Mutation screening in Rett syndrome patients

Fengqing Xiang, Silvia Buervenich, Piero Nicolao, Mark E S Bailey, Zhiping Zhang, Maria Anvret

Abstract
Rett syndrome (RTT) was first described in 1966. Its biological and genetic foundations were not clear until recently when Amir et al reported that mutations in the MECP2 gene were detected in around 50% of RTT patients. In this study, we have screened the MECP2 gene for mutations in our RTT material, including nine familial cases (19 Rett girls) and 59 sporadic cases. A total of 27 sporadic RTT patients were found to have mutations in the MECP2 gene, but no mutations were identified in our RTT families. In order to address the possibility of further X chromosomal or autosomal genetic factors in RTT, we evaluated six candidate genes for RTT selected on clinical, pathological, and genetic grounds: UBE1 (human ubiquitin activating enzyme E1, located in chromosome Xp11.23), UBE2I (ubiquitin conjugating enzyme E2I, homologous to yeast UBC9, chromosome 16p13.3), GdX (ubiquitin-like protein, chromosome Xq28), SOX3 (SRY related HMG box gene 3, chromosome Xq26-q27), GABRA3 (γ-aminobutyric acid type A receptor α3 subunit, chromosome Xq28), and CDR2 (cerebellar degeneration related autoantigen 2, chromosome 16p12-p13.1). No mutations were detected in these coding regions of these six genes in 10 affected subjects and, therefore, alterations in the amino acid sequences of the encoded proteins can be excluded as having a causative role in RTT. Furthermore, gene expression of MECP2, GdX, GABRA3, and L1CAM (L1 cell adhesion molecule) was also investigated by in situ hybridisation. No gross differences were observed in neurones of several brain regions between normal controls and Rett patients.

Keywords: Rett syndrome; mutation screening; in situ hybridisation; candidate gene

Rett syndrome (RTT) is a childhood neurodevelopmental disorder that almost exclusively affects females, with a prevalence of approximately 1 in 10 000-15 000. It had been suggested to be a genetic disorder because of the existence of a small number of familial cases, and because it shows a higher degree of concordance in monozygotic (MZ) than dizygotic (DZ) twins. Rett syndrome is characterised by stereotypical hand movements, loss of acquired skills, breathing dysfunction, gait abnormalities, autistic behaviour, and mental retardation. Besides the “classical RTT” described by Hagberg et al and Diagnostic Criteria Working Group, there also exist non-classical RTT or Rett variants, known as “forme fruste RTT”. The fact that only females are affected and most cases are sporadic suggested that RTT could be an X linked dominant disease with lethality in affected males. Alternatively, Thomas’ hypothesised that a locus on the X chromosome that was prone to a higher rate of mutation during spermatogenesis than during oogenesis could explain these features. Our finding of linkage of RTT to the telomeric region of Xq is consistent with this model and has recently been validated by other groups. Recently, mutations in the X linked MECP2 gene were identified in some RTT patients by Amir et al. Methyl-CpG-binding protein 2 (MeCP2) is an abundant chromosomal protein that binds specifically to methylated CpG dinucleotides. The protein consists of a methyl-CpG binding domain (MBD), a transcriptional repression domain (TRD), and a corepressor interacting region. MeCP2 protein selectively binds to methylated CpG dinucleotides in the genome and mediates transcription repression through interaction with histone deacetylase and the corepressor SIN3A.

Because MECP2 mutations have not been shown in all affected subjects, there is still the possibility of other genetic factors involved in the pathogenesis of RTT, as earlier suggested by Bühler et al including the X chromosome and autosomal loci.

We have therefore confirmed the frequent presence of MECP2 mutations in our own RTT material and adopted a candidate gene approach for further exclusion of genes involved in RTT. One group of candidates was chosen to examine the possibility that enzymes involved in ubiquitination of proteins might be involved in RTT. Previous studies have shown that mutations in the ubiquitin protein ligase E3A (UBE3A) gene cause Angelman syndrome (AS). which is characterised by mental retardation, seizures, hypotonia, epilepsy, absence of speech, and abnormal gait; these phenotypic characteristics overlap those of RTT to a considerable extent. Consequently, we screened three genes involved in the ubiquitination pathways (UBE1, UBE2I, and GdX). UBE1 catalyses the first step in ubiquitin conjugation. UBE2I is essential for cell viability. GdX is a housekeeping gene that encodes a protein similar to ubiquitin. Further genes (SOX3, GABRA3, CDR2, and L1CAM) were selected based on their expression patterns or their known roles in neural function. Among all these, GdX, GABRA3, and L1CAM, are located in Xq28. We have also found evidence for regions on several autosomes that are concordant in

Candidate gene screening in Rett syndrome

affected patients from several familial Rett cases (unpublished data). Two of the genes examined here, UBE2I and CDR2, are located in one of these candidate regions, chromosome 16p12-p13.

SOX3 is closely related to SRY, the mammalian sex determining gene on the Y chromosome. The mouse homologue of SOX3 is expressed at high levels in neuronal tissues during development. The possibility of compensatory expression of the SRY gene in males in the presence of a mutated SOX3 allele on the X chromosome could account for the lack of the RTT phenotype in males and could lead to disease presentation in females only.

GABRA3 is expressed during development of the brain. It also constitutes a good a priori candidate for the site of the biological lesion in the nervous system. Mutations in L1CAM have already been shown to lead to X linked mental retardation syndromes, including X linked hydrocephalus (HSAS), MASA syndrome (Mental retardation, Aphasia, Shuffling gait, and Adducted thumbs), X linked complicated spastic paraparesis (SP1), and X linked corpus callosum agenesis (ACC).

Materials and methods

Ten unrelated patients with “classical” Rett syndrome and normal controls were used in the mutation screening for six candidate genes. For mutation screening of the MECP2 gene, 59 sporadic cases and nine familial RTT cases were included. Genomic DNA was extracted from blood samples using standard methods. Total RNA was isolated from lymphoblastoid cell lines using the Ultracell™ RNA kit (Bioteck). First strand cDNA was synthesised using the First-Strand cDNA Synthesis Kit (Pharmacia Biotech). PCR amplification was carried out using standard protocols. The coding region of SOX3 is GC rich; therefore, a higher temperature (98°C) was used in the denaturing step during each PCR cycle. Primer pairs were designed from published sequences to amplify the entire coding regions of the seven candidate genes (table 1). Taq PCR Core Kit (Gigen) was used for MECP2-3-2 in order to increase PCR sensitivity. PCR products were isolated from low melting temperature agarose gels and purified using Wizard™ PCR Preps DNA Purification System (Promega). Mutation analysis was performed by direct sequencing (Thermo Sequenase radiolabelled terminator cycle sequencing Kit, BioRad). Materials and methods

Ten unrelated patients with “classical” Rett syndrome and normal controls were used in the mutation screening for six candidate genes. For mutation screening of the MECP2 gene, 59 sporadic cases and nine familial RTT cases were included. Genomic DNA was extracted from blood samples using standard methods. Total RNA was isolated from lymphoblastoid cell lines using the Ultracell™ RNA kit (Bioteck). First strand cDNA was synthesised using the First-Strand cDNA Synthesis Kit (Pharmacia Biotech). PCR amplification was carried out using standard protocols. The coding region of SOX3 is GC rich; therefore, a higher temperature (98°C) was used in the denaturing step during each PCR cycle. Primer pairs were designed from published sequences to amplify the entire coding regions of the seven candidate genes (table 1). Taq PCR Core Kit (Gigen) was used for MECP2-3-2 in order to increase PCR sensitivity. PCR products were isolated from low melting temperature agarose gels and purified using Wizard™ PCR Preps DNA Purification System (Promega). Mutation analysis was performed by direct sequencing (Thermo Sequenase radiolabelled terminator cycle sequencing Kit, BioRad).

Additionally, gene expressions of GdX, GABRA3, L1CAM, and MECP2 were investigated by in situ hybridisation using postmortem brain tissue. The L1 adhesion molecule is expressed mainly on neural cells and is involved in neuro-neurone adhesion and has an important function in the development of the nervous system. Mutations in L1CAM have already been shown to lead to X linked mental retardation syndromes, including X linked hydrocephalus (HSAS), MASA syndrome (Mental retardation, Aphasia, Shuffling gait, and Adducted thumbs), X linked complicated spastic paraparesis (SP1), and X linked corpus callosum agenesis (ACC).

Materials and methods

Ten unrelated patients with “classical” Rett syndrome and normal controls were used in the mutation screening for six candidate genes. For mutation screening of the MECP2 gene, 59 sporadic cases and nine familial RTT cases were included. Genomic DNA was extracted from blood samples using standard methods. Total RNA was isolated from lymphoblastoid cell lines using the Ultracell™ RNA kit (Bioteck). First strand cDNA was synthesised using the First-Strand cDNA Synthesis Kit (Pharmacia Biotech). PCR amplification was carried out using standard protocols. The coding region of SOX3 is GC rich; therefore, a higher temperature (98°C) was used in the denaturing step during each PCR cycle. Primer pairs were designed from published sequences to amplify the entire coding regions of the seven candidate genes (table 1). Taq PCR Core Kit (Gigen) was used for MECP2-3-2 in order to increase PCR sensitivity. PCR products were isolated from low melting temperature agarose gels and purified using Wizard™ PCR Preps DNA Purification System (Promega). Mutation analysis was performed by direct sequencing (Thermo Sequenase radiolabelled terminator cycle sequencing Kit, BioRad).
Results

Mutations in the coding region of MECP2 were found in 35% of our cases. The MECP2 gene consists of three exons. The whole coding regions including the exon-intron junction were amplified by PCR (~1.5 kb, Genbank accession Nos AF030876 and X96866). Primer information is shown in table 1. Altogether, we identified six deletions (from 7 bp up to 101 bp), one insertion (10 bp), six missense mutations, and three nonsense mutations (table 3) in 27 of 78 patients. Five of the six deletions and the insertion cause frameshifts, and the same mutations were not found in their parents or normal sibs. R106W, R255X, and R294X are located in MBD, R306C, R255X, and R294X are located in TRD, and R168X is in the corepressor interacting region. One patient has a 7 bp deletion at nt 1256 and a 70 bp deletion at nt 1172. Another patient has a 10 bp insertion at nt 1127 and a 55 bp deletion at nt 1215. All the deletions and the insertion detected in our patients are clustered in a 150 bp region (nt 1127-1275) which contains tandem repeats. This indicated that the region could be a hot spot for deletion owing to DNA polymerase slippage.

The UBE1 gene transcript (GenBank accession No M58028) is 3.5 kb in length and we used six pairs of primers to amplify the whole coding sequence (3.2 kb). In contrast, a single pair of primers was used to cover the entire coding region (spanning 476 bp) of UBE2I (GenBank accession No U45328). The gene structure of G4X is known (Genbank accession No J03589) and primers were therefore designed using intronic sequence to amplify two PCR fragments, each containing two complete exons and the intervening intron. In order to increase the PCR specificity and sensitivity for G4X, a two round, nested PCR was used (GDX1-2 in table 1), and in addition, the Taq PCR Core kit was used. No mutations were

Table 3 Mutations of MeCP2 found in this study

<table>
<thead>
<tr>
<th>Patients No</th>
<th>Mutation type</th>
<th>Region or domain</th>
<th>Nucleotide change</th>
<th>Amino acid change</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>Deletion</td>
<td>Poly H+poly P</td>
<td>nt1233(del43)</td>
<td>Frameshift</td>
</tr>
<tr>
<td>57</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>57</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>62</td>
<td>—</td>
<td>Poly H+poly P</td>
<td>nt1217(del101)</td>
<td>—</td>
</tr>
<tr>
<td>72</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>86</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>91</td>
<td>—</td>
<td>Poly H+poly P</td>
<td>nt1215(del85)</td>
<td>Frameshift</td>
</tr>
<tr>
<td>91</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>30</td>
<td>Insertion</td>
<td>MBD</td>
<td>390C→T</td>
<td>R106W†</td>
</tr>
<tr>
<td>70</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>80</td>
<td>—</td>
<td>365C→A</td>
<td>D97E</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>—</td>
<td>471C→T</td>
<td>R133C</td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>—</td>
<td>537T→A</td>
<td>F157I</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>Nonsense</td>
<td>TRD</td>
<td>547C→T</td>
<td>T158M†</td>
</tr>
<tr>
<td>42</td>
<td>—</td>
<td>857C→T</td>
<td>R255X†</td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>43</td>
<td>—</td>
<td>954C→T</td>
<td>R294X</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>55</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>33</td>
<td>—</td>
<td>CRIR</td>
<td>576C→T</td>
<td>R168X§</td>
</tr>
<tr>
<td>52</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

*Nucleotide number according to GenBank accession No X99686.
†Amino acid number according to GenBank accession No CA68001.
‡The same as above.
§Patients 28 and 29 are a twin pair.
**Patient 18 is from a Rett “cluster” family.
††Corepressor interacting region.
detected in the coding regions of UBE1, UBE2I, or GdX.

SOX3 comprises a single exon. Three pairs of primers were used to amplify and sequence the 1331 bp region containing the coding sequence. An 8 bp insertion (TGCCGCGCG) in the gene was identified (between bases 1327 and 1328 of the sequence contained in the Genbank entry with accession number X71135). However, the insertion was present in all 12 subjects tested (10 patients and two normal controls) and the deletion variant may therefore represent an error in the original database sequence.

The coding region of GABRA3 spans ~1.4 kb (Genbank accession No A28102) and three primer pairs were used to amplify the coding region. The two most 5’ fragments were amplified using a two round, nested PCR strategy to increase PCR specificity and sensitivity. Three pairs of primers were designed to amplify the entire cDNA sequence of CDR2 (~1.3 kb, Genbank accession No M63256). No mutations were found in either gene.

From the in situ hybridisation experiments, representative photomicrographs are shown in figs 1 and 2. L1CAM mRNA was detected at relatively high levels in neurones and lower levels in glia in all brain regions investigated (prefrontal cortex shown in figs 1 and 2, motor cortex, thalamus, and brain stem). A strong ubiquitous signal was generated by MECP2 specific oligonucleotides (fig 2). Low levels of GABRA3 and GdX expression were found in sections containing tissue from these regions (fig 1). There were no detectable differences in

Figure 1 Bright field microscopy of neurones in the frontal cortex showing GABRA3 and L1CAM expression. Panels show representative fields from sections from a normal control (A-D) and a Rett patient (E-H). The first (A, E) and the second (B, F) probe for GABRA3 showed similarly low levels of expression. Relatively strong expression levels were detected with the probe against L1CAM. (C, G) Hybridisation with a random oligonucleotide resulted in low background signal (D, H). No differences could be observed between patients and controls for any of the probes used.

Figure 2 Bright field microscopy of neurones in the frontal cortex showing MECP2 and L1CAM expression. Panels show representative fields from sections from a normal control (A-C) and a Rett patient (D-F). MECP2 (A, D) was expressed at similar levels as L1CAM (B, E). The negative control hybridisation (C, F) displays very low background signal. MECP2 was expressed ubiquitously in patients and controls in all brain tissues tested (besides the prefrontal cortex shown in this figure we also investigated expression in the motor cortex, thalamus, and brain stem, data not shown). No differences in expression between different brain regions or patients versus controls could be detected.
expression levels or patterns for any of these four genes between Rett patients and controls.

**Discussion**

Recent evidence from genetic mapping using familial Rett cases has strongly indicated the presence of a gene for RTT in Xq28, finally leading to the discovery of mutations in MECP2. Because not all patients investigated so far carry mutations in MECP2 and because of the clinical variability seen in RTT, it is possible that this disorder may be genetically heterogeneous. In this study, we therefore determined the percentage of our Rett patients with MECP2 mutations to be 35% and screened further X chromosomal and autosomal candidate genes on the basis of their genomic locations or physiological or genetic characteristics. For some of the candidate genes investigated here, we chose to sequence cDNA products, because full gene structures and sequences were not available for all the genes. In this way, we expected to be able to find mutations in the coding sequence that would be predicted to alter the amino acid sequence of the proteins. This strategy cannot identify mutations in the promoter or 5' or 3' untranslated regions. In addition, truncating mutations that reduce RNA stability might not be apparent. We can be sure that the latter problem was not encountered for most of the genes investigated here, because we were able to show the presence of heterozygous polymorphisms in several subjects (not shown), suggesting that effective biallelic amplification was occurring. It would clearly be desirable to extend these studies to include the remaining parts of these genes, to investigate whether other common types of mutation in these regions play a role.

By screening the MECP2 gene, 16 mutations were identified, including missense mutations, nonsense mutations, an insertion, and deletions. Most of the point mutations are C→T transitions, which are located in the MBD or TRD or corepressor interacting region. The six deletions and one insertion are exclusively clustered within a 150 bp region of exon 3. Surprisingly, we detected two de novo deletions in the “hot spot” region in one patient and one deletion and one small insertion in the same region in another patient. Since it is extremely rare that two de novo mutations happen to occur in a small region and occur twice (two patients), the region is likely to be a hotspot for deletion or insertion. This region contains several tandem repeats; therefore, the region could be unstable owing to DNA polymerase slippage.

The missense mutation R306C was found in seven out of 78 patients (9%). In order to rule out the possibility of a common polymorphism, we used HhaI restriction enzyme to test 47 ethnically matched normal controls. None of the 47 normal controls carried this polymorphic site (data not shown). After completely sequencing the entire coding region, mutations were detected in only 35% of patients. There are some mutations that may be missed if they are located in introns, the promoter region, or the 3' untranslated region or consist of huge deletions. However, we have not identified any mutation in our nine familial RTT cases. Our results indicated that there might be more than one gene for RTT. MECP2 may play a key role in the development of the disease. Other genes that may interact with MECP2 or in the same pathway may contribute to the subset of Rett patients remaining unexplained so far.

Skewed X chromosome inactivation has been found in some of our familial RTT cases (paper submitted). The XIST gene (X inactive specific transcript), located at the X inactivation centre, XIC, and expressed exclusively from the inactive X chromosome, is thought to be the major gene involved in initiating the process of X chromosome inactivation. Previous studies of two unrelated families with skewed X chromosome inactivation have shown a promoter mutation in the XIST in nine females, indicating that there may be an association between alterations in the regulation of XIST expression and X chromosome inactivation. We then used the patients with skewed X chromosome inactivation and some other RTT patients without MECP2 gene mutation to test if there is any mutation in the promoter region of XIST. We could not find any mutation in the promoter region of XIST in any of the 10 patients tested (data not shown).

We also used in situ hybridisation to investigate the expression levels of a set of candidate genes. We predicted that mutations in these genes that caused gross alterations in transcription rates or in mRNA stability or localisation would be detected by this technique. The ability to detect such changes might be of use in narrowing down the search for candidates on which to perform a mutation scan. In the case of MECP2, changes in regional or global expression levels may be important clues for understanding of the pathogenetic mechanisms underlying the syndrome. We observed no gross alterations of expression levels in the four genes tested (GABRA3, GdX, LICAM, and MECP2), but two of them, GABRA3 and GdX, were expressed at relatively low levels and subtle changes in expression pattern or levels in patients would therefore be difficult to detect. When using postmortem material, there is the added danger that global degrading processes acting on mRNA will have proceeded too far to be able to detect transcripts. In this study, we detected high levels of LICAM and MECP2 transcripts, suggesting that this problem did not occur in our samples.

We thank all the patients and their families for their participation, all doctors who provided the sample materials, Dr Nessa Carey and Dr Mark G Darlison for helpful comments, Dr Laurent Villard for helping design primers for MECP2 exon 1, and Mrs Ann-Christin Thelander for excellent technical assistance. This work is supported by the Axel and Margaret Axson Johnson Foundation, the Emil and Wera Cornell Foundation, the Ragnar Söderberg Foundation, the International Rett Syndrome Association, and Karolinska Institutet, NIMH and USPHS funds. Brain tissues were provided by the Harvard Brain Tissue Resource Center, which is supported in part by PHS grant number M01RR00274.

12 Nan X, Campoy FJ, Bird A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chroma-
27 Dagerlind Å, Friborg K, bean AG, Hökfelt T. Sensitive mRNA detection using unfixed tissue: combined radio-