ORIGINAL INVESTIGATION

Esteban Ballestar · Santiago Ropero · Miguel Alaminos
Judith Armstrong · Fernando Setien · Ruben Agrelo
Mario F. Fraga · Michel Herranz · Sonia Avila
Mercedes Pineda · Eugenia Monros · Manel Esteller

The impact of *MECP2* mutations in the expression patterns of Rett syndrome patients

Abstract Rett syndrome (RTT), the second most common cause of mental retardation in females, has been associated with mutations in MeCP2, the archetypical member of the methyl-CpG binding domain (MBD) family of proteins. MeCP2 additionally possesses a transcriptional repression domain (TRD). We have compared the gene expression profiles of RTT- and normal female-derived lymphoblastoid cells by using cDNA microarrays. Clustering analysis allowed the classification of RTT patients according to the localization of the MeCP2 mutation (MBD or TRD) and those with clinically diagnosed RTT but without detectable MeCP2 mutations. Numerous genes were observed to be overexpressed in RTT patients compared with control samples, including excellent candidate genes for neurodevelopmental disease. Chromatin immunoprecipitation analysis confirmed that binding of MeCP2 to corresponding promoter CpG islands was lost in RTT-derived cells harboring a mutation in the region of the *MECP2* gene encoding the MBD. Bisulfite genomic sequencing demonstrated that the majority of MeCP2 binding occurred in DNA sequences with methylation-associated silencing. Most importantly, the finding that these genes are also methylated and bound by MeCP2 in neuron-related cells suggests a role in this neurodevelopmental disease. Our results provide new data of the underlying mechanisms of RTT and unveil novel targets of MeCP2-mediated gene repression.

E. Ballestar · S. Ropero · M. Alaminos · F. Setien · R. Agrelo
M. F. Fraga · M. Herranz · S. Avila · M. Esteller (✉)
Cancer Epigenetics Laboratory,
Molecular Pathology Programme,
Spanish National Cancer Centre (CNIO),
Melchor Fernández Almagro 3, 28029 Madrid, Spain
Tel.: +34-912246940
Fax: +34-912246923
E-mail: mesteller@cnio.es
URL: http://grupos.cnio.es/epigenetica/

J. Armstrong · E. Monros
Genetics Section, University Hospital Sant Joan de Deu,
Esplugues, Barcelona, Catalonia, Spain

M. Pineda
Neurology Service, University Hospital Sant Joan de Deu,
Esplugues, Barcelona, Catalonia, Spain

Introduction

Rett syndrome (RTT, MIM 312750) constitutes an archetypical example of a disease in which both genetic and epigenetic defects coincide. Clinical features of this disorder, which occurs with an incidence of one in 10,000–15,000 women, do not develop until 6–18 months of age and include deceleration of head growth, loss of purposeful hand movements, and autistic features among others (Armstrong 1997). The identification of mutations in the *MECP2* gene (MIM 300005) as a common feature in RTT patients represented a key breakthrough in the understanding of the disease (Amir et al. 1999; Webb et al. 2001). Mutations in the *MECP2* gene are found in approximately 70–80% of classic RTT patients (Van den Veyver et al. 2000) and the remaining cases result from large deletions of the *MECP2* gene (Laccone et al. 2004), mutations in non-coding regions of *MECP2* gene, and perhaps, from mutations in other genes (locus heterogeneity).

The *MECP2* gene encodes a protein that preferentially binds methylated CpG dinucleotides and, in turn, mediates transcriptional repression through the recruitment of histone deacetylases and other corepressors (Jones et al. 1998; Nan et al. 1998).

CpG island methylation is associated with transcriptional repression (Bird and Tweedie 1995). Although very few examples of genes regulated by promoter methylation during development or tissue-specific expression have been reported, CpG methylation has been implicated in stable alterations in gene expression in cancer (Esteller 2000) and plays an important role in inactivation of the X-chromosome and genomic

imprinting (Nan and Bird 2001). Loss of function of MeCP2 in RTT has been postulated to involve the dysregulation of genes that are regulated upon interaction of MeCP2 to their methylated promoters (Willard and Hendrich 1999).

MeCP2 contains two well-defined domains: a methyl-CpG binding domain (MBD), common to all the methyl-CpG-binding proteins, and a transcriptional repression domain (TRD), involved in recruiting Sin3A and histone deacetylase to repress transcription (Jones et al. 1998; Nan et al. 1998). Most RTT-associated missense mutations in the *MECP2* gene are clustered within the region encoding the MBD, while deletion and insertion mutations are concentrated in the C-terminal region (Kriaucionis and Bird 2003). Overall, eight mutational hotspots account for 67% of all mutation-positive cases, four at the MBD (R106, R133, T158, R168) and four at the TRD (R255, R270, R294 and R306). Mutations within the MBD impair MeCP2 for binding to methylated DNA (Ballestar et al. 2000) and both MBD and TRD mutations affect the ability of MeCP2 to repress transcription when tethered to a promoter (Yusufzai and Wolffe 2000). A hotspot of mutations at the C-terminus of MeCP2 indicates the existence of an additional functional domain (Vacca et al. 2001). In particular, a domain with the ability to bind WW domains has been recently found at the C-terminus of MeCP2 (Buschdorf and Stratling 2004). Finally, recent data have shown that the *MECP2* gene suffers an event of alternative splicing that produces two variants with different first exons, MeCP2A and MeCP2B. The latter is the only one that has deletions associated at its first exon in RTT patients (Mnatzakanian et al. 2004).

The identification of genes whose expression is regulated by MeCP2 is essential to understanding the pathway of this disease. Recent reports have investigated the transcriptional profile of RTT-derived samples or mouse models that mimic RTT (Colantuoni et al. 2001; Traynor et al. 2002; Tudor et al. 2002). The results of these experiments have revealed that expression changes are subtle, indicating that regulation exerted by MeCP2 could be very specific. However, pure expression studies on MeCP2-deficient samples do not inform whether those changes are a consequence of a direct MeCP2 lack of function over certain gene promoters or a secondary effect due, for instance, to the indirect disregulation of a transcription factor. It is, therefore, an issue of inherent interest to distinguish between direct or indirect MeCP2 targets as it can be potentially helpful in the design of new therapeutical strategies.

In order to examine the putative distorted expression profile of RTT patients, we established six lymphoblastoid cell lines derived from patients with different mutations in the MBD and TRD regions of MeCP2. Furthermore, four patients with clinically diagnosed RTT, but without apparent mutations in the coding region of *MECP2*, were also included in the study. Relative gene expression profiles, compared with normal females, were determined by cDNA microarray analysis.

Both global and specific disregulation of the expression pattern was observed, with overexpression of a significant number of genes in RTT patients. Most importantly, the vast majority of upregulated genes exhibited MeCP2/methylation-mediated silencing in normal cells that was lost in the RTT cells. This was confirmed by combining chromatin immunoprecipitation analysis (ChIP), bisulfite genomic sequencing and treatment with demethylating agents. ChIP assays, bisulfite sequencing and expression analysis of neuron-related cell lines indicate that aberrant overexpression of these genes may be relevant to the neurodevelopmental disease. In summary, our results demonstrate an essential role for MeCP2 in the specific repression of a set of methylated genes and the consequences of the lack of functional MeCP2 in RTT patients.

Materials and methods

Establishment of B lymphoblastoid RTT cell lines and neuroblastoma cell cultures

Samples of peripheral blood were obtained from six different RTT patients for which mutations in MeCP2 had been determined (Monros et al. 2001) and four patients with clinically diagnosed RTT but without detectable MeCP2 mutations in the coding region (Monros et al. 2001). The mutation analysis had been previously characterized by direct sequencing of the coding region of the *MECP2* gene (including the exon 1 open reading frame of MeCP2B) (Monros et al. 2001; Mnatzakanian et al. 2004). We first isolated peripheral blood mononuclear cells (PBMCs) from blood by density sedimentation with Ficoll/Hypaque (Amershan) and used a standard protocol of infection of B lymphocytes with Epstein–Barr virus (EBV) to produce immortalized lymphoblastoid cell lines (Kempkes et al. 1995). The selected RTT cell lines used in the study expressed more than 50% of the mutant *MECP2* allele according to direct cDNA sequencing. The neuroblastoma cell lines SK-N-AS and IMR-32 were cultured in Dulbecco's modification of Eagle's medium: with 4.5 g/l glucose and L-glutamine supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin.

HUMARA assay for clonality

The analysis of X-chromosome inactivation was performed by the human androgen-receptor (HUMARA) assay as reported (Allen et al. 1992). Briefly, DNA samples were first digested with a methylation sensitive restriction enzyme, *Hpa*II, in order to cleave the unmethylated, active alleles of the *HUMARA* gene. After digestion, the DNA samples were amplified by PCR of the *HUMARA* gene. The sequences of the *HUMARA* primers were as follows: forward primer, 5′-GCTGTG-AAGGTTGCTGTT-3′, and reverse primers, 5′-TCCA-

GAATCTGTTCCAGAGCGTGC-3′. Samples were analyzed in a non-denaturing 6% 29:1 acrylamide/bis-acrylamide gel and stained with ethidyum bromide.

Microarray analysis

The profiles of gene expression were determined by cDNA microarray analysis with the CNIO Oncochip. In this microarray, cDNAs were printed onto chemically activated glass slides (CMT-GAPS; Corning, Corning, N.Y., USA) using the spotter Multigrid II (BioRobotics, Woburn, Mass., USA). The cDNA microarray consists of 7,237 sequence validated I.M.A.G.E. clones, including 5,253 clones representing known genes and the remaining 1,984 clones representing expressed sequence tags (ESTs). Human cDNA clones were purchased from Research Genetics (Huntsville, Ala., USA). The list of genes on the array can be found at: http://bioinfo.cnio.es/data/oncochip. The CNIO Oncochip has been previously validated for the study of the patterns of gene expression in different tumor types (Tracey et al. 2002; Moreno-Bueno et al. 2003).

In our study, total RNAs, extracted with the RNeasy kit (Qiagen) from each of the control and RTT cell lines, were converted to double-stranded cDNA using the superscript choice system (Life Technologies) with an oligo(dT) primer containing a T7 RNA polymerase promoter. Fluorescent first-strand cDNA was made in the presence of 50 μmol/l of Cy5-dCTP (red) for each sample or Cy3-dCTP (green) for Universal Human Reference RNA (Stratagene). This reference comprises a collection of RNA pooled from ten human cell lines for optimal broad gene coverage and provides the ability to cross-compare data sets from multiple experiments as a single, common control. Slides were simultaneously hybridized with labeled sample and references in the presence of the following competitors: human Cot1 DNA (Invitrogen), yeast total tRNA (Sigma) and poly(A) DNA (Amershan). Slides were then scanned for Cy3 and Cy5 fluorescence using Scanar ray 5000 XL (GSI Lumonics Kanata, ON, Canada) and quantified using the Quantarray software (GSI Lumonics) and GenePix Pro 4.0 software (Axon Instruments, Union City, Calif., USA).

Fluorescence intensity measurements from each array probe were compared with local background (measurement outside the spots) and subtraction of this was performed; spots that this value was lower than 500 fluorescence units were excluded. Scanned data were adjusted to a matrix generated by the Gene-Pix 4.0 program for the identifications of all the clones. Data were preprocessed (Herrero et al. 2003) in the following way: (1) log-transformation to obtain symmetrical ratios, (2) replicate handling (removing inconsistent replicates and merging the remaining ones), (3) missing value management (Troyanskaya et al. 2001), (4) flat pattern filtering by standard derivation, and (5) pattern standardization by subtracting the pattern average and dividing the values by the standard derivation. We performed cluster analysis employing the self-organizing hierarchical neural network SOTA (Dopazo and Dopazo 1997), an unsupervised neural network with a binary tree topology, combining the advantages of divisive and customizable methods. The analysis was performed blind, whereby the MeCP2 mutation type corresponding to each sample was decoded at the end of the process.

In order to identify genes associated with specific groups of samples, i.e. genes differentially expressed in RTT cases versus control cells or in RTT cases with TRD mutations versus RTT cases with a different kind of alteration, and to take both magnitude and rank into account, we used the following two-step approach. First, the average fold-change was calculated as a ratio of the means of the comparison groups. Then, a statistical test was carried out to identify genes with a strong, significant differential expression between the groups of comparison. Two different tests were used: significant analysis of microarray (SAM) (Tuscher et al. 2001) and Mann–Whitney non-parametric test. SAM computes a statistic d_i for each gene i, measuring the strength of the relationship between gene expression and the response variable. It uses repeated permutations of the data to determine if the expression of any genes is significantly related to the response. The cut-off for significance is determined by a tuning parameter delta, chosen by the user based on the false discovery rate (FDR). On the other hand, the non-parametric test of Mann–Whitney is universally used because it applies rank statistics and, thus, ignores the magnitude of the expression values. Along the text, genes with an average fold-change of more than 2 and a Mann–Whitney P value < 0.01 were considered as differentially expressed between both groups of comparison (Alaminos et al. 2003), as well as genes with an average fold-change of more than 2 and FDR $< 1\%$ for the SAM. A complete list of the identities and expression levels of all the genes studied in the microarray analysis is available upon request.

Chromatin immunoprecipitation (ChIP) assay

In order to investigate the existence of MeCP2 association to the selected promoters, standard ChIP assays were performed as previously described (Fournier et al. 2002). Commercial antibodies against MeCP2 (C-terminus) (Stancheva et al. 2003) and antisera raised against the N-terminal portion of MeCP2 were used (Fournier et al. 2002). Additionally, antibodies against other MBD proteins were also used (Fournier et al. 2002). In all cases, chromatin was sheared to an average length of 0.25–1 kb for this analysis. The ChIP assays were performed with the two normal immortalized lymphocytes and the RTT cell line harboring the MeCP2 T158M mutation in the MBD region that reduces its affinity for methylated DNA (Yusufzai and Wolffe 2000). We first evaluated the sensitivity of the PCR amplification on serial dilutions (0–100 ng) of total DNA collected after sonication (input DNA) with

specific primers for the promoter region of all selected genes, in order to obtain conditions in which a linear PCR amplification occurs. Primer sequences are available upon request.

Global 5-methylcytosine quantification

Quantification of the 5-methylcytosine (mC) content was carried out by high performance capillary electrophoresis as previously described (Fraga and Esteller 2002). In brief, immunoprecipitated DNA samples were speed-back preconcentrated to 0.1 mg/ml and enzymatically hydrolyzed in a final volume of 5 µl. Samples were then directly injected in a Beckman MDQ high performance capillary electrophoresis apparatus and mC content was determined as the percentage of mC of total cytosines: mC peak area×100/(C peak area + mC peak area).

Bisulfite sequencing

We established the methylation status of CpG islands of selected genes by sequencing bisulfite-modified genomic DNA. We carried out bisulfite modification of genomic DNA, followed by PCR amplification, DNA isolation and direct sequencing of both strands as previously described (Clark and Warnecke 2002). We designed all the bisulfite genomic sequencing primers according to GenBank data around the presumed transcription start sites of the investigated genes. Primer sequences and PCR conditions for methylation analysis are available upon request. The degree of methylation for a particular CpG site was estimated with the relative height of the peak C versus the peak T in the sequencing electropherogram. Each fragment was sequenced three times.

Semiquantitative RT-PCR expression analysis

We reverse-transcribed total RNA (2 µg) treated with DNase I (Ambion) using oligo(dT) primer with Superscript II reverse transcriptase (Gibco/BRL). We carried out PCR reactions in a 25-µl volume containing 1×PCR buffer (Gibco/BRL), 1.5 mM of MgCl$_2$, 0.3 mM of dNTP, 0.25 µM of each primer and 2 U of *Taq* polymerase (Gibco/BRL). We used 100 ng of cDNA for PCR amplification, and we amplified all of the genes with multiple cycle numbers (20–35 cycles) to determine the appropriate conditions for obtaining semiquantitative differences in their expression levels. RT-PCR primers were designed between different exons to avoid any amplification of genomic DNA. Primer sequences are available upon request.
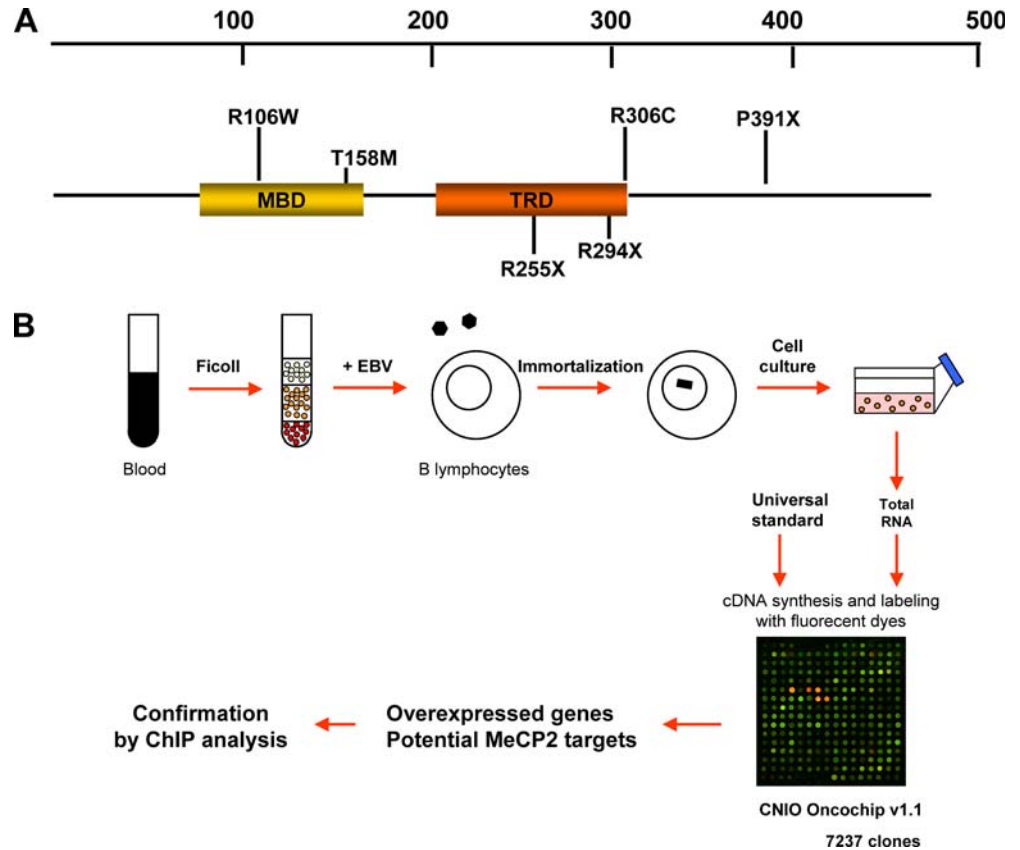
Results

In order to characterize the spectrum of gene disregulation due to MeCP2 loss of function, we generated B lymphoblastoid cell lines from patients representing subtypes of RTT with and without mutations in the MeCP2 coding region. To this end, we first obtained blood samples from female patients who had been previously clinically diagnosed with classical RTT and for which the spectrum of MeCP2 mutations was known (Monros et al. 2001). We selected patients with mutations within the different domains of MeCP2: two patients with mutations on the MBD of MeCP2 (R106W, T158M), three patients with mutations on the TRD (R255X, R294X, R306C), one patient with a C-terminal mutation (P391X) and four patients without mutations, at least within the coding region (Fig. 1a). The age of the affected donors relative to the onset of the disease ranged between 12 and 20 months, with an average of 16.5 months. The age of the patients ranged between 10 and 16 years old. As controls, we also obtained blood samples from two normal female individuals with ages within this range. Lymphocytes were transformed with EBV to produce immortalized polyclonal lymphoblastoid cell lines. We decided to maintain polyclonality, in contrast with monoclonal systems that other groups have investigated (Traynor et al. 2002) as it may represent a situation closer to the biological situation that is found in RTT patients. The X-chromosome inactivation status of the RTT cell lines, studied by HUMARA assay (Fig. 2), as well as direct sequencing of cDNA, indicated that a significant proportion of both paternal and maternal alleles were indeed represented in most of all the samples (Fig. 2) (those with a marked skewed pattern were enriched with the mutant form). Following immortalization, total RNA was extracted, hybridized to cDNA microarrays and the data analyzed, as described in Materials and methods (Fig. 1b).

Hierarchical cluster analysis of all samples demonstrated that the expression profiles corresponding to normal controls are globally different to that of RTT patients (Fig. 3). As a consequence of that, all RTT cases, even when harboring different kinds of mutations, cluster together in the same dichotomic branch of the tree, suggesting that they are more similar to one another than to normal controls (Fig. 3). These findings are consistent with the fact that all RTT patients develop relatively similar phenotypes and clinical features regardless of the type of mutations they exhibit. However, cells with different MeCP2 mutations tended to segregate into different subbranches of the secondary cluster tree (Fig. 3). In short, both MBD mutations (T158M and R106W) show a differential pattern of expression, and one mutation seems to be different to the other one. On the other hand, the C-terminal mutated cell type P391X clustered together, with the three cases showing TRD mutations (R255X, R294X and R306C), suggesting a common expression profile for both types of MeCP2 mutations. Finally, the four cases with no detectable mutations for MeCP2 were grouped in the same branch of the tree (Fig. 3), implying that these cases might share a common pattern of gene expression and, perhaps, a common molecular pathway. Taking

Fig. 1 a Diagram showing the list of Rett mutations used in this study. Additionally we obtained cell lines from four RTT patients without mutations in the *MECP2* coding sequence. **b** Scheme showing the generation of B lymphoblastoid cell lines from Rett blood samples. Lymphocytes are isolated from whole blood by gradient centrifugation with Ficoll and transformations are initiated by Epstein–Barr (EBV) infection. They can then be propagated indefinitely in culture for experiments. Total RNA is then extracted. Copy DNA is synthesized from total RNA and subsequently labeled with Cy3. In parallel, a universal standard is used to synthesize a control cDNA that is labeled with Cy5. Hybridizations are performed in a cDNA microarray. Overexpressed genes are potential candidates to identify MeCP2 targets

into account the results of all the experiments, 2.55% of all the genes represented in the microarray (259 cDNA probes) were upregulated in the RTT patients (Mann–Whitney $P < 0.01$ and average fold-change > 2), a value in agreement with other human (Colantuoni et al. 2001; Traynor et al. 2002) and mouse (Tudor et al. 2002) experimental RTT models, and 257 probes (2.53%) were downregulated. This finding is consistent with the hypothesis that RTT-associated mutations of MeCP2 compromise its repressing activity, leading to target gene activation. In the case of genes that are downregulated, it is likely that this behavior is an indirect consequence of the misregulation of transcriptional factors that had their levels increased as a direct or indirect consequence
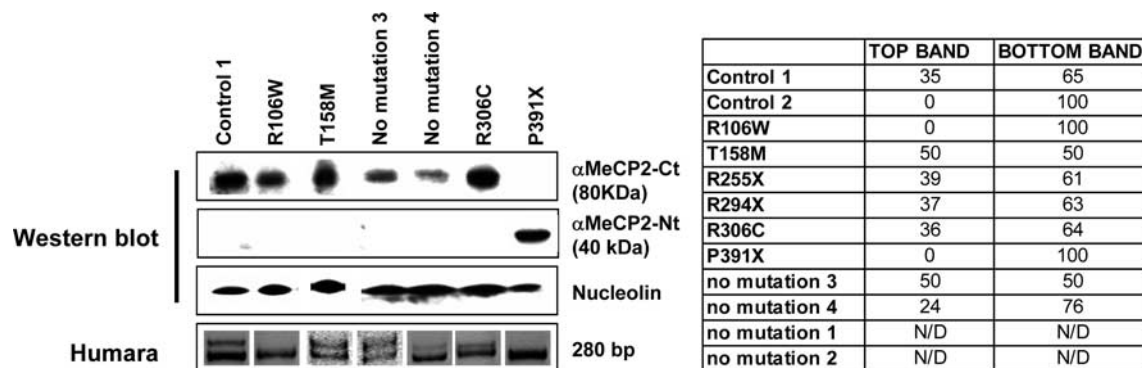

	TOP BAND	BOTTOM BAND
Control 1	35	65
Control 2	0	100
R106W	0	100
T158M	50	50
R255X	39	61
R294X	37	63
R306C	36	64
P391X	0	100
no mutation 3	50	50
no mutation 4	24	76
no mutation 1	N/D	N/D
no mutation 2	N/D	N/D

Fig. 2 Western blot and HUMARA assay. The *left panel* shows representative examples of both Western blots and HUMARA assays corresponding to different normal and RTT-derived lymphoblastoid cells. The antibody against the C-terminus of MeCP2 recognizes the protein at around 80 kDa in all lanes except P391X in which the C-terminus is truncated. Conversely, the antibody against the N-terminus of MeCP2 only produces a band at around 40 kDa in P391X cells. PCR products of HUMARA assays (Allen et al. 1992) are around 280 bp. The relative proportion of the *top* and *bottom band* indicates the relative proportions of the X-inactivation levels of paternal and maternal alleles. The *table on the right* summarizes the results of HUMARA assays
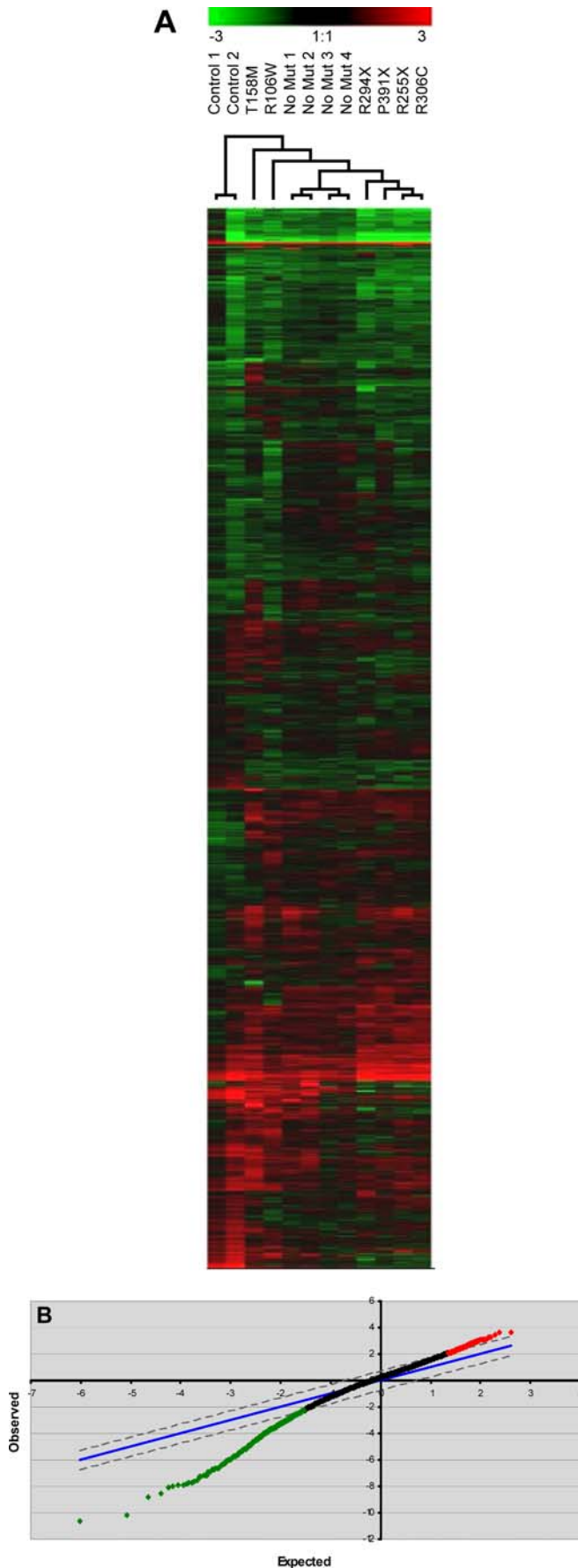
of MeCP2 misfunction. Possibly both phenomena, direct and indirect upregulation of genes and indirect down-regulation of genes, may have relevance to the RTT phenotype. In the case of overexpressed genes, the degree of upregulation ranged between two- and tenfold in most cases, and only a few genes were overexpressed more than tenfold. Most interestingly, although some upregulated genes are common among RTT patients, many genes upregulated in those RTT patients without mutations identified in the MeCP2 coding region were different to those of cases with characterized mutations. These findings, in agreement with the segregation of the cases in the cluster analysis, suggest a common alternative causal molecular pathway for the RTT phenotype in the absence of mutation. In addition, the SAM analysis (see Materials and methods) showed that 4,086 genes were differentially expressed between the group of RTT cases and the controls, with a Δ value of 0.36282 corresponding to a FDR of 30.65322 (see Fig. 2) and 83 genes differentially expressed for a Δ value of 2.30422 (FDR of 0.81436).

Of all the genes included in the microarray analysis, 30 genes were selected for further study (Table 1). As shown in Table 1, 21 of these genes were selected because of their relative overexpression in RTT cases in comparison with the control samples (average fold-change expression > 2 and Mann–Whitney and/or SAM analyses statistically significant, see Materials and methods), whereas five genes were chosen due to their significant differential expression (Mann–Whitney $P < 0.01$) in cases with TRD mutations when compared with the rest of the RTT samples (the genes *AGTR1*, *CSE1L*, *CYPD*, *PAM* and *BIRC2*). An additional set of four genes was also selected because of the relevant role of their products in neural development [*NET1* (Chan et al. 1996) and *MYCN*], skeletal muscle development (*FHL1*) or both [*PGK1* (Sugie et al. 1994)], despite they did not match statistical significance.

Among the 21 genes overexpressed in RTT patients when compared with the control cells, excellent candidates for the neurodevelopmentally altered phenotype of RTT patients were found. Key examples include: the guanine nucleotide binding protein alpha (*GNAS*), involved in the modulation of voltage-dependent calcium channels of neuroendocrine cells (Hescheler and Schultz 1994); the myosin regulatory light chain interacting protein (*MIR*), a novel ERM-like protein that interacts with myosin regulatory light chain and inhibits

Table 1 Selected genes from the microarrays analysis of ten RTT-derived lymphoblastoid cell lines. *No. cells with FC > 2* the number of cases showing a significant increase in comparison with the average controls (third column). The *d* and *q* values assigned by the SAM algorithm to each significant gene (for the delta value selected) are shown in columns 6 and 7. Columns 8–10 represent the average fold-change (*FC*) for the comparison of each case with the average controls (*FC RTT/C*), cases with a characterized mutation versus cases with no mutation detected (*FC MUT/NOMUT*) and RTT cells with a TRD (*T*) mutation vs cases without this mutation [cells harboring MBD (*M*) or C-terminal (*C*) mutations and cells lacking any mutations] [*FC 3/(1 + 2 + NOMUT)*]. The Mann–Whitney *P* values (*MW*) for the same comparisons are presented in the last three columns (a value below 0.01 is considered as significant)

Gene	MIM	No. cells with FC > 2	Function	Significant for delta 0.36282	Score (*d*)	*q*-value (%)	FC RTT/C	FC MUT/NOMUT	FC T/(M + C + NOMUT)	MW R/C	MW MUT/NOMUT	MW T/(M + C + NOMUT)
IGF2	147470	7	Cell proliferation	Yes	1.7712	13.8790	2.2143	1.2725	0.8354	0.0303	0.1143	0.0333
MPP1	305360	7	Signal transduction	Yes	1.6019	15.6818	3.3281	1.6139	0.9465	0.0303	0.1143	0.1167
ITGA3	605025	7	Cell adhesion	Yes	3.0018	7.5493	2.4838	1.0076	0.8430	0.2727	0.2571	0.2667
MIR	29116	5	Links actin to membrane-bound proteins	Yes	1.9393	13.8790	1.7813	1.2517	2.3836	0.0610	0.2570	0.0170
GNAS	139320	1	Guanine nucleotide binding protein	Yes	0.8905	27.7571	1.3222	0.9564	1.8829	0.0303	0.1143	0.0333
TOP2A	126430	1	DNA transcription and replication	Yes	0.9387	27.7571	1.3441	1.0866	1.0858	0.0303	0.2571	0.2667
DNAJB1	604572	3	Heat shock protein activity	Yes	1.3442	17.7123	1.7028	1.4029	1.0236	0.0303	0.1714	0.1167
SLC2A3	138170	5	Glucose transporter	Yes	1.3522	17.7123	2.5672	1.5747	0.8785	0.0303	0.4762	0.6667
ITGA2	192974	1	Adhesion and cell-surface signaling	Yes	1.0199	25.6803	1.4048	0.8359	1.1031	0.0303	0.2571	0.8333
PGK1	311800	1	Induces multidrug resistance	–	–	–	1.2294	1.2834	1.3457	0.6061	0.7619	0.0667
IGFBP2	146731	0	Cell proliferation	–	–	–	1.0020	0.7082	1.3564	0.0303	0.3524	0.2667
NET1	606450	1	Guanine nucleotide regulatory protein	–	–	–	1.3184	1.1189	0.8490	0.4848	0.0667	0.3833
PTPRCAP	601577	0	Key regulator of lymphocyte activation	Yes	1.7373	14.0865	1.5719	0.8784	0.9798	0.0606	0.3524	0.6667
CDH1	192090	0	Cell–cell adhesion	Yes	1.4111	17.1376	1.3262	1.0327	1.0121	0.0303	0.2571	1.0000
ZNF183	7737	7	Development?	Yes	1.9128	13.8790	2.4289	0.8238	0.7143	0.1212	0.0381	0.0667
RBBP7	602922	6	Regulation of cell proliferation	Yes	1.4289	16.8321	3.0564	1.0066	1.1265	0.0606	0.0095	0.0167
CSE1L	601342	1	Apoptosis and in cell proliferation	Yes	1.3702	17.7123	1.4867	1.0503	0.8966	0.0606	0.7619	0.0167
PAM	170270	0	Modification of neuroendocrine peptides	–	–	–	0.8639	1.0666	1.5750	0.0303	0.6095	0.2667
SMARCA1	300012	2	Nucleosome remodeling	Yes	1.5283	16.3491	1.8097	1.3800	1.4315	0.0303	0.4762	0.6667
CTSS	116845	1	Degradation of antigenic proteins	Yes	1.5768	15.6818	1.5858	1.1328	0.8078	0.0303	0.2571	0.2667
BIRC2	601712	0	Apoptosis inhibition	–	–	–	0.6195	1.0349	0.6610	0.6061	0.1714	0.0167
FMR2	309548	3	Fragile mental retardation	Yes	1.3749	17.7123	2.0643	1.2271	0.9668	0.0303	0.0381	0.3833
CDC10	603151	3	Structural component of chromatin	Yes	1.2584	18.4945	1.7257	1.1429	1.0002	0.0303	0.0095	0.1833
AGTR1	106165	1	Vasopressor hormone	–	–	–	4.8373	7.4641	14.5442	0.2727	0.4762	0.0167
MYCN	164840	0	Transcription factor in neuroblastoma	–	–	–	1.1708	1.0581	1.1792	0.7576	0.7619	0.5167
FHL1	300163	0	Cell growth and differentiation	–	–	–	0.9210	1.0814	0.5950	0.1212	0.3524	0.0667
S100A10	114085	0	Exocytosis and endocytosis	–	–	–	1.2338	0.7498	0.9207	0.0303	0.0381	0.2667
CLK2	602989	9	Cell cycle, apoptosis, and telomere length	Yes	3.4426	6.7197	2.3075	0.8901	0.8097	0.4848	0.9143	0.3833
LHCGR	152790	10	Hormone receptor	Yes	2.5308	12.1831	7.5432	1.6441	1.1766	0.0303	0.0667	0.2667
CYPD	601753	8	Protective role against apoptosis	Yes	1.8827	13.8790	4.0938	1.7737	1.3688	0.0606	0.1714	0.0167

Table 2 The majority of

Gene	Name	Location	Bisulfite sequencing
NET1	Neuroepithelial cell transforming gene 1	10p15	Methylated
IGFBP2	Insulin-like growth factor binding protein 2	2q33-q34	Methylated
CSE1L	Chromosome segregation 1-like	20q13	Methylated
CDC10	Yeast CDC10 homolog	7p14.3-p14.1	Unmethylated
PAM	Peptidylglycine alpha-amidating monooxigenase	5q14-21	Unmethylated
BIRC2	Baculoviral IAP repeat-containing 2	11q22	Unmethylated
IGF2	Insulin-like growth factor 2	11p15.5 (Imp.)	Methylated
PGK1	Phosphoglycerate kinase 1	Xq13	Methylated
FHL1	LIM and zinc finger domain protein	Xq26	Methylated
MPP1	Palmitoylated membrane protein 1	Xq28	Methylated

overexpressed genes that are direct *MECP2* targets are methylated

neurite outgrowth (Strausberg et al. 2002); the palmitoylated membrane protein 1 (*MPP1*), a membrane-associated guanylate kinase that mediate the anchoring of proteins at synapses (Caruana 2002), which was up-regulated in seven out of ten cell lines; and the gene *IGFBP2*, encoding the insulin-like growth factor-binding protein 2 (see Table 1). Interestingly, we found that some of the selected genes, were previously known to be silenced in normal cells as a result of CpG island methylation. This group included imprinted genes (*IGF2*) or X-chromosome genes (*PGK1, FHL1*, etc) and genes representative of the tissue type used in the model system (lymphocytes); for example, the protein tyrosine phosphatase receptor type C-associated protein (*PTPRCAP*), a key regulator of T- and B-lymphocyte activation (Schraven et al. 1994). Also included were genes involved in chromosome condensation and relieve the torsional stress that occurs during DNA transcription and replication, both processes directly linked to DNA methylation (Cobb et al. 1999), such as the DNA topoisomerase II alpha (*TOP2A*) (Lang et al. 1998).

Once an aberrant expression profile for RTT patients was defined, we undertook the task to elucidate the underlying mechanism involved. Two possible scenarios exist: the overexpression of a particular gene in RTT-derived samples can be interpreted as the result of the loss of association of MeCP2 or, alternatively, the overexpression of a regulatory protein due to the loss of control by MeCP2 that leads to the disregulation of a number of downstream targets. It is, therefore, a key issue to distinguish between genes directly and indirectly regulated by MeCP2 in the context of RTT and micro-array analyses provide a powerful source of candidate genes. To this end, we performed ChIP analysis (Orlando 2000). This assay can determine whether a particular transcriptional factor (in this case MeCP2) is bound to a specific promoter in its natural chromatin state (in this case the CpG islands of the genes overexpressed in our microarray analysis). In this regard, we have already used ChIP analysis to determine that MeCP2 binds to the methylated sequences of the imprinted gene *U2af1-rs1* (Fournier et al. 2002) or several tumor suppressor genes (Ballestar et al. 2003).

First, we immunoprecipitated DNA from the normal female lymphoblastoid cell lines with two different MeCP2 antibodies (indicated in Materials and methods) and total 5-methylcytosine DNA content was analyzed by high performance capillary electrophoresis (Fraga and Esteller 2002). We observed an average two- to threefold enrichment in 5-methylcytosine DNA in this MeCP2-ChIP DNA versus the input-control DNA (data not shown), supporting the notion that MeCP2 is associated in vivo with methylated DNA sequences.

Second, we performed PCR amplification reactions for the promoter of the 30 genes shown in Table 1 using this MeCP2-immunoprecipitated DNA. The presence of a PCR band for that particular DNA region indicates binding of MeCP2 and its absence means the lack of any MeCP2 binding. Our MeCP2-ChIP analysis of the genes upregulated in RTT cell lines allowed to distinguish two groups: those genes where MeCP2 is bound in normal cells, while MeCP2 is lost in RTT cells, leading to the release of gene silencing (such as the neurodevelopmental genes *NET1* or *MPP1*); and genes that are not directly regulated by the action of MeCP2 over their promoters (such as *GNAS* or *NMYC*). The first group was composed of ten genes (see Fig. 1a and Table 2) and the second consisting of the remaining 20 genes. In the latter case, an indirect effect is implied, such as the disregulation of other transcription factors that are themselves directly regulated by MeCP2. The ten genes whose direct interaction with MeCP2 was demonstrated by ChIP assays are shown in Table 2. Interestingly, when we performed ChIP assays with the cells derived from a patient harboring mutations in the MBD of MeCP2 (T158M) we observed a total loss of association of MeCP2 in four out of the ten genes (*IGFBP2, CSE1L, PGK1, IGF2*) and a significant decrease of association in three out of the ten genes (*MPP1, FHL1, NET1*) (Fig. 4a). The existence of partial association of MeCP2 in some of the genes can be explained considering that there is still a significant fraction of the wild-type allele that can be expressed in T158M cells (as deduced from the HUMARA analysis shown in Fig. 2). Curiously enough CDC10, PAM and BIRC2 did not exhibit any significant difference with respect to MeCP2 association regardless of the presence of a mutation in the MBD of MeCP2 (this issue will be discussed below).

In order to investigate whether other MBDs were targeted to these MeCP2 target genes, we also performed

ChIP assays with antibodies against the remaining members of the MBD family for which a role in transcriptional repression has been reported (*MBD1, MBD2 and MBD3*). We found that in some cases, additional MBD proteins were bound to a few promoters among these genes (Fig. 4a). In particular, we found binding of other MBD proteins to the promoter of the X-chromosome genes *FHL1*, *MPP1* and *PGK1* and the imprinted gene *IGF2* (in agreement with data reported in references Lang et al. 1998 and Martinowich et al. 2003). We also found binding of other MBDs to the promoter of *NET1* (see Fig. 4a). Loss of MeCP2 recruitment due to mutations in the MBD encoding region of MeCP2 did not affect in general the recruitment of MBD1, MBD2

and MBD3 to these genes. Despite the presence of additional MBDs in some cases, overexpression of those genes upon loss of functional MeCP2 can still occur as part of the repressing transcriptional machinery has been lost.

The current model proposes that MeCP2 mainly represses transcription through interaction with methylated promoters. If this assumption proves to be correct, the ten upregulated genes for which interaction with MeCP2 has been demonstrated should be methylated. Therefore, we investigated the promoter-associated CpG island methylation status of all ten genes in normal lymphoblastoid cells using bisulfite genomic sequencing (Clark and Warnecke 2002) (Fig. 5). The majority of genes (seven out of ten) was methylated at these regions supporting the concept of MeCP2 methylation-mediated gene repression. In some genes, such as *MPP1*, *FHL1* and *PGK1*, the presence of methylation was highly likely because these genes are located in the X-chromosome, and all cells analyzed were from females that undergo inactivation of one X-chromosome by methylation. The same result was also expected for the *IGF2* gene due to its well characterized imprinting (Reik et al. 2000). For both X-chromosome and imprinted genes, the CpG dinucleotide methylation had a 50% value, as expected due to the presence of one methylated allele and one unmethylated allele (Fig. 5). Most interestingly, when we performed bisulfite genomic sequencing of the DNA immunoprecipitated with the MeCP2 antibody, CpG dimethylation reached 100% (see *PGK1* in Fig. 5). Focusing again in the target genes, most interesting was the finding that there was methylation at the 5′-regulatory regions of the other three non-X-chromosome non-imprinted genes: *CSEL1*, *NET1*, and *IGFBP2*. It is noteworthy that the CpG methylation pattern observed for these genes was different from the one observed in the X-chromosome or imprinted genes: while these latter genes show dense methylation across the whole CpG island, *CSEL1*, *NET1* and *IGFBP2* show discrete sequences embedded in the CpG island where the hypermethylated CpG dinucleotides were clustered (Fig. 5). The lack of methylation in the three remaining genes (*CDC10, PAM and BIRC2*) is consistent with the absence of differences in association of MeCP2 between normal and RTT-derived cells harboring the T158M mutation (see Fig. 4a). We also performed bisulfite sequencing of the above genes in some of the lymphoblastoid cells from RTT patients, in order to discard the possibility of demethylation events in these cell lines. Identical patterns of methylation were observed in the three cell lines studied (R106W, T158M and R294X) as shown in Fig. 5 (where only data corresponding to the two normal controls and T158M are shown) and Table 2. A summary of all the methylation results, which are identical in normal and RTT-derived lymphoblastoid cells, are shown in Table 2.

On the other hand, the functional relevance of the MeCP2-mediated gene silencing of direct target genes was investigated by RT-PCR and a comparison between
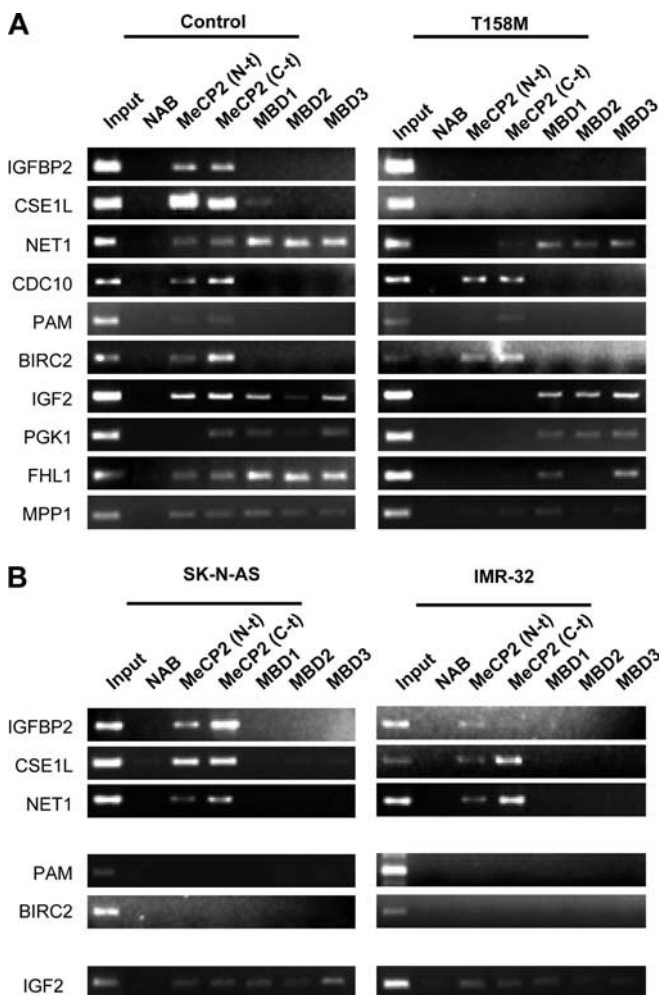

Fig. 4 a Rett mutations in the MBD of MeCP2 involve loss of recruitment to several overexpressed promoters. Chromatin immunoprecipitation (*ChIP*) analysis of the occupancy by MeCP2 and other MBD proteins of several promoters obtained from the microarray expression studies. Input fraction is designed as I. Only the 'bound' fractions corresponding to MeCP2 and the other MBDs are shown. The bound fraction of the 'no antibody' (*NAB*) control is also shown. ChIP assays were performed for control cell lines and a cell lines harboring a mutation at T158M. **b** ChIP assays with two neuroblastoma lines SK-N-AS and IMR-32

Fig. 5 DNA methylation analysis of the particular genes for which interaction with MeCP2 was demonstrated. Bisulfite genomic sequencing of the CpG island of the target genes: a fragment of the sequence is shown. Unmethylated Cs become Ts upon bisulfite modification. A schematic representation of some of the CpG sites included in the PCR fragment is shown. CpG sites are represented as *circles*: *black* (100% methylation), *gray* (50% methylation) and *white* (0% methylation). The samples correspond to two control lines (C1 and C2), the T158M cell line and two neuroblastoma cell lines SK-N-AS and IMR-32
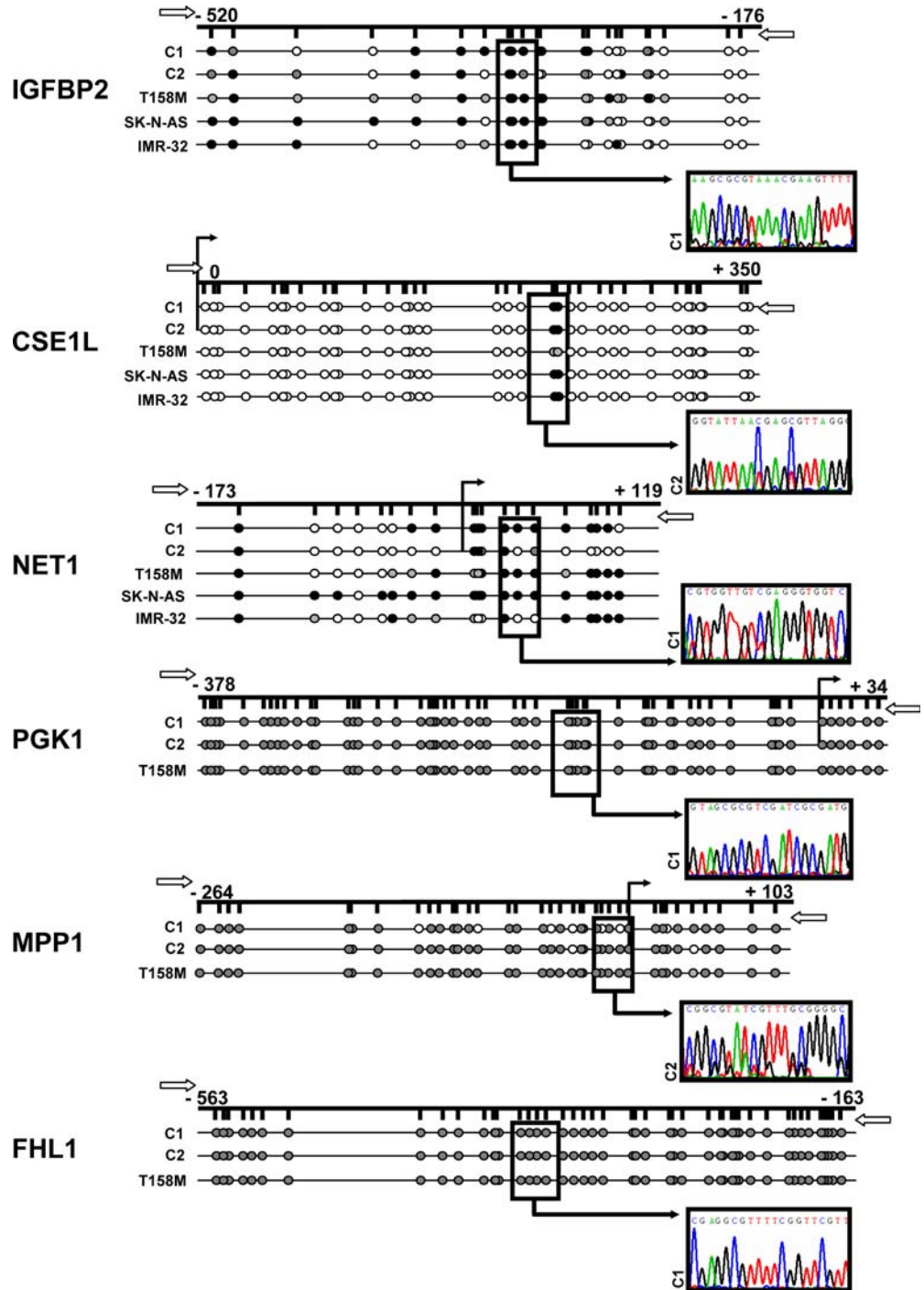
control and RTT-derived cells shows elevated expression levels of the above genes in the RTT patients (Fig. 6a). Moreover, the use of the demethylating agent 5-aza-2-deoxycytidine (5-aza-dC) in the control lymphoblastoid cells induced expression many of these genes (*FHL1, NET1, CSE1L, IGFBP2*), further supporting the role of DNA methylation in the silencing of these MeCP2-target genes (Fig. 6a). Expression of unmethylated gene *BIRC2* was not stimulated by 5-aza-dC as expected.

However, in the case of the unmethylated gene *CDC10* a small but significant overexpression after 5-aza-dC treatment was observed (Fig. 6a). This result might be due to indirect effects resulting of altered levels of transcription factors that are directly affected by 5-aza-dC. The lack of 5-aza-dC dependent stimulation of expression of some X-chromosome genes (*MPP1* and *PGK1*) may suggest additional chromatin mechanisms involved in the regulation of these genes.

As a next step to identify misregulated target genes in RTT patients, we decided to move from the lymphocytes to neuroblasts. Although our original cell type was not directly relevant to the brain phenotype in RTT, it is possible not only that those genes are misregulated in lymphocytes in RTT but also that these genes may also be misregulated in neuron-related tissue. Therefore, we investigated this possibility by using a neuron-related model: two neuroblastoma cell lines, SK-N-AS and

Fig. 6a,b Expression analysis monitored by RT-PCR of the MeCP2-target genes. *GAPDH* (*bottom panel*) was used as a control. All the genes, including *GAPDH*, produced a PCR product of about 400 bp. **a** RT-PCR of the two control (C1 and C2) and three RTT (R106W, T158M, R294X) samples. The control samples (C1 and C2) were additionally treated with 5-aza-dC. **b** RT-PCR for the SK-N-AS cell line in the absence (control) and presence of 5-aza-dC. Three illustrative examples are shown: *NET1* and *IGF2* are methylated and *CDC10* is unmethylated

IMR-32 . The same issues analyzed in the lymphocytes were addressed in these cells: (1) recruitment of MeCP2 and other MBD proteins to the above genes by using ChIP assays, (2) the methylation status of these genes by bisulfite sequencing, and (3) expression levels of these genes before and after treatment with 5-aza-dC.

Firstly, MeCP2 was found to be associated with the promoter of the autosomal genes *NET1*, *IGFBP2* and *CSE1L* (Fig. 4b). In contrast, MeCP2 did not appear to be associated with *CDC10*, *PAM* and *BIRC2*. When we performed bisulfite sequencing of these sequences, we observed that this association was methylation-dependent, since only NET1, IGFBP2 and CSE1L appeared to be methylated (Fig. 5 and Table 2). Finally, RT-PCR analysis of these genes indicated that treatment with the demethylating agent 5-aza-dC was able to increase their expression levels (Fig. 6b).
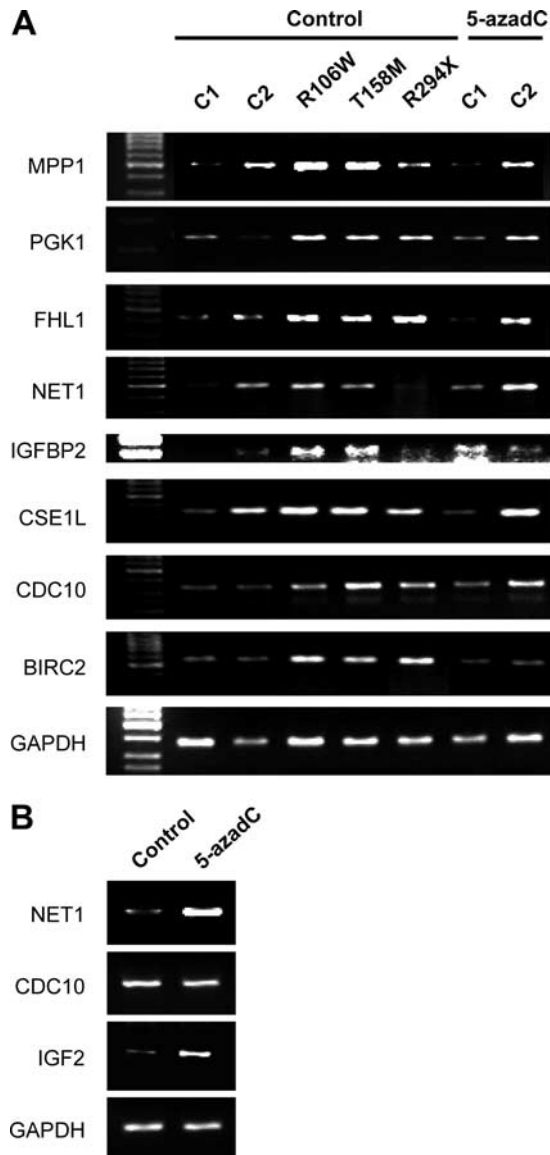
Discussion

Mutations in the *MECP2* gene are associated with RTT (Amir et al. 1999). Previous research has shown that MeCP2 is a methyl CpG binding protein that acts as a transcriptional repressor by preventing unscheduled transcription through the recruitment of histone de-acetylases and other corepressors to methylated DNA in the regulatory and promoter regions of genes (Jones et al. 1998; Nan et al. 1998). The answer to the question as to whether this function is phenotypically relevant to the phenotype of RTT patients now depends on the research progress in determining the degree and specificity of the disturbance of gene expression in RTT patients, followed by the confirmation of bona fide target genes for MeCP2 in normal cells.

In the present study, we have addressed these questions in a two-step manner. First, we have investigated the changes in the expression levels of different RTT-derived cell lines when compared with cell lines derived from normal individuals by a cDNA microarray approach. Secondly, the overexpressed genes identified by microarray analysis were then divided by ChIP assay into direct MeCP2 targets and genes whose expression is indirectly affected by MeCP2.

We have observed the existence of a specific aberrant gene expression signature in RTT patients compared to normal females. Moreover, clustering analysis was able to determine a distinct profile of expression according to the type of MeCP2 mutation (TRD and C-terminal region mutations versus MBD mutations) and also for those patients with clinically diagnosed RTT but without mutations in their coding regions. From the quantitative standpoint, the overall number of genes that are significantly deregulated in RTT-derived samples is relatively small, a finding that is in agreement with other studies conducted in humans (Colantuoni et al. 2001; Traynor et al. 2002) as well as data derived from MeCP2 knockout mouse brain tissues (Tudor et al. 2002). Various explanations may be proposed to explain

the absence of a massive effect over the entire transcriptome, including the existence of possible redundancy between the different members of the MBD family of proteins (MeCP2, MBD1, MBD2, MBD3) that may compensate the effects of MeCP2 loss. This notion could be supported by the fact that several MBDs can simultaneously occupy the same methylated promoter (Fournier et al. 2002; Koizume et al. 2002; Ballestar et al. 2003). In fact, we show that additional MBD proteins are recruited to some of the genes for which MeCP2 binding was demonstrated (Fig. 4b). Alternatively, it is possible that spatial and temporal factors play an essential role in determining major expression changes in RTT patients: for example, the deregulation of a cluster of genes expressed only in certain parts of the brain during a precise developmental phase may be crucial for the emergence of the phenotypical abnormalities observed in RTT patients. In our approach, it is worth noting that we have identified several genes, such as *NET1* and *MPP1*, which are important for neuronal homeostasis (Chan et al. 1996; Schmidt et al. 2002) and relieved from MeCP2-mediated transcriptional silencing in RTT patients.

To date, identification of MeCP2 targets has relied on a candidate gene approach, most of which are tumor suppressor genes hypermethylated in cancer (Nguyen et al. 2001; Koizume et al. 2002). These targets are, however, unlikely to be relevant to the RTT phenotype due to the aberrant DNA methylation profile of cancer cells compared with normal tissues (Esteller 2002). Recently, when the candidate gene strategy was applied to the *Xenopus laevis* model, a bona fide candidate target for MeCP2 in the context of RTT was identified in differentiating neuroectoderm: the neuronal repressor *xHairy2a* (Stancheva et al. 2003). Our microarray analysis unveils a comprehensive list of candidate genes for further investigation into their association with MeCP2 in normal cells, where the DNA methylation patterns have not been disturbed. Most importantly, our approach is validated by the finding that several previously reported MeCP2 target genes disregulated in RTT patients and mouse models (Colantuoni et al. 2001; Traynor et al. 2002; Tudor et al. 2002) also showed a significant misregulation in our RTT patients, such as the *TPD52L2*, *CRSP7* and AI216628 genes that were upregulated and the *TCEB2*, *RPL31*, *MAP3K14* and AI285186 genes that were downregulated. In our study, in addition to identifying new disregulated target genes, the combination of cDNA microarray with ChIP analysis, provides an strategy for distinguishing between direct and indirect targets of MeCP2 in RTT-derived samples.

We have demonstrated the presence of overexpressed genes in RTT-derived cells that exhibit release of MeCP2 transcriptional repression over methylated gene promoters. However, the exceptions to this general rule are worth noting. For example, a minority (30%) of the genes that had MeCP2 associated with their promoters were found to be unmethylated in our first screening. A technical explanation would be that, in these genes, MeCP2 is indeed binding to a methylated sequence, but outside of the region amplified in the bisulfite genomic PCR sequencing. However, since MeCP2 appears to be associated to its promoter in the mutant T158M form, an alternative explanation comes from the fact that MeCP2 may also be targeted to non-methylated sequences. In fact, it has been recently demonstrated that MeCP2 assembles novel secondary chromatin structures independent of the DNA methylation status (Georgel et al. 2003) and it has also been shown to bind unmethylated CpG islands in vitro, although with a lower affinity to that for methylated CpGs (Fraga et al. 2003). On the other hand, recent reports (Chen et al. 2003; Martinowich et al. 2003) have demonstrated that MeCP2 activity can be modulated by calcium-dependent phosphorylation that depends on membrane depolarizations. This connection between MeCP2 activity and signaling pathways could operate in other cell types and future studies will evaluate the contribution of these modifications for the regulation of each of the MeCP2 targets. For the three unmethylated genes that exhibit MeCP2 association, even when the MBD is mutated (see Fig. 4), it is likely that additional mechanisms of transcriptional regulation that are altered in RTT cells might be participating in their overexpression. Finally, it is important to take note of those genes overexpressed in RTT-derived samples (some of which possess clear neuron-related functions) where direct binding of MeCP2 to their promoters was not observed. Again, a technical explanation would be that MeCP2 is bound to another regulatory region outside the range of the ChIP primers used. However, a "biological" explanation is equally possible: their overexpression is an indirect effect that may reflect the direct release of MeCP2 from the promoter of an activating transcription factor. Here, we can cite the example of the gene encoding the transcriptional activator *FHL1*, where we have demonstrated that in RTT-derived cells MeCP2 is released from the *FHL1* methylated promoter leading to its overexpression. *FHL1* is a member of the LIM domain family of proteins, which are involved in the regulation of many neuronal genes, among them the cadherin and integrin genes that were found to be overexpressed in our RTT-derived cells (Bazan et al. 2002; Allan and Thor 2003).

In summary, we have demonstrated that RTT cells present a distinctive profile of gene expression that distinguishes them from corresponding normal cells. These changes are both quantitative in that they represent a trend for gene overexpression and qualitative through the presence of direct and indirect MeCP2-mediated mechanisms of abrogation of gene silencing. In this context, our results provide in vivo mechanistic evidence of the loss of MeCP2-mediated transcriptional repression in RTT patients and a comprehensive view of the aberrant expression defects underlying the RTT phenotype.

References

Alaminos M, Mora J, Cheung NK, Smith A, Qin J, Chen L, Gerald WL (2003) Genome-wide analysis of gene expression associated with MYCN in human neuroblastoma. Cancer Res 63:4538–4546

Allan DW, Thor S (2003) Together at last: bHLH and LIM-HD regulators cooperate to specify motor neurons. Neuron 38:675–677

Allen RC, Zoghbi HY, Moseley AB, Rosenblatt HM, Belmont JW (1992) Methylation of *Hpa*II and *Hha*I sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. Am J Hum Genet 51:1229–1239

Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY (1999) Rett syndrome is caused by mutations in X-linked MeCP2, encoding methyl-CpG-binding protein 2. Nat Genet 23:185–188

Armstrong DD (1997) Review of Rett syndrome. J Neuropathol Exp Neurol 56:843–849

Ballestar E, Yusufzai TM, Wolffe AP (2000) Effects of Rett syndrome mutations of the methyl-CpG binding domain of the transcriptional repressor MeCP2 on selectivity for association with methylated DNA. Biochemistry 20:7100–7106

Ballestar E, Paz MF, Valle L, Wei S, Fraga MF, Espada J, Cigudosa JC, Huang THM, Esteller M (2003) Methyl-CpG binding proteins identify novel sites of epigenetic activation in human cancer. EMBO J 22:6335–6345

Bazan NG, Palacios-Pelaez R, Lukiw WJ (2002) Hypoxia signaling to genes: significance in Alzheimer's disease. Mol Neurobiol 26:269–281

Bird A, Tweedie S (1995) Transcriptional noise and the evolution of gene number. Philos Trans R Soc Lond B Biol Sci 349:249–253

Buschdorf JP, Stratling WH (2004) A WW domain binding region in methyl-CpG-binding protein MeCP2: impact on Rett syndrome. J Mol Med 82:135–143

Caruana G (2002) Genetic studies define MAGUK proteins as regulators of epithelial cell polarity. Int J Dev Biol 46:511–518

Chan AM, Takai S, Yamada K, Miki T (1996) Isolation of a novel oncogene, *NET1*, from neuroepithelioma cells by expression cDNA cloning. Oncogene 12:1259–1266

Chen WG, Chang Q, Lin Y, Meissner A, West AE, Griffith EC, Jaenisch R, Greenberg ME (2003) Derepression of *BDNF* transcription involves calcium-dependent phosphorylation of MeCP2. Science 302:885–889

Clark SJ, Warnecke PM (2002) DNA methylation analysis in mammalian cells. Methods 27:99–100

Cobb J, Miyaike M, Kikuchi A, Handel MA (1999) Meiotic events at the centromeric heterochromatin: histone H3 phosphorylation, topoisomerase II alpha localization and chromosome condensation. Chromosoma 108:412–425

Colantuoni C, Jeon OH, Hyder K, Chenchik A, Khimani AH, Narayanan V, Hoffman EP, Kaufmann WE, Naidu S, Pevsner J (2001) Gene expression profiling in postmortem Rett sndrome brain: differential gene expression and patient classification. Neurobiol Dis 8:847–865

Dopazo J, Dopazo JM (1997) Phylogenetic reconstruction using an unsupervised growing neural network that adopts the topology of a phylogenetic tree. J Mol Evol 44:226–233

Esteller M (2000) Epigenetic lesions causing genetic lesions in human cancer: promoter hypermethylation of DNA repair genes. Eur J Cancer 36:2294–2300

Esteller M (2002) CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. Oncogene 21:5427–5440

Fournier C, Goto Y, Ballestar E, Delaval K, Hever A, Esteller M, Feil R (2002) Allele-specific histone lysine methylation marks regulatory regions at imprinted mouse genes. EMBO J 21:6560–6570

Fraga MF, Esteller M (2002) DNA methylation: a profile of methods and applications. Biotechniques 33:632–649

Fraga MF, Ballestar E, Montoya G, Taysavang P, Wade PA, Esteller M (2003) The affinity of different MBD proteins for a specific methylated locus depends on their intrinsic binding properties. Nucleic Acids Res 31:1765–1774

Georgel PT, Horowitz-Scherer RA, Adkins N, Woodcock CL, Wade PA, Hansen JC (2003) Chromatin compaction by human MeCP2. Assembly of novel secondary chromatin structures in the absence of DNA methylation. J Biol Chem 278:32181–32188

Herrero J, Diaz-Uriarte R, Dopazo J (2003) Gene expression data preprocessing. Bioinformatics 19:655–656

Hescheler J, Schultz G (1994) Heterotrimeric G proteins involved in the modulation of voltage-dependent calcium channels of neuroendocrine cells. Ann N Y Acad Sci 733:306–312

Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J, Wolffe AP (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat Genet 19:187–191

Kempkes B, Pich D, Zeidler R, Hammerschmidt W (1995) Immortalization of human primary B lymphocytes in vitro with DNA. Proc Natl Acad Sci USA 92:5875–5879

Koizume S, Tachibana K, Sekiya T, Hirohashi S, Shiraishi M (2002) Heterogeneity in the modification and involvement of chromatin components of the CpG island of the silenced human *CDH1* gene in cancer cells. Nucleic Acids Res 30:4770–4780

Kriaucionis S, Bird A (2003) DNA methylation and Rett syndrome. Hum Mol Genet 12 (Spec 2):R221–R227

Laccone F, Junemann I, Whatley S, Morgan R, Butler R, Huppke P, Ravine D (2004) Large deletions of the MeCP2 gene detected by gene dosage analysis in patients with Rett syndrome. Hum Mutat 23:234–244

Lang AJ, Mirski SE, Cummings HJ, Yu Q, Gerlach JH, Cole SP (1998) Structural organization of the human *TOP2A* and *TOP2B* genes. Gene 221:255–266

Martinowich K, Hattori D, Wu H, Fouse S, He F, Hu Y, Fan G, Sun YE (2003) DNA methylation-related chromatin remodeling in activity-dependent *BDNF* gene regulation. Science 302:890–893

Mnatzakanian GN, Lohi H, Munteanu I, Alfred SE, Yamada T, MacLeod PJ, Jones JR, Scherer SW, Schanen NC, Friez MJ, Vincent JB, Minassian BA (2004) A previously unidentified MeCP2 open reading frame defines a new protein isoform relevant to Rett syndrome. Nat Genet 36:339–341

Monros E, Armstrong J, Aibar E, Poo P, Canos I, Pineda M (2001) Rett syndrome in Spain: mutation analysis and clinical correlations. Brain Dev 23 (Suppl 1):S251–S253

Moreno-Bueno G, Sanchez-Estevez C, Cassia R, Rodriguez-Perales S, Diaz-Uriarte R, Dominguez O, Hardisson D, Andujar M, Prat J, Matias-Guiu X, Cigudosa JC, Palacios J (2003) Differential gene expression profile in endometrioid and non-endometrioid endometrial carcinoma: *STK15* is frequently overexpressed and amplified in nonendometrioid carcinomas. Cancer Res 63:5697–5702

Nan X, Bird A (2001) The biological functions of the methyl-CpG-binding protein MeCP2 and its implication in Rett syndrome. Brain Dev 23 (Suppl 1):S32–S37

Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 393:386–389

Nguyen CT, Gonzales FA, Jones PA (2001) Altered chromatin structure associated with methylation-induced gene silencing in cancer cells: correlation of accessibility, methylation, MeCP2 binding and acetylation. Nucleic Acids Res 29:4598–4606

Orlando V (2000) Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation. Trends Biochem Sci 25:99–104

Reik W, Constancia M, Dean W, Davies K, Bowden L, Murrell A, Feil R, Walter J, Kelsey G (2000) Igf2 imprinting in development and disease. Int J Dev Biol 44:145–150

Schmidt JT, Morgan P, Dowell N, Leu B (2002) Myosin light chain phosphorylation and growth cone motility. J Neurobiol 52:175–188

Schraven B, Schoenhaut D, Bruyns E, Koretzky G, Eckerskorn C, Wallich R, Kirchgessner H, Sakorafas P, Labkovsky B, Ratnofsky S et al (1994) LPAP, a novel 32-kDa phosphoprotein that interacts with CD45 in human lymphocytes. J Biol Chem 269:29102–29111

Stancheva I, Collins AL, Van den Veyver IB, Zoghbi H, Meehan RR (2003) A mutant form of MeCP2 protein associated with human Rett syndrome cannot be displaced from methylated DNA by Notch in Xenopus embryos. Mol Cell 12:425–432

Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF et al (2002) Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. Proc Natl Acad Sci USA 99:16899–16903

Sugie H, Sugie Y, Tsurui S, Ito M (1994) Phosphoglycerate kinase deficiency. Neurology 44:1364–1365

Tracey L, Villuendas R, Ortiz P, Dopazo A, Spiteri I, Lombardia L, Rodriguez-Peralto JL, Fernandez-Herrera J et al (2002) Identification of genes involved in resistance to interferon-alpha in cutaneous T-cell lymphoma. Am J Pathol 161:1825–1837

Traynor J, Agarwal P, Lazzeroni L, Francke U (2002) Gene expression patterns vary in clonal cell cultures from Rett syndrome females with eight different MeCP2 mutations. BMC Med Genet 3:12

Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, Botstein D, Altman RB (2001) Missing value estimation methods for DNA microarrays. Bioinformatics 17:520–525

Tudor M, Akbarian S, Chen RZ, Jaenisch R (2002) Transcriptional profiling of a mouse model for Rett syndrome reveals subtle transcriptional changes in the brain. Proc Natl Acad Sci USA 99:5536–5541

Tusher V, Tibshirani R, Chu C (2001) Significance analysis of microarrays applied to ionising radiation response. Proc Natl Acad Sci USA 98:5116–5121

Vacca M, Filippini F, Budillon A, Rossi V, Della Ragione F, De Bonis ML, Mercadante G, Manzati E, Gualandi F, Bigoni S, Trabanelli C, Pini G, Calzolari E, Ferlini A, Meloni I, Hayek G, Zappella M, Renieri A, D'Urso M, D'Esposito M, Macdonald F, Kerr A, Dhanjal S, Hulten M (2001) MECP2 gene mutation analysis in the British and Italian Rett Syndrome patients: hot spot map of the most recurrent mutations and bioinformatic analysis of a new MeCP2 conserved region. Brain Dev 23 (Suppl 1): S246–S250

Van den Veyver IB, Zoghbi HY (2000) Methyl-CpG-binding protein 2 mutations in Rett syndrome. Curr Opin Genet Dev 10:275–279

Webb T, Latif F (2001) Rett syndrome and the MECP2 gene. J Med Genet 38:217–223

Willard H, Hendrich BD (1999) Breaking the silence in Rett syndrome. Nat Genet 23:127–128

Yusufzai TM, Wolffe AP (2000) Functional consequences of Rett syndrome mutations on human MeCP2. Nucleic Acids Res 28:4172–4179