fluorescence *in situ* hybridisation of this chromosome 7 breakpoint-spanning BAC clone, with a signal on the normal 7 and split signals on the der(X) and der(7), as indicated.

3.2.2 Full-length *ZNF41* transcripts are absent in the patient cell line

Further analysis of *ZNF41* transcripts in the patient lymphoblastoid cell line (Figure 15) confirmed the aberrant transcription observed earlier. Transcripts 5' to the breakpoint could be amplified readily, whereas transcripts 3' to the breakpoint were not detected even at low levels in the patient (Figure 15B). As expected, hypoxanthine phosphoribosyl transferase (*HPRT*), a housekeeping gene also located on the X chromosome, was present at normal levels in the patient. Interestingly, very low levels of *ZNF41* were also detected with high cycle numbers in reactions with breakpoint-spanning primers, suggesting that at least some

transcripts are transcribed from the normal X chromosome in the patient, and that they can be detected with specific amplification conditions.

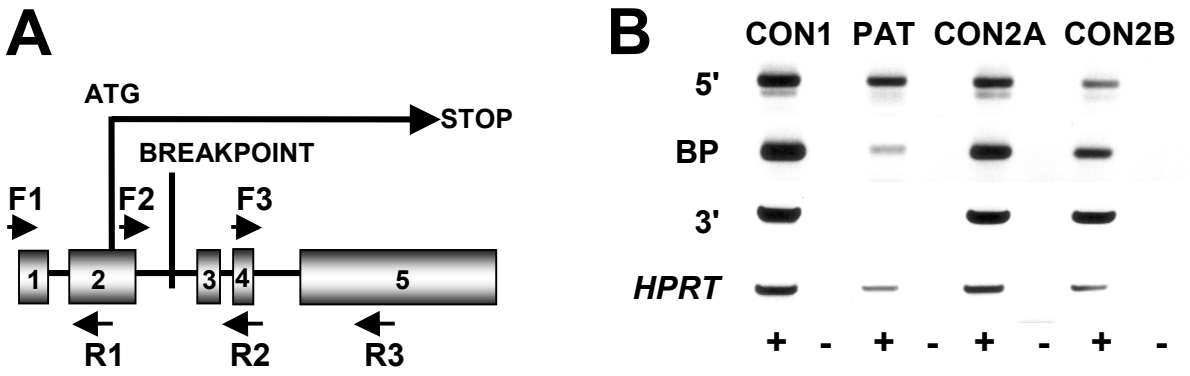


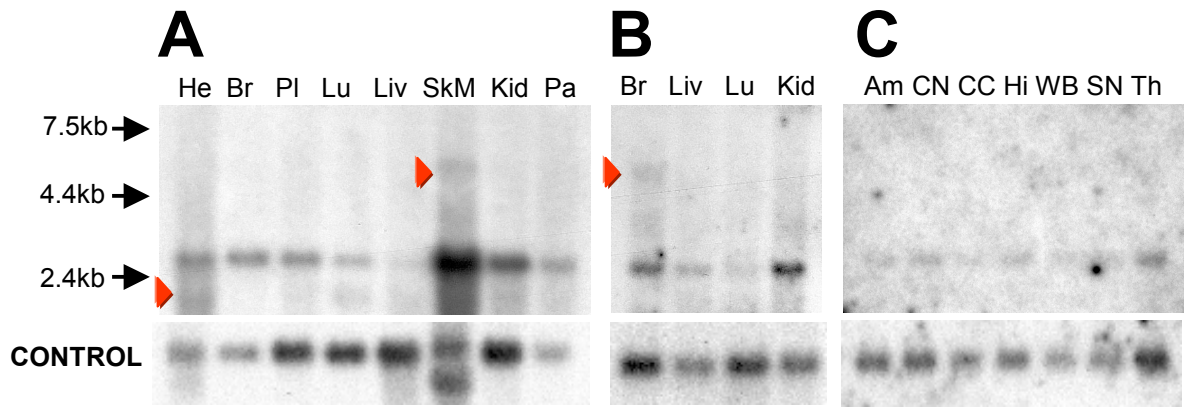
Figure 15: Further RT-PCR analysis of *ZNF41* transcripts in Patient 3
A: Schematic diagram (not to scale) indicating the exon-intron structure of *ZNF41* relative to the breakpoint location. Arrows indicate approximate location of primers used for RT-PCR analysis. **B:** RT-PCR analyses using *ZNF41* primers 5' to the breakpoint (F1 and R1), spanning the breakpoint (F2 and R2), and 3' to the breakpoint (F3 and R3). *ZNF41* products described above correspond respectively to nucleotides 41 to 344, 407 to 702, and 621 to 1099 of Accession # NM_007130. Amplification of hypoxanthine phosphoribosyltransferase (*HPRT*) served as a control. Female control (CON1), patient (PAT), and male control (CON2; A and B refer to different RT reactions from the same RNA sample) are from lymphoblastoid cell lines; + and – refer to reaction with reverse transcriptase or mock reaction, respectively.

3.2.3 *ZNF41* is expressed in multiple tissues including brain

Little is known about the *ZNF41* gene; we therefore investigated its expression pattern in both adult and foetal tissues. As expected, a predominant transcript of approximately 3 kb was observed. The transcript is present in all adult tissues tested, with relatively low expression in lung and liver (Figure 16A). *ZNF41* is also transcribed in all foetal tissues tested, and shows a similar pattern (Figure 16B). Hybridisation of the same *ZNF41* 3' cDNA probe to a blot with RNA from several adult brain tissues also confirmed ubiquitous expression in these brain structures (Figure 16C). Interestingly, a larger transcript of approximately 6 kb was detected at significant levels specifically in foetal brain. A transcript of the same size was observed at low levels in adult skeletal muscle; however, due to the high levels of RNA present in this sample, no conclusions can be drawn about the tissue specificity of this transcript in the adult. There was also a transcript of approximately 2 kb observed exclusively in adult heart. Expression of both the expected 3 kb transcripts and the

uncharacterised 6 kb transcript in foetal brain support a role for the *ZNF41* gene in early cognitive development.

Figure 16: *ZNF41* northern blot hybridisations
Analysis of *ZNF41* transcripts by Northern blot hybridisations using a cDNA probe corresponding to nucleotides 621-1099 of Accession # NM_007130. **A:** Adult tissues (L-R) heart, brain, placenta, lung, liver, skeletal muscle, kidney, pancreas. **B:** Foetal tissues (L-R) brain, lung, liver, kidney. **C:** Adult brain structures (L-R) amygdala, caudate nucleus, corpus callosum, hippocampus, whole brain, substantia nigra, thalamus. Red arrows indicate a novel tissue-specific transcripts. *Actin* (**A** and **C**) or *GAPDH* (**B**) served as a control for RNA loading.



3.2.4 Identification of two novel *ZNF41* mutations in a panel of patients with XLMR

Given the absence of this potentially important zinc finger protein in the translocation patient investigated, we considered *ZNF41* to be a candidate disease gene in families presenting with XLMR. We screened the complete *ZNF41* coding region, as well as neighbouring intronic sequence, for disease-causing mutations in 210 unrelated patients with suspected XLMR. The predominant *ZNF41* coding sequence (henceforth also referred to as *ZNF41.1* or variant 41.1) is present in two documented mRNA sequences (Accession #s NM_007130 and NM_153380, respectively), which differ only in their 5'UTRs. Additional documented transcripts include transcripts *ZNF41.2- ZNF41.8* (Rosati et al. 1999). Transcript variants 41.2, 41.3, and 41.6 include additional coding sequence that is not present in 41.1, which was also included in our mutation screening.

We confirmed the presence of two previously published single nucleotide polymorphisms (SNPs), both present in approximately 1% of the patients screened, and identified four novel nucleotide exchanges. For a complete description of nucleotide

exchanges identified, see Table 2. Of these four exchanges, one (780T>G, Accession # NM_007130 resulting in I125R, Accession # NP_009061) was confirmed to be a polymorphism; it was present in 3 out of 210 patients, and in 8 out of 340 control individuals. The 1864A>G silent mutation found in one patient was not investigated in controls because the presence of a G at the corresponding nucleotide position in *ZNF81*, a close homologue of *ZNF41*, suggests that this alteration is also a polymorphism. More importantly, we identified one proline to leucine amino acid exchange (P111L, Accession # NP_009061) and one intronic splice site mutation (479–42A>C, Accession # NM_007130) that were not found in over 400 control X chromosomes screened. These mutations are examined further in the next sections.

Table 17: Summary of sequence alterations identified in *ZNF41* DNA in an MRX patient panel

Exon	Nucleotide Exchange (NM_007130)	Amino Acid Exchange (NP_009061)	Number of Patients with Exchange	Number of Control Chromosomes	Previously Identified?
3	**479–42A>C ***	**N/A**	**1**	**0/405**	**No**
5a	**738C>T** 780T>G	**P111L** I125R	**1** 3	**0/401** 8/401	**No** No
5b	1192G>A	Q262Q	3	NA	Yes
5c	1351T>G	D315E	3	NA	Yes
5e	1864A>G	Q486Q	1	ND	No

NA = not applicable, ND = not determined
Potentially disease-causing mutations highlighted in bold type
*at position –2 with respect to the corresponding exon in variants 41.3 and 41.9

3.2.5 Studies on family P13 with proline to leucine mutation

In MRX family P13 (depicted in Figure 17), a C to T exchange at position 738 of *ZNF41.1* results in a proline to leucine exchange at amino acid 111 (P111L, Accession # NP_009061). Given the unique structure of proline, and the absence of a similar exchange in over 400 control X chromosomes screened, we consider this alteration to be of potential significance. DNA from non-nuclear family members is not available for further studies, but sequence analysis of DNA from both the mother of the affected individual (II:1), and the only brother of the patient (III:1), who is also mentally disabled, confirmed the presence of the

same mis-sense mutation (see Figure 17). Both affected boys (III:1 and III:2) in this family presented with non-specific mental retardation. The index patient (III:2) was walking at the age of 17 months, and at 5 years and 8 months, he was functioning at an intellectual level of 3 years (by Brunet-Lezine's test). He exhibited a language retardation, avoided social contact, and was hyperactive. He had no dysmorphic features or additional neurological abnormalities. His older brother (III:1) had a similar history. The parents are unrelated, and two brothers on the maternal side are similarly affected (see Figure 17). There is some data suggesting that one female cousin on the maternal side (III:4, indicated with a *) may also be affected in some way, but no clinical data is available.

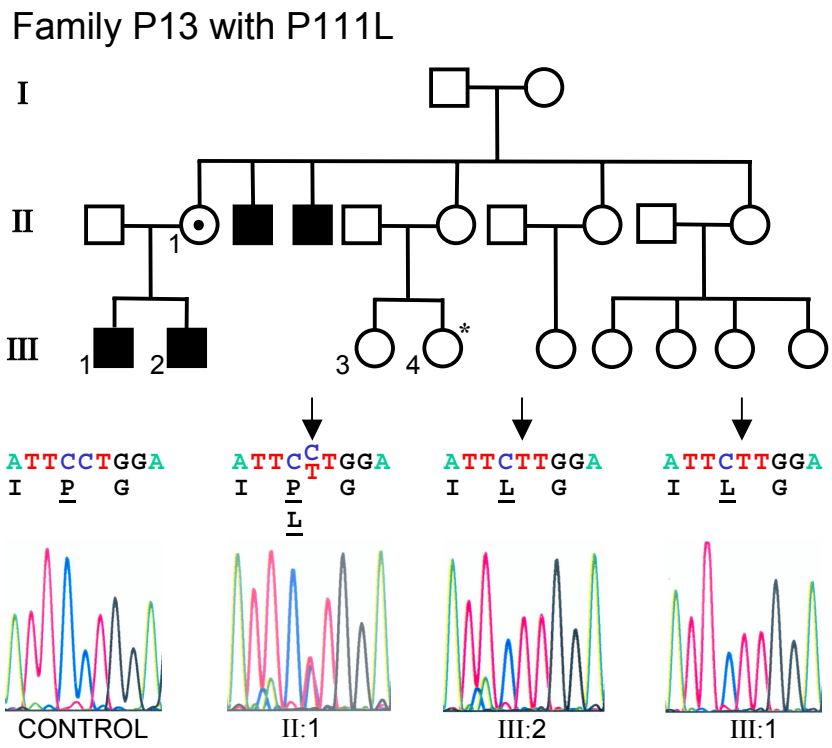


Figure 17: MRX family P13 with P111L mutation
Pedigree for family P13, with sequence corresponding to the proline to leucine mutation: (L-R) unrelated control, mother (II:1), index patient (III:2), and brother of the index patient (III:1). For the potentially affected female cousin (individual III-4, indicated with a *), no clinical data is available. Affected nucleotides are indicated by black arrows.

3.2.6 Studies on documented *ZNF41* transcripts in family P42 with splice-site mutation

In MRX family P42 (depicted in Figure 18), an adenine to cytidine intronic mutation (479-42A>C, Accession # NM_007130) (Figure 18, lower panel) that results in aberrant

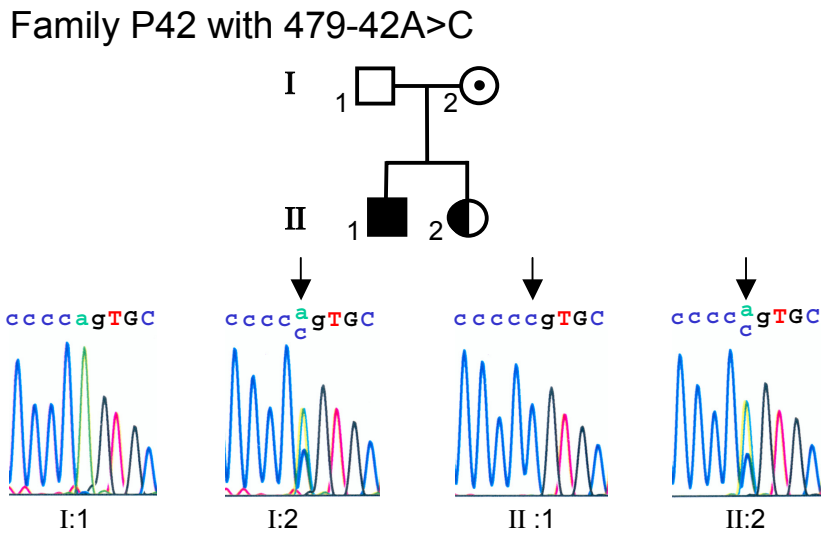


Figure 18: MRX family P42 with intronic splice-site mutation
Pedigree for family P42, with sequence chromatograms indicating the splice site mutation in affected individuals: (L-R) father (I:1), mother (I:2), index patient (II:1), and mildly affected sister (II:2). Upper case letters indicate coding sequence; affected nucleotides are indicated with black arrows.

splicing of *ZNF41* was detected. More precisely, two documented *ZNF41* transcripts, variants 41.3 and 41.6 (corresponding to Accession #s AJ010018 and AJ010021, respectively; see also Rosati et al. 1999), contain 5' extensions of exon 3 (referred to as exon 2 by Rosati et al. 1999, but now updated in NCBI), that make use of the consensus splice site that is mutated in family P42. Our initial investigations on the resulting splicing defects in the patient cell line focussed primarily on transcript variant 41.3 (depicted schematically in Figure 19A, together with variants 41.1 and 41.2), as variant 41.6 is expressed at low levels (see Rosati et al. 1999) and could not be readily detected in lymphoblastoid cell lines. As shown in Figure 19B, transcript variant *ZNF41.3* can be specifically amplified with primers F2 and RX (depicted in Figure 19A) in control cell lines but is absent in the patient cell line. Moreover, exclusively in

the patient cell line, larger mis-spliced transcripts are observed (Figure 19B). In variant *ZNF41.3* the intron 2 acceptor splice-site follows predicted amino acid 20, and subsequent skipping of exon 3 would result in the presence of a premature stop codon in exon 5 (Figure 19A bottom). Like other RNAs with premature termination codons, such a transcript would likely be targeted for nonsense-mediated mRNA decay (for review, see Hentze and Kulozik 1999). This is supported by the fact that we did not detect *ZNF41.3* transcripts lacking exon 3 in the patient cell line using external primers (see Figure 19C). With primers F2 and R2 in exons 2 and 4, respectively (depicted in Figure 19A), variants 41.1 and 41.2 could be detected in both control and patient cell lines (Figure 19C). Variant 41.3, which is weakly expressed, was also observed in the control, whereas it could not be amplified at all in the patient cell line (see Figure 19C).

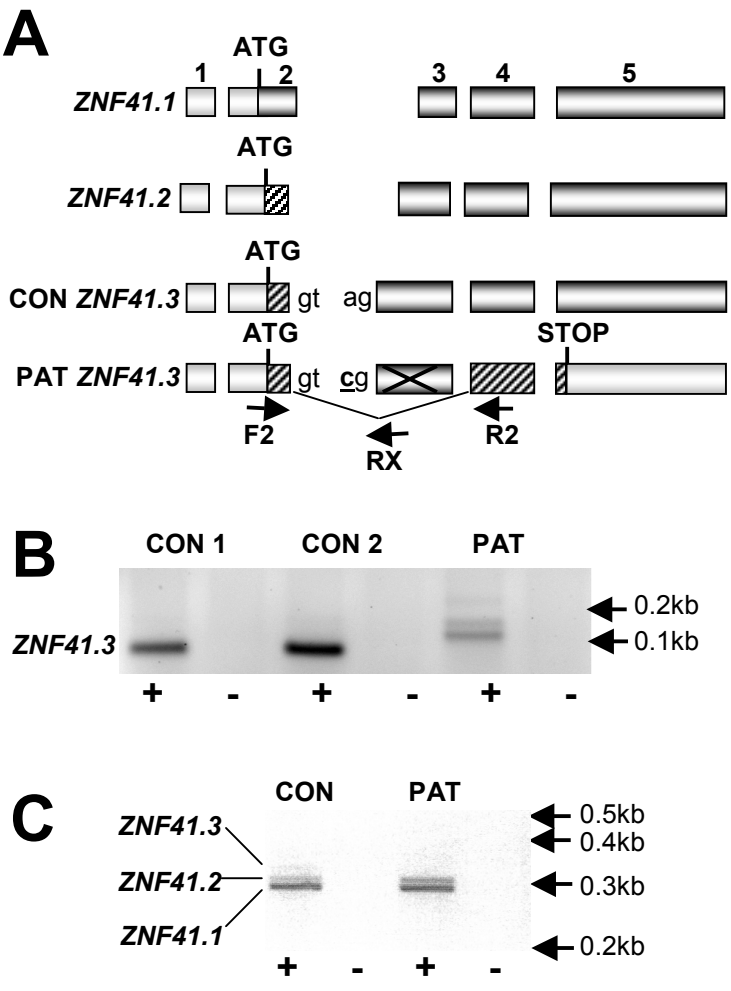


Figure 19: Analysis of *ZNF41.3* transcripts in family P42 index patient

A: Schematic diagram depicting *ZNF41* transcript variants 41.1, 41.2, and 41.3. The predominant open reading frame is indicated by dark shading; alternate 5' coding sequence present in variants 41.2 and 41.3 is indicated with diagonal line pattern; light shading indicates predicted UTR. In transcript 41.3, shown for both control (CON) and patient (PAT), the a to c intronic splice site mutation is highlighted in bold type and underlined. Arrows indicate the approximate locations of primers F2, R2, and RX used for RT-PCR analyses. **B:** RT-PCR specific for *ZNF41.3* in male control (CON1), female control (CON2), and patient (PAT) lymphoblastoid cell lines, indicating the absence of normal 41.3 transcripts in the patient. **C:** RT-PCR on lymphoblastoid cell lines with external primers F2 and R2 designed to simultaneously amplify the three splice variants depicted in **A**. CON indicates female control, and PAT indicates patient. + and – refer to reaction with reverse transcriptase or mock reaction, respectively.

Like the proline to leucine amino acid exchange, this mutation was not found in other patients screened, nor was it detected in controls. The patient has a diagnosis of mild mental retardation. He was born at term (by Cesarean section), with a birth weight of 3000 grams (10^{th}-25^{th} percentile) and length of 51 centimetres (50^{th} percentile). He walked at twelve to thirteen months and reached early milestones within the normal time frame; however, he exhibited a severe language delay. He first made two-word associations at three years of age and was first speaking in simple phrases at the age of four and a half years. At the age of 8 years, he was 135 centimetres tall (90^{th} percentile) and had a head circumference of 53 centimetres (75^{th} percentile). He had no additional dysmorphic or neurological symptoms, and screening for FRAXA was negative. At the age of 10 years and 3 months, the patient was unable to read and had essentially no acquisition of mathematics skills. He had mood disorders with both hyperactivity and aggressiveness, and he required special education. At this time his verbal IQ score was 83, and his performance IQ score was 59 (by Wechsler's scale). Cerebral MRI was normal. The younger sister of this patient was recently diagnosed with a borderline form of mental retardation. She also had school difficulties associated with impulsive behaviour and hyperactivity. Sequencing confirmed that she is heterozygous for the splice-site mutation (see Figure 18). The father does not carry the mutation; it was inherited from the mother, who is also a carrier (Figure 18). We have no clinical data suggesting that the mother is affected, but this question was never directly addressed.

3.2.7 Identification of a novel *ZNF41* splice variant also affected in family P42

Interestingly, we have also identified a novel splice variant of *ZNF41* (designated 41.9 and depicted in Figure 20A with respect to related *ZNF41* transcripts 41.3 and 41.4), which lacks exon 2, and therefore, like variant *ZNF41.4* (Rosati et al. 1999), has a predicted open reading frame (ORF) that does not contain the transcriptional repressor domain encoded by exon 3. Unlike variant *ZNF41.4*, however, this variant contains the 5' extension of exon 3 that is present in variant *ZNF41.3*, and like variant *ZNF41.3*, *ZNF41.9* is absent in the patient cell line (Figure 20B).

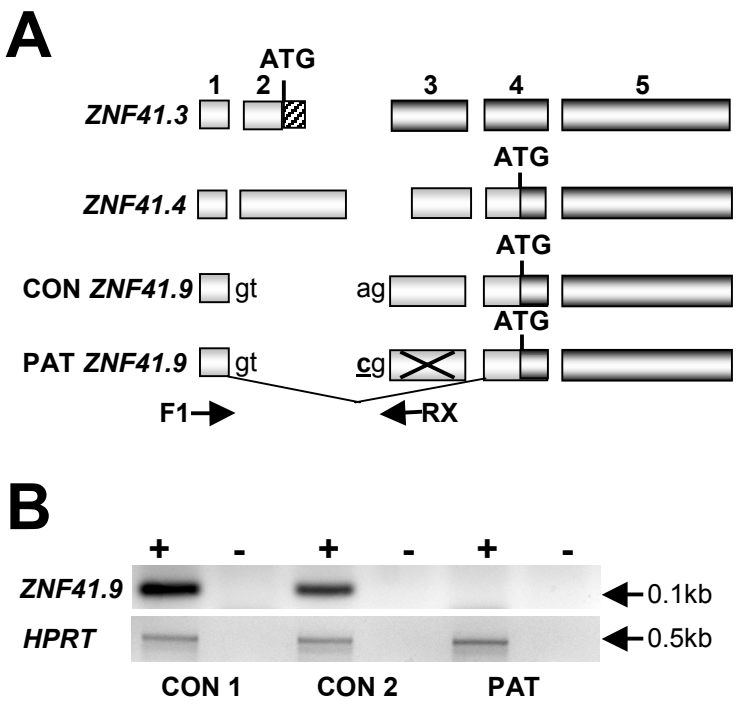


Figure 20: Analysis of *ZNF41.9* transcripts in family P42 index patient

A: Schematic diagram depicting the novel transcript variant *ZNF41.9*, with respect to the related transcripts 41.3 and 41.4. In transcript 41.9, shown for both control (CON) and patient (PAT), the a to c intronic splice site mutation is highlighted in bold type and underlined. Predicted open reading frames are indicated with dark shading; light shading indicates predicted UTR sequence. Black arrows indicate approximate locations of primers F1 and RX used for amplification of variant 41.9, shown below. **B:** RT-PCR for the novel transcript variant *ZNF41.9* in male control (CON 1), female control (CON2), and patient (PAT) lymphoblastoid cell lines, indicating its absence in the patient cell line. Amplification of hypoxanthine phosphoribosyltransferase (*HPRT*) served as a control; + and – refer to reaction with reverse transcriptase or mock reaction, respectively.

3.2.8 Expression analysis of *ZNF41* variants affected by the splice-site mutation in family P42

Simultaneous PCR amplification of multiple *ZNF41* transcripts affirmed that the various transcripts exhibit differential expression, and specifically that variant 41.3, relative to variants 41.1 and 41.2, is present at low levels in lymphoblastoid cell lines (see Figure 19C, CON lane). In order to investigate the expression pattern of both variants 41.3 and 41.9, which were previously confirmed by RT-PCR to be affected by the splice-site mutation in family P42, it was necessary to use a probe specific for these variants to avoid cross-hybridisation with other *ZNF41* transcript variants that are expressed at higher levels. We used an oligo corresponding to the 5'extension of exon 3 that is specific for variants 41.3, 41.6 and 41.9 for northern blot hybridisation (for a schematic diagram indicating the location of this probe relative to coding sequence of all *ZNF41* variants, see Figure 21A). We detected a ubiquitously expressed transcript (see Figure 21B). Interestingly, this transcript has a size of approximately 2kb, suggesting that the predominant variant harbouring the 5'-extension of

exon 3 is either the smaller variant 41.9, in which exon 2 is missing, or an as yet undiscovered variant related to 41.3 and 41.6, for which the 3'end of the gene differs from the 41.1 described by Rosati et al. (1999). Clearly, any *ZNF41* splice variants detected by this oligo are likely affected by the splice-site mutation, but further studies are required to understand how changes specifically within these variants can affect the function of the normal ZNF41 protein.

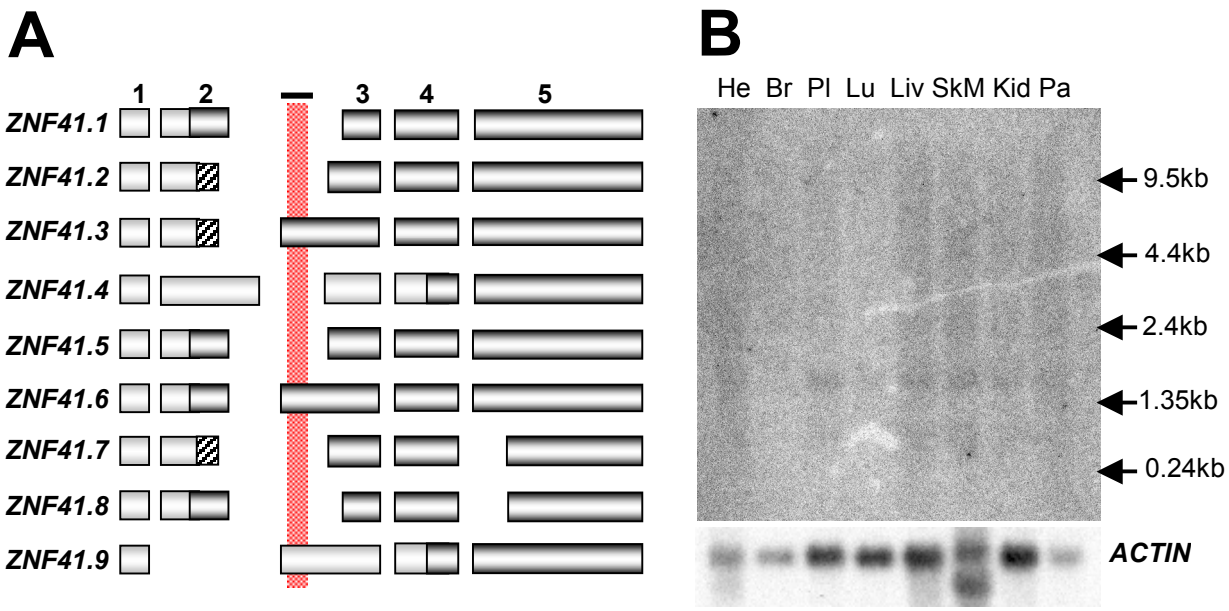


Figure 21: Analysis of transcripts affected by the splice-site mutation in MRX family P42
A: Schematic diagram depicting the 8 *ZNF41* transcript variants described by Rosati et al. (1999), as well as the novel *ZNF41.9*. The predominant ORF is indicated by dark shading; alternate 5' coding sequence unique to transcripts 41.2, 41.3, and 41.7 is indicated with a diagonal line pattern; light shading indicates predicted 5'UTR. The red shading highlights the position of the oligo probe (black bar) used for expression analysis of the three transcripts that are affected by the splice-site mutation in family P42. **B:** Hybridisation using the probe described in **A** (5'-CTG TTC AAG GCA GCA TGG TC-3'), specific for splice variants containing this 5'-extension of exon 3, to a northern blot with adult tissues as follows: (L-R) heart, brain, placenta, lung, liver, skeletal muscle, kidney, pancreas. Approximate sizes are indicated to the right. Actin served as a control for RNA loading.

3.3 Clinical and molecular studies in patients with disease-associated autosome translocations: the search for autosomal MR candidate genes

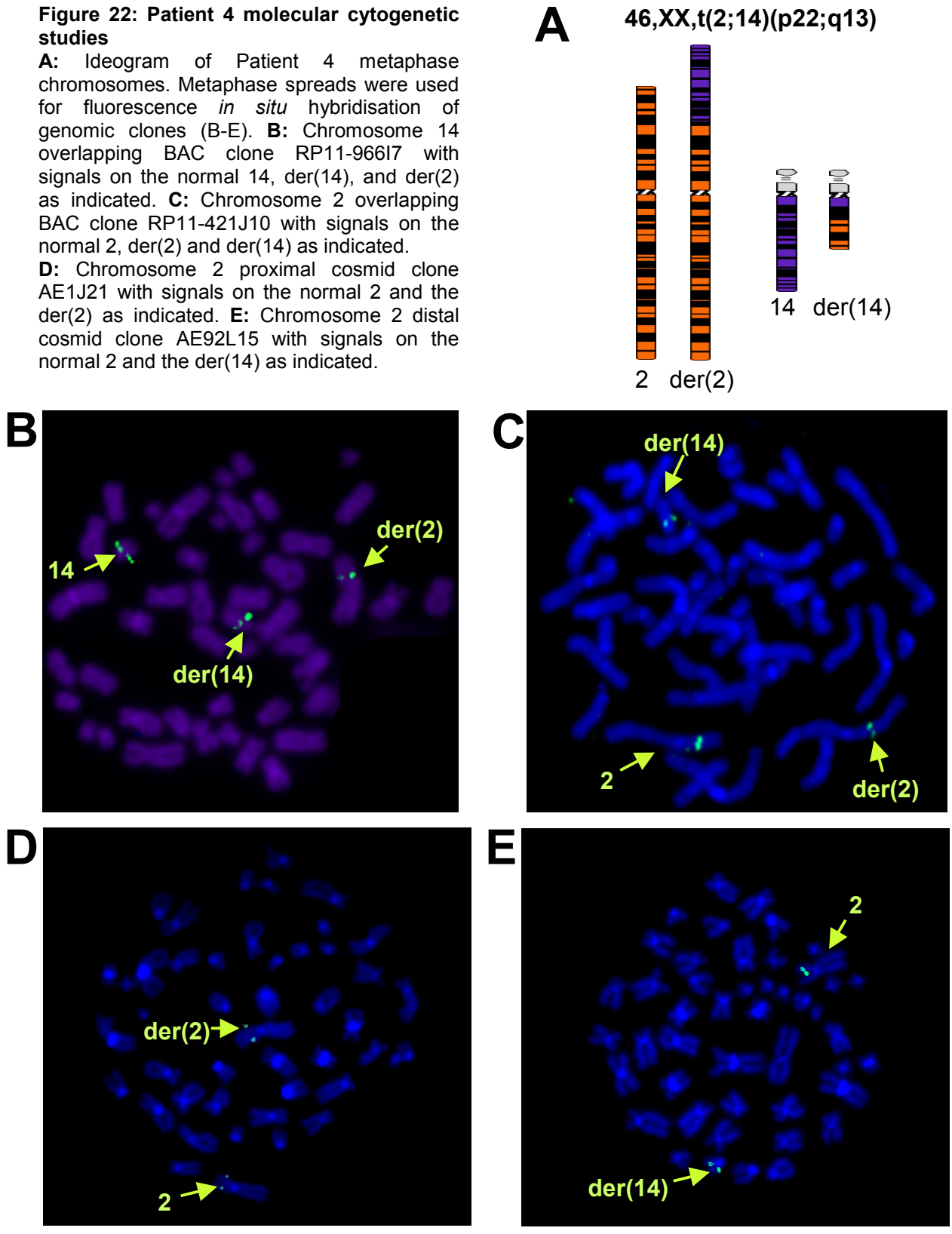
Two autosome translocation carriers with associated mental disabilities were also included in this study. Based on the severity of the phenotype in both cases, a significant genetic defect was considered probable. As for X chromosome translocation cases, after confirmation of the translocation and breakpoint localisation by FISH, initial molecular and *in silico* studies aimed to identify specific candidate disease genes, and in one of the two cases investigated, such a gene was identified; further studies on this candidate are described in section 3.4.

3.3.1 Patient 4 with translocation t(2;14)(p22;q13)

Patient 4 is a young female diagnosed with a severe and complex early-onset neurological disorder. Her birth was without complications (Apgar score 10/10) following a normal pregnancy. All birth parameters were in the normal range: mass 3350 grams (50^{th} percentile) length 50 centimetres (50^{th} percentile), and head circumference 36 centimetres (75^{th} percentile). Her disorder was first noted at the age of two weeks, when during feeding, severe muscle rigidity was observed. Subsequently, she became increasingly agitated, and sweating episodes were noted. Upon thorough clinical examination at three weeks of age, EEG was normal but brain sonography indicated a significant asymmetrical enlargement of the lateral ventricles, in particular in the region of the hindbrain and parietal lobes. The third ventricle was also somewhat enlarged, and complete agenesis of the corpus callosum was ascertained. Routine blood and urine tests for metabolic defects indicated no abnormalities.

At six months of age, she was unable to lift her head, and microcephaly was detected. Routine karyotyping indicated that she carries a *de novo* balanced translocation t(2;14)(p22;q13), depicted in Figure 22A. Fluorescence *in situ* hybridisation of genomic clones from both chromosome 2 and 14 led to identification of breakpoint-spanning clones from each chromosome. Clone RP11-966I7 (Accession # AL049777), from chromosome 14, hybridised to the normal chromosome 14, and to both derivative chromosomes (Figure 22B), indicating that it spans the chromosome 14 breakpoint; likewise, the chromosome 2 clone RP11-421J10 (Accession # AC009236) also gave split signals on both derivative chromosomes (Figure 22C), indicating that it spans the chromosome 2 breakpoint.

Figure 22: Patient 4 molecular cytogenetic studies
A: Ideogram of Patient 4 metaphase chromosomes. Metaphase spreads were used for fluorescence *in situ* hybridisation of genomic clones (B-E). **B:** Chromosome 14 overlapping BAC clone RP11-966I7 with signals on the normal 14, der(14), and der(2) as indicated. **C:** Chromosome 2 overlapping BAC clone RP11-421J10 with signals on the normal 2, der(2) and der(14) as indicated.
D: Chromosome 2 proximal cosmid clone AE1J21 with signals on the normal 2 and the der(2) as indicated. **E:** Chromosome 2 distal cosmid clone AE92L15 with signals on the normal 2 and the der(14) as indicated.



In silico analysis of fully-annotated sequence for both the chromosome 2 and 14 breakpoint-spanning clones indicated the presence of a promising candidate gene, the well-characterised forkhead transcription factor *FOXG1B* gene (known as *brain factor 1* in the mouse) near the centre of the chromosome 14 breakpoint-spanning BAC clone. This gene is almost exclusively expressed in the brain, and it is known in mice for its critical involvement in the development of the telencephalon (Tao and Lai 1992; Xuan et al. 1995); it is therefore a logical candidate for the structural changes observed in the brain in our patient, and indicated our starting point for the next phase of the breakpoint search. Restriction mapping of *FOXG1B* and neighbouring sequence (Figure 23A), and subsequent Southern hybridisations led to the detection of patient-specific aberrant fragments using a probe corresponding to nucleotides 101623-102882 of clone RP11-966I7, indicating that the breakpoint lies between nucleotides 100089 and 103386, just distal to the *FOXGIB* candidate gene, which spans nucleotides 92721-95305 (see Figure 23A).

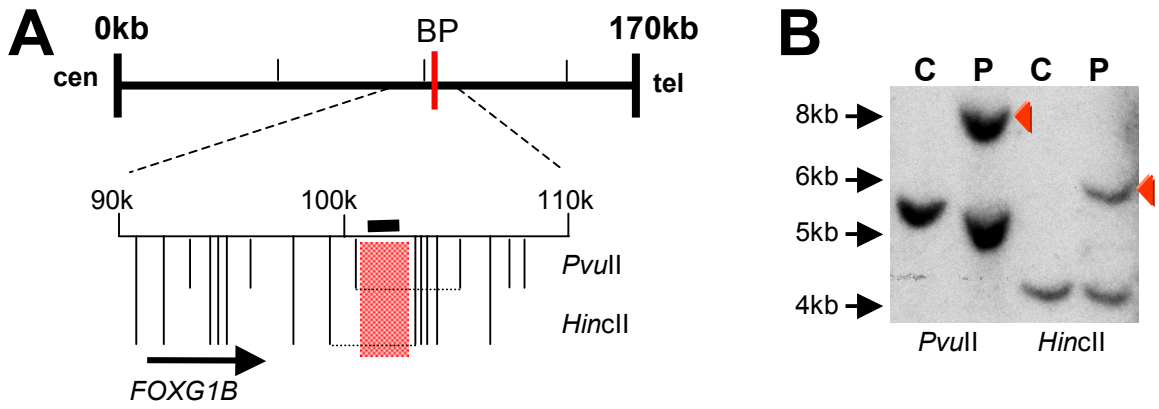


Figure 23: Patient 4 chromosome 14 breakpoint localisation
A: Schematic diagram of Patient 4 chromosome 14 overlapping BAC clone RP11-966I7 indicating the location and orientation of the known *FOXG1B* gene, relative to the centromeric (cen) and telomeric (tel) ends of the clone. The expanded insert depicts a restriction map of the breakpoint (BP) region. Horizontal dotted lines indicate expected control (C) restriction fragments that hybridise to the probe (black bar corresponding to nucleotides 101623-102882 of clone RP11-966I7) used for Southern blot hybridisation shown in **B**. Expected fragment sizes are 5299 and 4126 base pairs for *Pvu*II, and *Hinc*II, respectively, and they correspond to nucleotides 99260-103386 (*Hinc*II) and 100089-105388 (*Pvu*II) of clone 966I7. Aberrant restriction fragments of approximately 8kb and 6kb (red arrows) are observed exclusively in patient (P) samples in **B**, and indicate the approximate breakpoint (BP) location (red bars in **A**).

Clearly, the known *FOXG1B* coding sequence is not disrupted; it is also unlikely that the promoter region is affected, as the breakpoint lies 3' to the gene. However, the chromosome 14 breakpoint disrupts an uncharacterised transcript defined by Unigene cluster Hs.92556.

This cluster is composed of 17 EST sequences, 16 of which are derived from brain cDNA libraries. Together these sequences comprise three exons that map to nucleotides 98437-98496 (exon 1), 98999-99073 (exon 2), and 117520-118809 (exon 3) of clone RP11-966I7. The predicted open reading frame (ORF) is 170 amino acids; it spans exons 1 and 2, and is thus disrupted by the breakpoint. The termination codon lies within exon 3. For a schematic diagram of the predicted ORF and 3'UTR of this novel gene with respect to the breakpoint on chromosome 14, see Figure 24A.

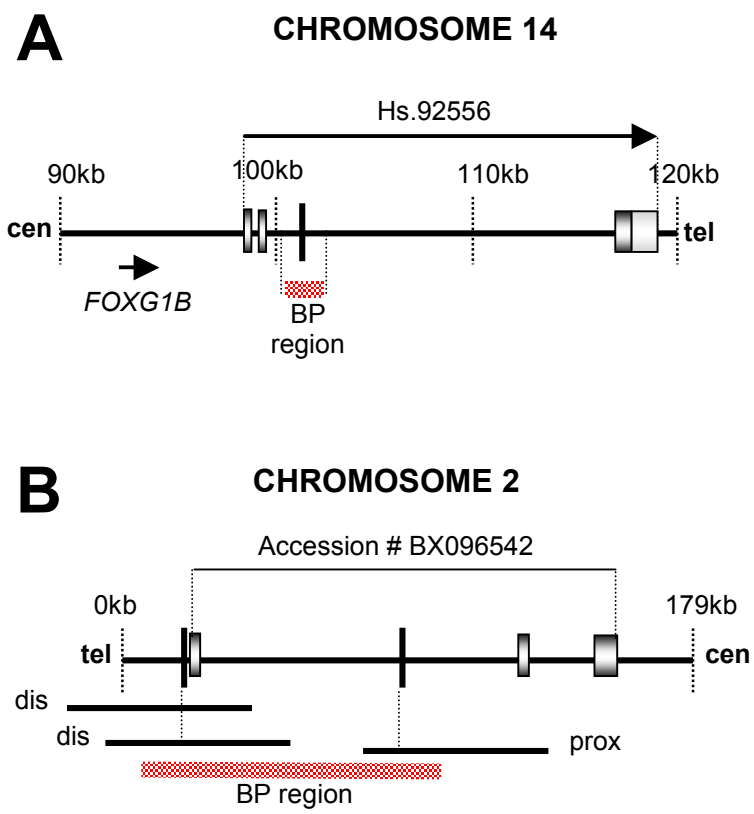


Figure 24: Summary of disrupted transcripts in Patient 4
A: Schematic diagram depicting the Patient 4 breakpoint region within the chromosome 14 breakpoint-spanning clone RP11-966I7. The black bar indicates the position of the probe used for identification of the breakpoint (BP) region (highlighted in red) by Southern hybridisation; shaded boxes indicate genomic location of known exon sequence from the breakpoint-spanning EST cluster. Dark shading represents predicted coding sequences; light shading indicates predicted 3'UTR. **B:** Schematic diagram depicting the Patient 4 chromosome 2 overlapping BAC clone RP11-421J10, and respective location of the intron-spanning EST that maps to the breakpoint region (exonic sequences indicated by shaded boxes). Black bars indicate location of probes within clone RP11-421J10 used for screening of a chromosome 2-specific cosmid library; approximate positions of proximal (prox) and distal (dis) FISH-mapped cosmid clones, and corresponding candidate breakpoint (BP) region are indicated below. In both **A** and **B**, centromeric (cen) and telomeric (tel) ends of clones are indicated.

The results of the chromosome 2 analyses are also complex. The availability of a chromosome 2-specific cosmid library enabled us to identify genomic clones for FISH-mapping. Using PCR products corresponding to base pairs 18035-19069 and 88322-89311 of clone RP11-421J10, positive clones from the chromosome 2-specific cosmid library (Livermore LLNL, available from RZPD) were selected. Clone AE1J21, which is positive for

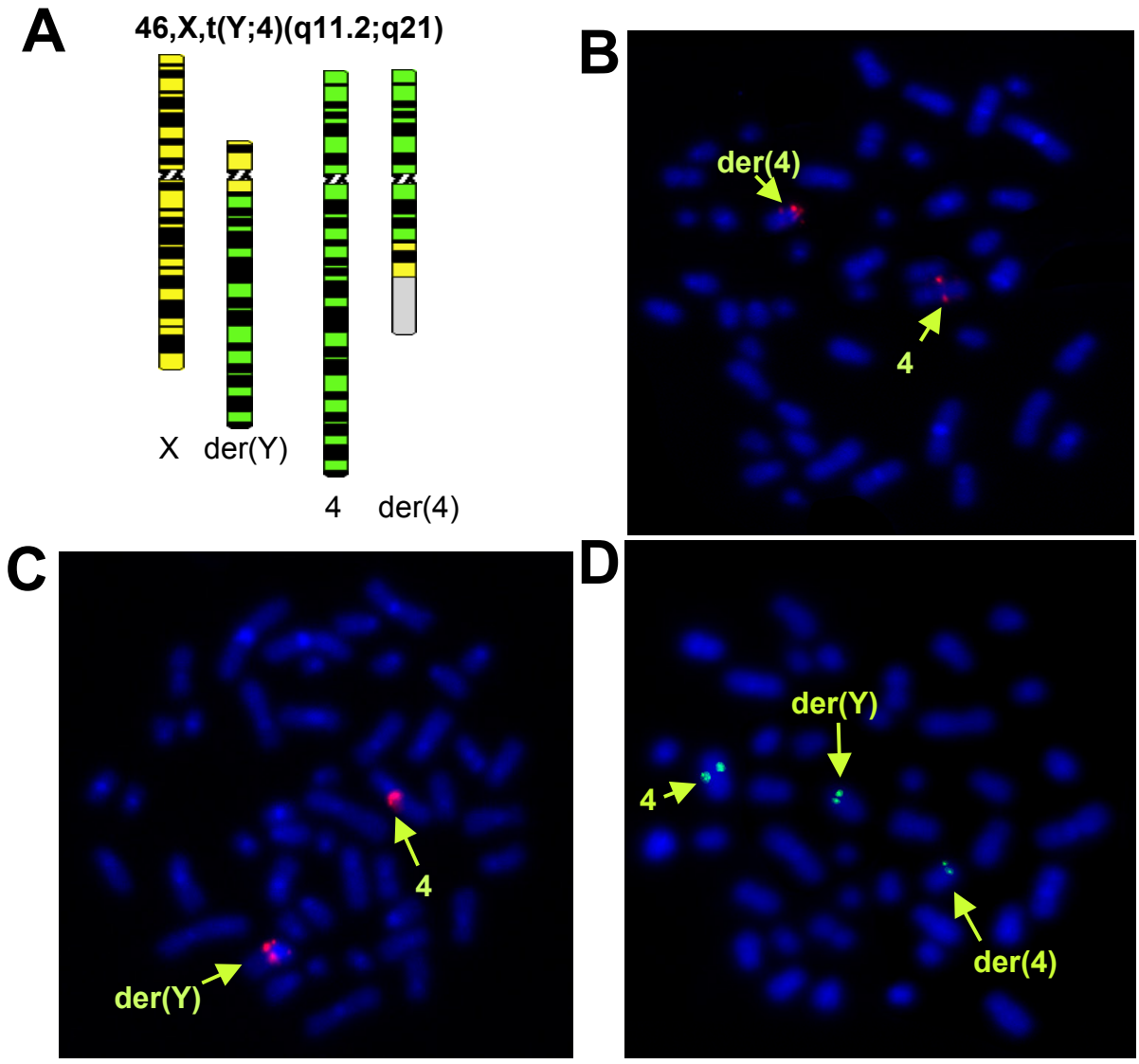
RP11-421J10 nucleotides 88322-89311 but not 18035-19069, hybridised to the normal 2 and the der(2) (Figure 22D), whereas clone AE92L15, which is positive for 421J10 nucleotides 18035-19069 but not 88322-89311, hybridised to the normal 2 and the der(14) (Figure 22E). Taken together, these results suggest that the breakpoint lies towards the middle of the overlapping BAC clone RP11-421J10, very likely between kilobase 18 and 88, or within a few kilobases to either side. *In silico* analysis of this region confirmed that at least one known EST (Accession # BX096452, derived from testis) spans this region. By sequence alignment, it maps to nucleotides 19050-19157, 127855-128144, and 149842-150166 of clone RP11-421J10 (depicted in Figure 24B), and is therefore likely disrupted by the breakpoint. Additionally, EST clusters map to both ends of this clone. The implications of these results are discussed in section 4.1.

3.3.2 Patient 5 with translocation t(Y;4)(q11.2;q21)

Patient 5 is a young male patient who was diagnosed with a severe neurodegenerative disorder associated with seizures that resist treatment. He was born at term after a normal pregnancy, with birth weight (4200 grams, 90th percentile) in the normal range. His early development was also normal: he could sit at 7 months, and he walked freely and also spoke a few words at 12 months. At thirteen months he presented with atonic seizures and ataxia. He remained interactive (social smile, fixation, and pursuit). EEG at this time showed diffuse asynchronous slow spike-wave complexes. Over the next few weeks, his interaction deteriorated, and he developed marked hypotonia and athetotic upper limb movements. He lost fine motor coordination, and he could no longer sit or walk unaided. At this time he also developed tonic seizures and partial complex seizures. Various therapies, including sodium valproate, lamotrigine, benzodiazepines, phenobarbitone, phenytoin, felbamate and carbamazepine, had little effect on seizures. He continued to lose motor coordination skills and became progressively less interactive. From the age of two, he had several episodes of convulsive and non-convulsive status epilepticus. Steroid therapy had some effect on non-convulsive status epilepticus, but from age 3, it was quasi-persistent despite therapy. He received a 5-day course of oral ketamine (2 mg/kg/day), and exhibited marked improvement after two days. He became more alert, and seizures ceased completely for one week. Throughout the ketamine course, EEG showed a background of low amplitude, which was not the case before treatment. There was an excess of diffuse irregular theta activity.

Discharges were rare but with very high amplitude (up to 1300 microvolts). After one week, seizures returned, with markedly diminished intensity, and the patient remained alert. His social smile returned, he recognised his mother, and he regained some coordination. Since then, however, he has made no neurodevelopmental progress. In contrast, he had episodes of

Figure 25: Patient 5 molecular cytogenetic studies
A: Ideogram of Patient 5 metaphase chromosomes. Metaphase spreads were used for fluorescence *in situ* hybridisation of (**B**) chromosome 4 proximal BAC clone RP11-375K17 which gives signals on the normal 4 and the der(4), (**C**) chromosome 4 distal BAC clone RP11-778J15**,** which gives signals on the normal 4 and the der(Y), and (**D**) overlapping cosmid clone D0537, which gives signals on the normal 4, and split signals on both derivative chromosomes, as indicated.



increased seizure activity. Two additional ketamine courses also had positive effects, but results were not as striking as the initial response described above. The third course brought about relatively little change. Topiramate and levetiracetam have since been administered without significant results. Following treatment, EEG still showed background activity, with occasional runs of fast activity (approximately 20/sec) in the anterior regions, and there were minimal discharges. The florid discharges that prevailed prior to the first ketamine course have never recurred, and the very high amplitude of discharges was observed only during the first two courses of ketamine. Further deterioration since this time has not been observed at the clinical level, but no further testing has been done. The picture is consistent with severe Lennox-Gastaut encephalopathy.

Routine karyotyping indicated the presence of a *de novo* balanced translocation t(Y;4)(q11.2;q21) (depicted in Figure 25A). FISH mapping of genomic clones from chromosome 4q21 to patient metaphase chromosomes led to breakpoint localisation between clones RP11-375K17 (Accession # AC139716) and RP11-778J15 (Accession # AC104827), which lie proximal and distal to the breakpoint, respectively (Figures 25B and C). The Y chromosome, which presumably harbours no genes of essential importance for basic brain development, was not considered; further studies in Patient 5, therefore, centred on this region of chromosome 4. Public genomic sequence for the region between the proximal and distal FISH-mapped BAC clones was not available; however, sequence data (available from Celera^{TM} in 2001) suggested that the region between the two clones was approximately 160 kb, and that the well-characterised c-Jun terminal kinase 3 gene, *JNK3,* spanned this region. Screening of a chromosome 4-specific cosmid library (LANL, available from RZPD) and subsequent FISH mapping of cosmid clones led to the identification of an overlapping clone 37D05 (Figure 25D), which confirmed that the breakpoint disrupts *JNK3* (for schematic diagram see Figure 26).

JNK3 is a member of the mitogen-activated protein (MAP) kinase family of signal transduction molecules, and it is expressed predominantly in the central nervous system. It plays an established role in both neuronal apoptosis and differentiation and therefore was an obvious candidate for further studies, which are described in the next section.

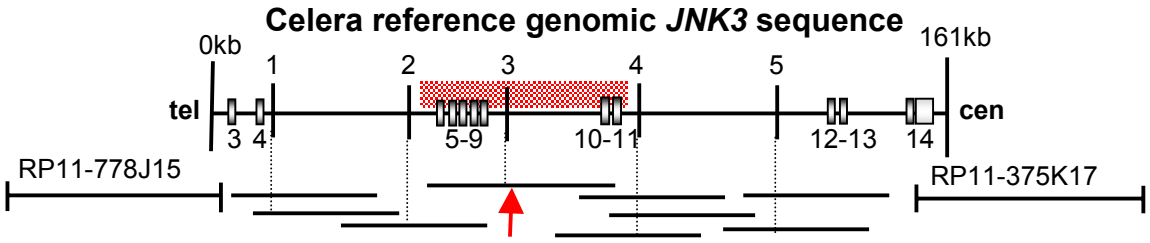


Figure 26: *JNK3* coding sequence with respect to FISH mapped genomic clones
Schematic diagram indicating the location of the probes used for screening the chromosome 4-specific cosmid library, with respect to the proximal (RP11-375K17) and distal (RP11-778J15) FISH mapped clones, and the Celera reference genomic sequence of *JNK3*, which spans the region in between. Centromeric (cen) and telomeric (tel) ends are indicated. Shaded boxes depict approximate location of *JNK3* exonic sequences, numbered according to Yoshida et al. (2001); black bars indicate location of probes 1-5, at respective nucleotides 14463-15560, 49930-50981, 74952-76052, 107399-108493, and 122858-123731, of Celera reference sequence used for mapping. Approximate locations of cosmids selected for FISH are depicted schematically below. The overlapping cosmid (positive for probe 3, but not 2 or 4) is indicated with a red arrow; the corresponding candidate breakpoint region is highlighted with red shading.

3.4 Further studies on JNK3

Preliminary studies on two patients with MR and associated balanced autosome translocations led to the identification of a promising candidate gene for MR, *JNK3*, which is the focus of section 3.4.

3.4.1 The breakpoint disrupts JNK3α isoforms after amino acid 267

FISH mapping of genomic clones to Patient 5 metaphase chromosomes indicated that *JNK3* is disrupted by the breakpoint; precise localisation of the breakpoint required restriction mapping and Southern blot hybridisations. Figure 27A depicts a restriction map of the chromosome 4 breakpoint region with respect to the genomic sequence of *JNK3*. Aberrant bands (Figure 27B, red arrows) in patient DNA from four restriction digests were identified by hybridisation of a probe corresponding to nucleotides 77714-79011 of the *JNK3* gene sequence used for mapping (Figure 27A). The breakpoint could therefore be localised to the

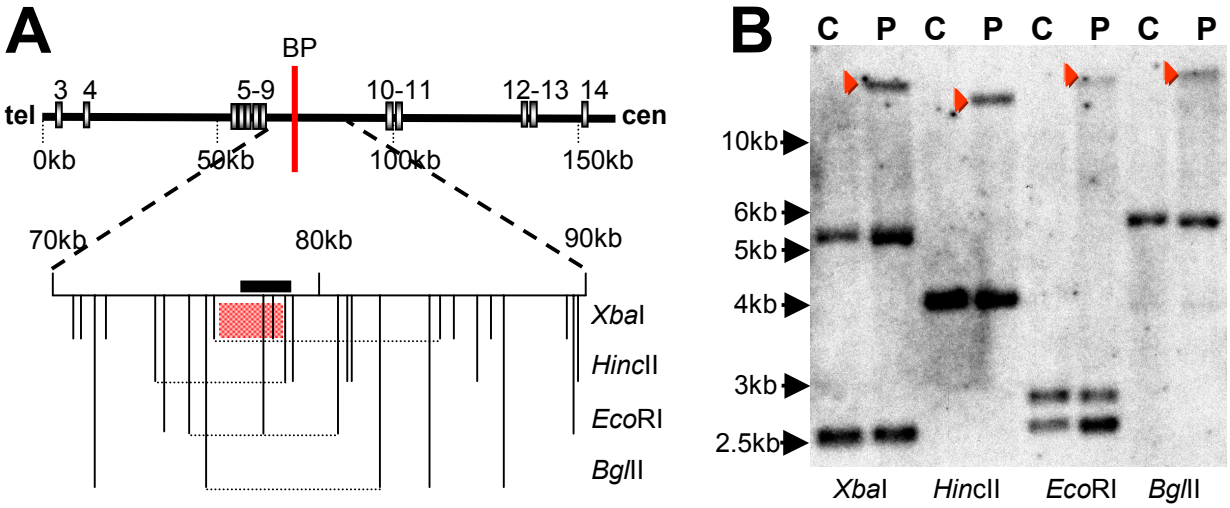


Figure 27: Patient 5 chromosome 4 breakpoint localisation
A: Schematic diagram of Patient 5 chromosome 4 breakpoint region relative to *JNK3* exonic sequence (shaded bars, numbers correspond to those used by Yoshida et al. 2001), and restriction map (expanded insert) of the breakpoint region, indicating the location of the probe used for Southern blot hybridisation in **B,** with respect to the reference genomic *JNK3* sequence. Centromeric (cen) and telomeric (tel) ends are are indicated. Dotted lines in A indicate expected control (C) restriction fragments that hybridise to the probe (black bar corresponding to nucleotides 77714-79011 of the reference *JNK3* sequence) used for Southern hybridisation shown in **B**. Expected fragment sizes are 2540 and 5681, 4219, 2820 and 2524, and 5837 nucleotides for *Xba*I, *Hinc*II, *Eco*RI, and *Bgl*II, respectively. Aberrant restriction fragments (red arrows) are observed in all patient (P) samples in **B**, thereby indicating the breakpoint region (highlighted in red in **A**).

region of sequence overlap within the relevant restriction fragments (highlighted in Figure 27). More specifically, the proximal and distal boundaries of the breakpoint region are marked by the *Hin*cII and *Bgl*II sites at respective nucleotides 78915 and 76023 of the reference genomic *JNK3* sequence used for mapping. The *JNK3* coding sequence is disrupted after exon 9 (exon numbering corresponds to that described by Yoshida et al. 2001).

For each of the *JNK* genes, several splice variants exist; specifically, there are two forms of each gene, referred to as α and β, that differ in their 5'coding sequence. Each of these forms has two variants, 1 and 2, which differ from each other in their N-termini, and thereby code for proteins of slightly different sizes. The longest of the four known human JNK3 isoforms is JNK3α2 (also known as MAPK10 variant 2) with a protein sequence of 464 amino acids. The JNK3α1 (also known as MAPK10 variant 1) has a length of 422 amino acids and is identical to JNK3α2 up to amino acid 417. The human JNK3α isoforms are better characterised than the β isoforms, and for the β isoforms there is conflicting data with respect to the true open reading frames; we therefore discuss the breakpoint location with respect to the longest JNK3α isoform. The corresponding transcript consists of 14 exons separated by 13 introns (see Yoshida et al. 2001). Exons 1 and 2 are solely untranslated sequence; the predicted start codon lies in exon 3. Exon 14 harbours the stop codon and the complete 3'UTR. The Southern blot described above and shown in Figure 27B indicated that the breakpoint sequence lies between exons 9 and 10 i.e. within intron 9, which spans approximately 30 kb. Both JNK3α1 and JNK3α2 (Accession #s NP_002744.1 and NP_620448.1, respectively) are therefore disrupted following amino acid 267; as these two isoforms differ only in their N-termini, the resulting truncated proteins are identical.

3.4.2 Truncated JNK3 lacks kinase domains X and XI and is expressed in the patient lymphoblastoid cell line

As described in detail by Hanks et al. (1988), the conserved region that is responsible for the enzymatic activity of protein kinases can be divided into several domains, referred to as kinase domains I-XI; Yoshida et al. (2001) mapped these kinase domains with respect to the 14 exons that make up JNK3α2. The chromosome rearrangement in Patient 5 results in a truncated gene, for which the predicted protein lacks the C-terminal end, including part of the

kinase domain. More precisely, this truncated open reading frame lacks kinase domains X and XI (for schematic diagram depicting the kinase domains with respect to *JNK3* exons, see Figure 28A). This gene is on chromosome 4; presumably therefore, one intact and functional copy of JNK3 is present in the patient. As expected, we were able to amplify the wild type JNK3 in the patient cell line by RT-PCR (data not shown). We also examined JNK3 in the patient cell line at the protein level, using a commercially available JNK3 antibody (Cell Signaling Technologies). Following cell fractionation of both patient and control lymphoblastoid cells, we identified a specific band at approximately 22 kD present in the microsomal fraction of the patient cell line (Figure 28B).

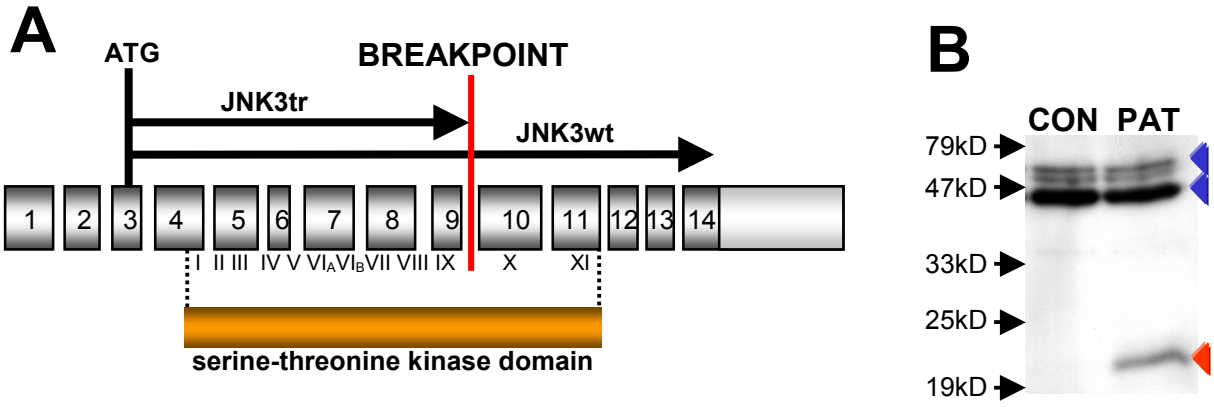


Figure 28: Analysis of truncated JNK3 protein present in Patient 5.
A: Schematic diagram indicating location of the breakpoint relative to the coding sequence and conserved kinase domain of JNK3. Wild type (JNK3wt) and predicted truncated JNK3 (JNK3tr) are depicted. **B:** Western blot on microsomal fraction of lymphoblastoid cells from control (CON) and patient (PAT) with anti-JNK3 (CST, specific for JNKs 1 and 3). Blue arrows indicate normal JNK isoforms; red arrow indicates truncated JNK3. Positions of molecular weight markers are indicated in kilodaltons (kD) to the left.

The presence of this patient-specific protein that binds to the JNK3 antibody strongly suggests that the truncated *JNK3* is in fact transcribed and translated in the patient, rather than being subject to nonsense-mediated mRNA decay, and that the phenotype may result from a dominant effect of the mutant protein rather than a simple partial loss of normal JNK3 function. It is worthy of note that, based on sequence analysis with respect to the breakpoint, truncated JNK3α and JNK3β proteins containing 267 and 229 amino acids are predicted. The expected sizes are 29.7 kD and 25.5 kD, both somewhat larger than the single band observed in our patient, suggesting that only one of these forms is expressed at detectable levels in

lymphoblastoid cell lines, and that it may be subject to some post-translation modification that results in a shorter protein.

3.4.3 Ectopic expression of truncated JNK3 differs from that of the wild type

In order to investigate the function of the truncated JNK3α protein, we over-expressed both the wild-type JNK3α2 and the mutant JNK3α proteins fused to the enhanced green fluorescent protein (EGFP) (for schematic diagram, see Figure 29A) in mouse neuroblastoma (Neuro2A) cells. Both mutant and wild type constructs were cloned into both N-terminal and C-terminal positions with respect to EGFP. Transiently transfected wild type JNK3 fusion proteins were distributed uniformly throughout the cell. This was true for both the C1 construct (C-terminal JNK3α2) and the N3 construct (N-terminal JNK3α2) (Figures 29B(i) and (iii) respectively). The mutant protein (JNK3tr), on the other hand, was not uniformly expressed throughout the cells; rather, it aggregated in those cells expressing significant levels of protein, perhaps due to aberrant protein folding and subsequent poor solubility. This result was observed for fusion proteins with both N-terminal EGFP (Figure 29B(ii)) and C-terminal EGFP (Figure 29B(iv)); however, the severity was more striking for fusion proteins with the truncated JNK3 at the C-terminus (Figure 29B(ii)).

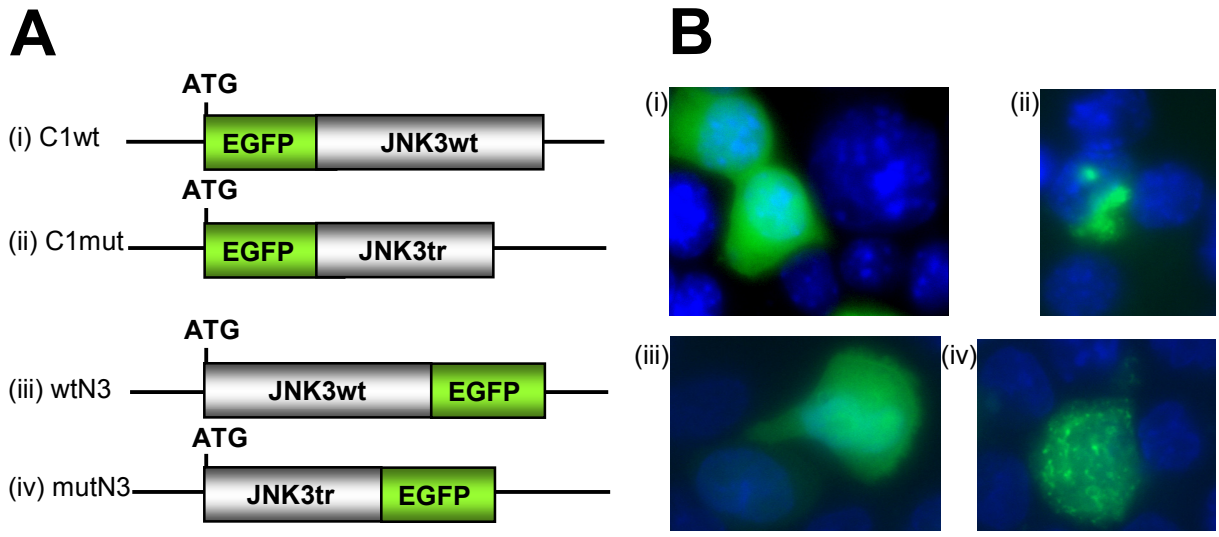


Figure 29: Over-expression of JNK3-EGFP fusion proteins in mammalian cells
A: Schematic diagram depicting (i-iv) JNK3-EGFP fusion constructs (constructed from EGFP C1 and N3 vectors, with respective cloning sites C-terminal and N-terminal of the EGFP tag) used for transfection and cellular localisation experiments. JNK3wt constructs (C1wt and wtN3) contain the 464 amino acids present in JNK3α2; JNK3tr constructs (C1mut and mutN3) contain the N-terminal 267 amino acids common to both JNK3α1and JNK3α2. **B:** Representative examples of Neuro2A cells transfected with constructs (i)-(iv) in **A**, 48 hours post-transfection.

In order to confirm that this result was not an artifact arising from the fusion of the truncated protein to the enhanced green fluorescent protein, which consists of more than 230 amino acids, we also created FLAG™-tagged constructs which harbour the N-terminal amino acid sequence DYKDDDDK as a tag (Figure 30A). This peptide binds to the FLAG™ monoclonal antibody, and presumably has little or no effect on the protein of interest. Immunofluorescence studies (Figure 30B) on HeLa cells over-expressing the tagged JNK3α2 wild type and mutant proteins corroborated the results observed in the EGFP-fusion protein experiments; the mutant protein aggregates within the cytoplasm rather than spreading uniformly throughout the cell as the wild type protein does. Correspondingly, accumulation of such an aberrant protein may have a dominant and detrimental effect on normal JNK signalling in the patient, particularly in neurons, where JNK3 is predominantly expressed. The implications of these results are discussed in detail in section 4.3.

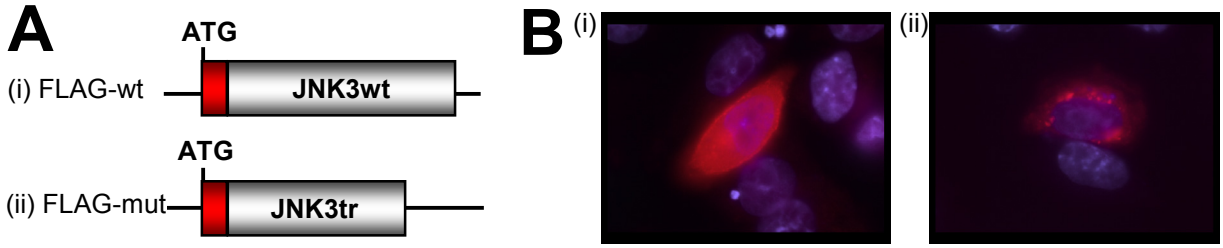


Figure 30: Over-expression of FLAG™-tagged JNK3 proteins in mammalian cells
A: Schematic diagram depicting N-terminal FLAG™-tagged JNK3 constructs used for transfection and cellular localisation experiments. The JNK3wt construct (FLAG-wt) contains the 464 amino acids present in JNK3α2; the JNK3tr construct (FLAG-mut) contains the N-terminal 267 amino acids of JNK3α2. In both constructs, the FLAG epitope (depicted in red) directly follows the start codon. **B:** Representative examples of HeLa cells transfected with constructs (i) and (ii) in **A**, following immunofluorescence with anti-FLAG monoclonal antibody 48 hours post-transfection.