

RESEARCH

Open Access



# Identification of molecular signatures and pathways involved in Rett syndrome using a multi-omics approach

Ainhoa Pascual-Alonso<sup>1,2†</sup>, Clara Xiol<sup>1,2†</sup>, Dmitrii Smirnov<sup>3,4</sup>, Robert Kopajtich<sup>3,4</sup>, Holger Prokisch<sup>3,4</sup> and Judith Armstrong<sup>2,5,6\*</sup>

## Abstract

**Background** Rett syndrome (RTT) is a neurodevelopmental disorder mainly caused by mutations in the methyl-CpG-binding protein 2 gene (*MECP2*). MeCP2 is a multi-functional protein involved in many cellular processes, but the mechanisms by which its dysfunction causes disease are not fully understood. The duplication of the *MECP2* gene causes a distinct disorder called *MECP2* duplication syndrome (MDS), highlighting the importance of tightly regulating its dosage for proper cellular function. Additionally, some patients with mutations in genes other than *MECP2* exhibit phenotypic similarities with RTT, indicating that these genes may also play a role in similar cellular functions. The purpose of this study was to characterise the molecular alterations in patients with RTT in order to identify potential biomarkers or therapeutic targets for this disorder.

**Methods** We used a combination of transcriptomics (RNAseq) and proteomics (TMT mass spectrometry) to characterise the expression patterns in fibroblast cell lines from 22 patients with RTT and detected mutation in *MECP2*, 15 patients with MDS, 12 patients with RTT-like phenotypes and 13 healthy controls. Transcriptomics and proteomics data were used to identify differentially expressed genes at both RNA and protein levels, which were further inspected via enrichment and upstream regulator analyses and compared to find shared features in patients with RTT.

**Results** We identified molecular alterations in cellular functions and pathways that may contribute to the disease phenotype in patients with RTT, such as deregulated cytoskeletal components, vesicular transport elements, ribosomal subunits and mRNA processing machinery. We also compared RTT expression profiles with those of MDS seeking changes in opposite directions that could lead to the identification of MeCP2 direct targets. Some of the deregulated transcripts and proteins were consistently affected in patients with RTT-like phenotypes, revealing potentially relevant molecular processes in patients with overlapping traits and different genetic aetiology.

**Conclusions** The integration of data in a multi-omics analysis has helped to interpret the molecular consequences of *MECP2* dysfunction, contributing to the characterisation of the molecular landscape in patients with RTT. The comparison with MDS provides knowledge of MeCP2 direct targets, whilst the correlation with RTT-like phenotypes highlights processes potentially contributing to the pathomechanism leading these disorders.

<sup>†</sup>Ainhoa Pascual-Alonso and Clara Xiol have contributed equally to this work.

\*Correspondence:

Judith Armstrong

judith.armstrong@sjd.es

Full list of author information is available at the end of the article



**Keywords** Rett syndrome, MECP2 duplication syndrome, Rett-like phenotypes, Multi-omics, Transcriptomics, Proteomics, Differential expression

## Background

Rett syndrome (RTT, OMIM#312750) is a severe neurodevelopmental disorder characterised by psychomotor regression after a period of normal early development. It mainly affects girls, who typically present with loss of purposeful hand use and expressive language, gait abnormalities and stereotypic hand movements. In addition, the main symptoms can be accompanied by a variety of other dysfunctions, such as seizures, breathing disturbances, scoliosis, impaired sleep patterns and abnormal muscle tone [1]. The diagnosis of RTT is mainly clinical and is based on a set of criteria that differentiate between typical RTT and three atypical variants with distinctive features: the preserved speech variant, the congenital variant and the early-onset seizure variant [2].

Most typical and atypical RTT cases are caused by loss-of-function mutations in the methyl-CpG-binding protein 2 (*MECP2*, OMIM\*300005) gene, located on the X chromosome [3]. MeCP2 is a chromatin-associated protein that acts as a transcriptional regulator, both repressing and activating transcription, and is also involved in maintaining heterochromatin structure, regulating splicing through interaction with splicing factors and miRNA processing by binding to microprocessor components [1]. MeCP2 is expressed ubiquitously, but is especially abundant in mature neurons. MeCP2 has proved to be crucial for neuronal maturation, dendritic arborisation and synaptic plasticity [1].

Mutations in other genes have also been identified in patients with RTT. Pathogenic variants in cyclin-dependent kinase-like 5 (*CDKL5*, OMIM\*300203) and forkhead box G1 (*FOXG1*, OMIM\*164874) have been detected in a substantial number of cases with early-onset seizure and congenital RTT variants, respectively [4, 5]. Moreover, with the generalisation of next-generation sequencing (NGS), the number of genes associated with RTT has increased remarkably [6–8]. Some of these are novel findings whilst others have already been associated with different neurodevelopmental disorders or epileptic encephalopathies. Patients with overlapping phenotypes with RTT but who do not fulfil established clinical criteria are termed 'RTT-like'. Therefore, any patient with a combination of RTT features can be described as RTT-like [6, 7].

MeCP2 levels are tightly regulated in the cells and not only a loss of function can have pathogenic effects. Chromosomal duplications at Xq28 encompassing the *MECP2* and *IRAK1* genes, leading to their gain of function,

cause *MECP2* duplication syndrome (MDS), a neurological disorder characterised by intellectual disability (ID), infantile hypotonia, seizures, speech impairment and recurrent respiratory infections [9]. It mainly affects males, whilst penetrance in females is highly dependent on X-chromosome inactivation (XCI). Phenotypic variability is high in patients with MDS and potentially related to the size and content of the duplication, which is unique for each family [10]. However, a clear genotype–phenotype correlation has not yet been found.

One of the drawbacks in studying the downstream molecular effects of MeCP2 dysfunction is the lack of accessibility to samples of the primarily affected tissue, the brain. In the search for new tissues, skin fibroblasts have demonstrated greater consistency in gene expression and include more OMIM and neurologically relevant genes compared with whole blood [11, 12].

Around 70 experimental and repurposed drugs have been investigated for RTT, but there is no approved treatment yet [13]. In RTT clinical trials, the success of the tested drugs is evaluated by measuring the improvement in the symptomatology and quality of life of the patients. The lack of a biomarker complicates an objective quantification of the improvements derived from drug treatments. An efficient way to extract huge amounts of molecular data in order to find biomarkers could be by analysing the RNA profiles and proteome of the patients using multi-omics technology.

To date, no multi-omics analysis has been performed with RTT human samples and only one has been published with 4 RTT mice samples [14]. Here, we aim to fill that knowledge gap by studying a cohort of 22 patients with RTT, 12 patients with RTT-like and 15 patients with MDS. Integration of transcriptomics and proteomics data could be a promising approach to find new potential therapeutic targets and biomarkers.

## Material and methods

### Clinical and molecular characterisation

The study was approved by the Hospital Sant Joan de Déu (HSJD) ethical committee, Comitè d'Ètica d'Investigació Clínica-Fundació Sant Joan de Déu (CEIC; internal code: PIC-219-20). Sixty-two individuals (49 patients and 13 healthy age-matched controls) participated in this study and provided written informed consent. Patients were recruited after clinical and genetic confirmation of their pathology as described elsewhere [15]. Eleven out of the fifteen

MDS patients were described in Pascual-Alonso et al. [16], and the four new patients were characterised in the same way. We studied 22 patients with RTT and mutations in *MECP2* (21 females and 1 male); 15 male patients with MDS; 12 patients with RTT-like phenotypes and mutations in *CDKL5* (1 female and 3 males), *FOXP1* (1 female and 1 male), *NR2F1* (1 female), *GRIN2B* (1 female), *AHDC1* (1 female) and 3 female patients without molecular diagnosis; and 13 healthy controls (7 females and 6 males) (Table 1). Clinical severity of patients with RTT and RTT-like phenotypes was measured using the clinical severity score designed by Dr Pineda [17].

Skin biopsies from the 62 individuals were obtained, and primary fibroblast cell lines were established. Fibroblast lines were grown on plates with Dulbecco's modified Eagle's medium high glucose with glutamine, supplemented with 10% heat-inactivated foetal bovine serum and 1% penicillin, streptomycin and B amphotericin (all from Thermo Fisher, Waltham, MA, USA). Cultures were kept at 37°C with 5% CO<sub>2</sub> in a humidified atmosphere. When the cells reached 70–80% confluence, they were trypsinised and either re-sowed on new plates or harvested for subsequent RNA or protein extraction. Frozen vials from all the fibroblast lines were entrusted to the Biobanc 'Hospital Infantil Sant Joan de Déu per a la Investigació', which is integrated into the Spanish Biobank Network of ISCIII for the sample and data procurement.

DNA was extracted from fibroblast cell lines using the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany)

**Table 1** Composition of the studied cohort, which consists of individuals with Rett syndrome (RTT) with mutations in *MECP2*, *MECP2* duplication syndrome (MDS) and Rett-like (RTT-like) with mutations in different genes that are not *MECP2* and healthy controls. 'Age' and 'Duration of disease' are given in years; Mean ± Standard Deviation with available data (Additional file 5: Table S2)

Individuals	Age	Duration of disease	Female	Male	Total
RTT	9 ± 6	12 ± 2	21	1	22
MDS	7 ± 6	6 ± 6	–	15	15
RTT-like	12 ± 7	11 ± 7	8	4	12
<i>CDKL5</i>	9.5 ± 10	9 ± 10.5	1	3	4
<i>FOXP1</i>	9.5 ± 3.5	9 ± 3.5	1	1	2
<i>NR2F1</i>	9	8.7	1	–	1
<i>GRIN2B</i>	9	8.7	1	–	1
<i>AHDC1</i>	13	11	1	–	1
Unknown mutation	18 ± 5	17 ± 5	3	–	3
Healthy controls	18 ± 14	–	7	6	13

according to manufacturer's instructions. XCI was performed in all female samples as described by Allen et al. [18]. XCI was considered skewed with an allele ratio of 80:20 or greater (Additional file 4: Table S1).

### RNA sequencing

RNA was extracted from cultured fibroblast pellets using the RNeasy Fibrous Tissue Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. The obtained RNA was then measured with a NanoDrop spectrophotometer and examined in an agarose gel to check its purity and integrity.

To further confirm the quality of the isolated RNA and to diminish undesirable gene alterations due to cell stress conditions [19], we performed reverse transcription-quantitative polymerase chain reaction (RT-qPCR) of five genes that are part of the oxidative respiratory chain: *MT-CO1*, *MT-CO2*, *MT-CYB*, *MT-ND4* and *MT-ATP6*. First, 500 ng of total RNA was processed according to the manufacturer's instructions, and double-stranded complementary DNA (cDNA) was generated in the presence of random hexamers using the SuperScript III First-Strand Synthesis SuperMix for qRT-PCR kit (Thermo Fisher, Waltham, MA, USA). Primers for the five mitochondrial genes and two additional housekeeping genes (*RPLP0* and *ALAS1*) were designed with Primer3 software [20] (Additional file 5: Table S2). The qRT-PCR was performed with PowerUp SYBR Green Master Mix (Thermo Fisher, Waltham, MA, USA) in an QuantStudio 6 Flex Real-Time PCR System (Applied Biosystems, Waltham, MA, USA). All reactions were conducted in triplicate and the average of each triplicate group was used for the analysis, which is based on the  $\Delta\Delta C_t$  relative quantification method. Three non-stressed control samples were added to the experiment to get the normalised values. Amplified product specificity was assessed via melting curve. All samples that overexpressed two or more genes more than 1.5-fold the values of non-stressed controls were discarded (Additional file 2: Fig. S2).

For each sample, 2500 ng of RNA was used for library preparation. Illumina's TruSeq Stranded mRNA kit (Illumina, San Diego, CA, USA) was used following the manufacturer's protocol. Libraries were quantified in a 4200 TapeStation (Agilent Technologies, Santa Clara, CA, USA) and their integrity was checked. Sequencing was performed on an Illumina NextSeq 500 sequencer and 75 base pair (bp) paired-end reads with around 40 million paired reads per sample successfully mapped to the reference genome. At least two healthy controls of the same sex as the patients were included in all the runs to enable normalisation and to control the batch effect.

To validate the RNAseq experiment, we chose 22 significantly differentially expressed genes (DEGs) and

performed RT-qPCR as explained above with a total of 23 different samples (Additional file 2: Fig. S2).

### Differential expression analysis

RNAseq reads were aligned to the human reference genome (GRCh37/hg19) using STAR (v2.4.2a) in a strand-specific manner. Uniquely mapped reads were counted for each gene using the HTSeq package (v2.0.2) [21] with gene models from GENCODE release 29. A final count matrix for analysis was generated by averaging the values of raw counts from different replicates of the same sample. Counts per million mapped reads (CPM) were computed and only genes where more than 50% of samples had at least 1 CPM were kept.

We first inspected age, sex, biopsy origin and batch as possible covariates in the differential expression study by principal component analysis (PCA) and cluster analysis, but found no clear patterns in our samples (data not shown). PCA identified the primary sources of variation in our data. The first three principal components, explaining 18.8%, 16.1% and 7.4% of the variance, were subsequently used in the model construction for differential expression analysis with DESeq2 (v1.34.0) [22]. We used a Benjamini–Hochberg (BH)-corrected p-value of 0.05 to consider significant differences.

### Enrichment and upstream regulation analysis

Enrichment analysis was performed using the clusterProfiler (v4.2.2) [23] and ReactomePA (v1.38.0) [24] R packages. Both overrepresentation analysis (ORA) and gene set enrichment analysis (GSEA) were carried out, using only significant DEGs and all expressed genes, respectively. Potentially enriched terms were searched in Gene Ontology (GO) [25], the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database [26], WikiPathways (WP) [27] and the Reactome pathway database (RP) [28]. All genes with CPM greater than 1 in at least 50% of samples and with an existing EntrezID were used as background (a total of 11,904 genes). The cut-off value for considering a significantly enriched term was 0.05 in BH-corrected p-value.

We considered upstream transcription factors (TFs) responsible for some of the differential expression changes observed in our data and used the ChIP-X Enrichment Analysis 3 (ChEA3) tool to identify them [29]. ChEA3 contains information about TF gene co-expression, association in ChIP-seq studies and co-occurrence in gene lists, which is used to predict upstream TFs involved in the regulation of the user inputted gene lists. The lists of DEGs resulting from differential expression analysis were fed to ChEA3 to predict the possible involvement of TFs in their dysregulation.

### Proteomics

Proteomics experiments were performed at the BayBioMS core facility at the Technical University of Munich (TUM) in Germany. Fibroblast cell pellets containing around 0.5 million cells were sent frozen. These cells were thawed and lysed with urea containing buffer and quantified using BCA Protein Assay Kit (Thermo Scientific, Waltham, MA, USA).

For the proteomics experiment, 15 µg of protein extract was reduced, alkylated and digested using Trypsin Gold (Promega, Madison, WI, USA). Digests were acidified, desalted and TMT-labelled following the protocol described by Zecha et al. [30] using the TMT 11-plex labelling reagent (Thermo Fisher, Waltham, MA, USA). TMT batches were organised to always include one reference sample that is common to all batches in order to enable normalisation. Liquid chromatography–mass spectrometry (LC–MS) measurements were run on a Fusion Lumos Tribrid mass spectrometer (Thermo Fisher, Waltham, MA, USA) operated in data-dependent acquisition mode and multi-notch MS3 mode. MaxQuant version 1.6.3.4 [31] was used for peptide identification, and protein groups were obtained. Missing values were imputed with the minimal value across the dataset [32].

### Differential expression in proteomics data

Prior to any analysis, MS data were adjusted with respect to one identical control sample that was present in each MS batch as described previously [32]. Recalibrated intensities were log-transformed for normalisation and proteins that were not detected in all samples were removed. An initial exploratory inspection of data by PCA and cluster analysis revealed that samples were grouped by MS batch (data not shown). Therefore, we carried out the differential expression analysis using the limma (v3.50.3) package [33] in R, including the MS batch as a covariate in the model to adjust for this confounding factor. We used the `removeBatchEffect` function from limma to recapitulate the exploratory analysis after batch correction and we observed no more clustering by MS batch. Finally, we took a nominal p-value of 0.05 as a threshold to define differentially expressed proteins (DEPs).

We carried out enrichment analysis just like we did for transcriptomics data. As a background, we considered the proteins that we detected in all samples with a valid EntrezID (a total of 5894 genes).



## Results

### Transcriptomic profiles in primary fibroblast cell cultures

First of all, we examined the similarity between the transcriptomic profiles obtained from primary fibroblast cell cultures and those from several brain areas, in order to understand how many of the molecular alterations that we identify could be extrapolated to neural tissues. We used publicly available data from the Genotype-Tissue Expression (GTEx) project and compared mean TPM (Transcripts per Kilobase Million) in fibroblast cultured cells and 11 brain areas: amygdala, anterior cingulate cortex, caudate basal ganglia, frontal cortex, cerebellar hemisphere, substantia nigra, hippocampus, hypothalamus, nucleus accumbens basal ganglia, putamen basal ganglia and spinal cord. 98.5% of detected transcripts (TPM > 0.5) in GTEx cultured fibroblasts RNAseq samples correspond to genes with some degree of expression in at least one neural tissue (Additional file 3: Fig. S3). More than 99% of the transcripts detected in our analysis are also reliably detected in GTEx cultured fibroblasts samples, indicating that the vast majority of the data that we are analysing may be extrapolated to biological processes occurring in the brain and therefore may impact neurological phenotypes.

### Characterisation of RTT-MECP2 versus controls

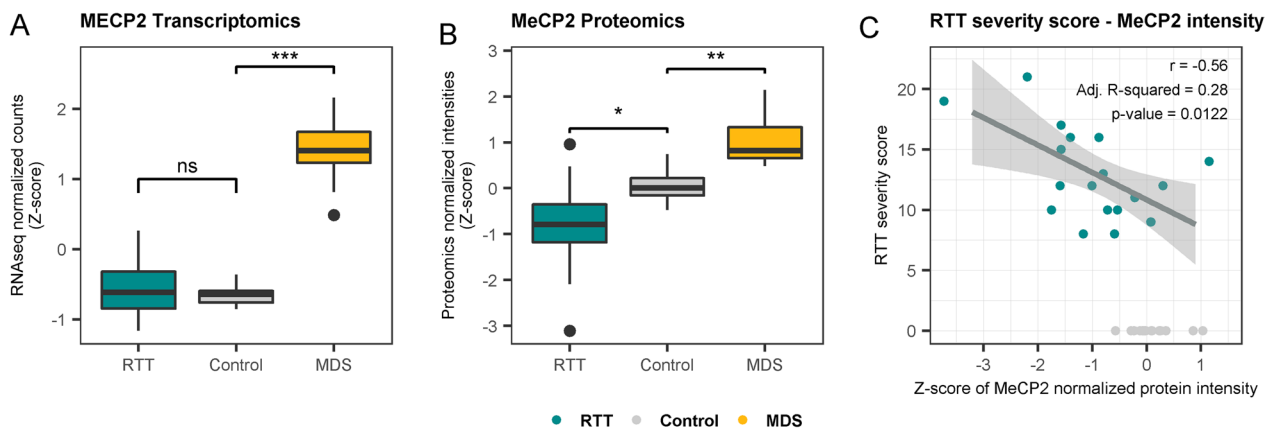
There were similar *MECP2* mRNA amounts in patients with RTT and controls, whereas MeCP2 protein amount was significantly reduced in patients with RTT (Fig. 1A, B). We found a significant correlation between MeCP2 levels and the Pineda clinical severity score of our patients with RTT, indicating that more severely affected patients present lower amounts of MeCP2 protein (Fig. 1C).

### Transcriptomics and upstream regulator analysis

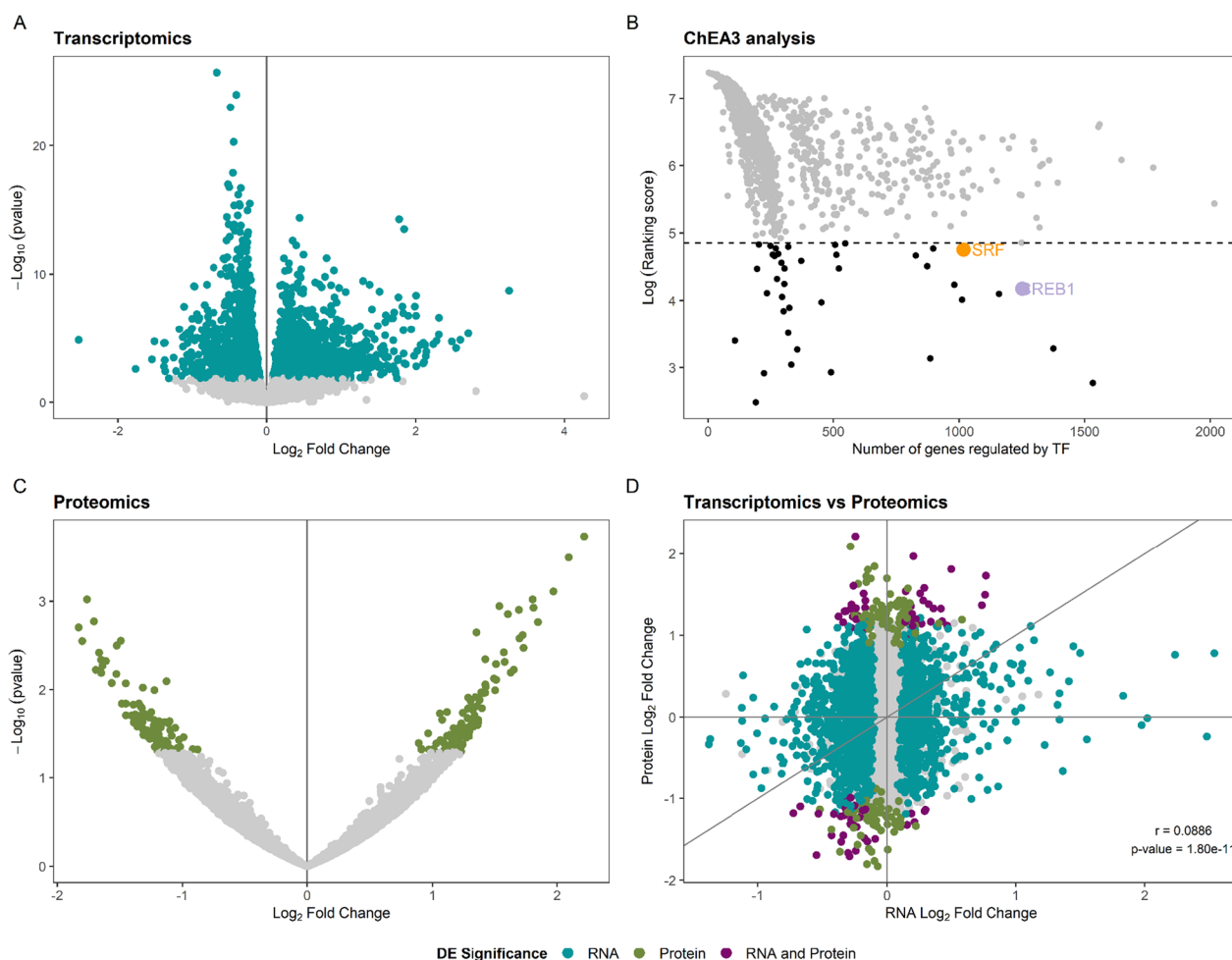
Differential expression analysis of patients with RTT carrying *MECP2* mutations versus healthy controls showed 3446 DEGs: 1713 upregulated and 1733 downregulated (Fig. 2A). We subsequently used these DEGs as input for upstream regulator analysis with ChEA3. We inspected the top 40 ranked TFs searching for proteins that regulate a large number of the identified DEGs, since they would potentially be driving some of these transcriptomic alterations. The list of DEGs was significantly enriched in *CREB1* and *SRF* targets (Fisher’s exact test  $p < 0.05$  in 5 of the 6 primary libraries in ChEA3). These two TFs have remarkable functions in neural tissues and could regulate the expression of 1253 and 1017 of the identified DEGs, respectively (Fig. 2B, Additional file 8: Table S5a). More than 98% of these potential targets have some degree of expression in at least one region of the nervous system, indicating that the alterations in transcriptomic networks identified in primary fibroblast cell cultures may affect the nervous system as well.

### Multi-omics: integrating transcriptomics and proteomics data

Proteomics differential expression analysis revealed 224 DEPs, 123 upregulated and 101 upregulated (Fig. 2C, Additional file 6: Table S3a,b). Thirty-three and 28 of these are *CREB1* and *SRF* targets, respectively. The number of DEPs is markedly lower than the number of DEGs identified in transcriptomics, probably in part due to the fact that mass spectrometry identified roughly half (5918) of the number of genes mapped in the RNAseq experiment (12,448). Almost 97% of the genes detected via mass spectrometry were also identified in RNAseq. Although the correlation between transcriptome and proteome differential expression findings was not high [Pearson



**Fig. 1** **A** MECP2 expression levels for RTT, MDS and control individuals obtained by RNAseq. **B** MECP2 intensity levels for RTT, MDS and control individuals obtained by proteomics. **C** Pearson’s correlation between the severity score of RTT patients and MeCP2 levels



**Fig. 2** Summary of the results of the RTT-MECP2 versus healthy controls analysis. **A** RNAseq DE analysis results. The coloured genes are considered differentially expressed, passing a threshold of BH < 0.05. **B** Upstream TF ChEA3 analysis for the DEGs. The 40 TFs that were further studied are coloured in black. **C** Proteomics DE analysis results. The coloured proteins are considered differentially expressed, passing a threshold of nominal *p* value < 0.05. **D** An integrated view of the transcriptomics and proteomics results. The genes that are significant at both analyses are coloured in purple

correlation coefficient = 0.09, *p* = 1.8e-11, Fig. 2D], we found 75 genes deregulated at both the RNA and protein levels in patients with RTT (Additional file 6: Table S3C). The overlap between DEGs and DEPs is not significantly higher than expected by chance (Fisher’s exact test *p* = 0.1397, OR = 1.18), but some of the concordant genes constitute strong candidates for deciphering some of the pathomechanisms behind RTT, as well as for establishing biomarkers of this disorder (Table 2).

Enrichment analysis uncovered significant overrepresentation of genes and proteins involved in several cellular functions and processes, some of which may be extrapolated to neuronal tissues and thus are especially interesting when trying to elucidate the pathomechanisms underlying RTT (Additional file 7: Table S4a,b).

The most remarkable pathways that repeatedly appeared significantly enriched with DEGs and DEPs were cytoskeletal processes, vesicular activity, rRNA processing and mRNA splicing (Fig. 3, Table 2). The vast majority of the consistent DEGs and DEPs driving this enrichment have some degree of expression in at least one brain area according to GTEx publicly available data.

**Patients with RTT versus patients with MDS**

We compared the results of the DE analysis performed with patients with classical RTT and patients with MDS to identify common gene expression dysregulations that could shed some light into the pathomechanism underlying both syndromes.

**Table 2** Genes with concordant differential expression in transcriptomics and proteomics that are involved in the main biological processes identified via enrichment analysis

Gene	Direction	Biological process	Potential TF
<i>AFAP1</i>	Upregulated	Cytoskeletal processes	SRF
<i>FMNL2</i>	Upregulated	Cytoskeletal processes	CREB1
<i>FNBP1L</i>	Upregulated	Cytoskeletal processes	CREB1
<i>KIF3A</i>	Upregulated	Cytoskeletal processes	–
<i>MARCKSL1</i>	Upregulated	Cytoskeletal processes	–
<i>PLS3</i>	Upregulated	Cytoskeletal processes	SRF
<i>ARMC9</i>	Downregulated	Cytoskeletal processes	SRF
<i>ARHGEF1</i>	Downregulated	Cytoskeletal processes	–
<i>CDC42EP1</i>	Downregulated	Cytoskeletal processes	–
<i>IQGAP3</i>	Downregulated	Cytoskeletal processes	SRF
<i>PLXNB2</i>	Downregulated	Cytoskeletal processes	CREB1, SRF
<i>EIF4G3</i>	Upregulated	RNA processing	–
<i>NUDT12</i>	Upregulated	RNA processing	–
<i>SART1</i>	Downregulated	RNA processing	CREB1, SRF
<i>DDX31</i>	Downregulated	RNA processing	CREB1, SRF
<i>DDX54</i>	Downregulated	RNA processing	SRF
<i>MYBBP1A</i>	Downregulated	RNA processing	SRF
<i>NCALD</i>	Upregulated	Vesicular activity	–
<i>PREPL</i>	Upregulated	Vesicular activity	CREB1
<i>TMED1</i>	Downregulated	Vesicular activity	SRF
<i>ZFPL1</i>	Downregulated	Vesicular activity	CREB1, SRF
<i>AGPAT3</i>	Downregulated	Metabolism	–
<i>AACS</i>	Downregulated	Metabolism	CREB1
<i>CTBS</i>	Upregulated	Metabolism	–
<i>DCAKD</i>	Downregulated	Metabolism	CREB1
<i>HS2ST1</i>	Upregulated	Metabolism	–
<i>ORMDL2</i>	Downregulated	Metabolism	–
<i>PCK2</i>	Downregulated	Metabolism	–
<i>PI4KB</i>	Downregulated	Metabolism	–
<i>UAP1L1</i>	Downregulated	Metabolism	CREB1
<i>COMT</i>	Downregulated	Metabolism	CREB1

Transcriptomics DE analysis of male patients with MDS versus male controls and female patients with RTT versus female controls revealed 2465 and 3716 DEGs, respectively. Proteomics DE analysis returned 300 and 238 DEPs, respectively. Of these, 721 DEGs and 12 DEPs are shared between both groups, but only 2 genes are dysregulated with both omics in both syndromes (Fig. 4F). Those common DE genes are *MYOIC* and *HARS2*. *MYOIC* (OMIM\*606538) is a myosin involved in cytoskeletal organisation and vesicle trafficking to the plasma membrane and is consistently downregulated. *HARS2* (OMIM\*600783) is a mitochondrial histidyl-tRNA synthetase 2. At the RNA levels, it is consistently upregulated in patients with MDS

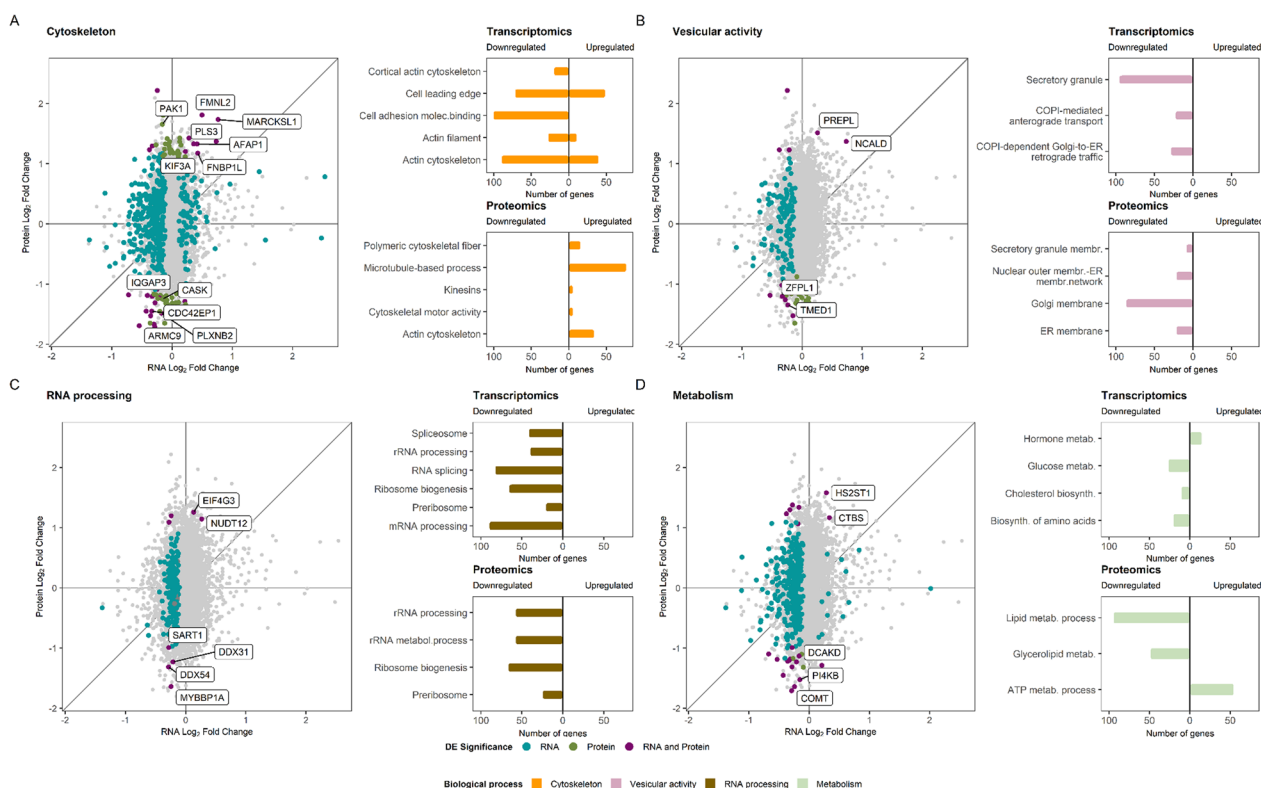
and downregulated in patients with RTT. At the protein level, however, it is upregulated in both sets of patients.

Because *MECP2* expression is decreased in RTT and increased in MDS, we wondered whether they share DEGs that are expressed in opposite directions. In our cohort, 82 DEGs were positively correlated with MeCP2 expression levels (hence, upregulated in MDS and downregulated in RTT), and 100 DEGs were negatively correlated with MeCP2 expression (upregulated in RTT and downregulated in MDS) (Additional file 9: Table S6a,b). Enrichment analysis of those two gene sets revealed that pathways related to cytoskeleton and mRNA processing are altered. In addition, we found other molecular functions and pathways commonly altered between the RTT and MDS cohorts, some of which could help to understand why these two syndromes share clinical traits (Fig. 4A,B; Additional file 7: Table S4c).

The 82 DEGs downregulated in patients with RTT and upregulated in patients with MDS are overrepresented in terms related to mRNA processing and cell cycle (Fig. 4A, Additional file 7: Table S4c). mRNA splicing-related genes appear dysregulated in both analyses. Interestingly, 8 of the 82 DEGs are part of spliceosome complexes and another four are related to mRNA stability, processing and maturation functions. When looking at the ChEA3 TF enrichment analysis that regulates the same 82 DEGs, we found several TFs, most of them zinc finger proteins, described as cell cycle regulators and also *SRF*, which we found in the RTT ChEA3 analysis (Fig. 4C, Additional file 8: Table S5b). These results are consistent with our findings in transcriptomics enrichment.

The 100 DEGs upregulated in RTT and downregulated in MDS enrich processes related to neurogenesis regulation; signalling cascades, such as Wnt, BMP and TGF $\beta$ ; and the cytoskeleton (Fig. 4B, Additional file 7: Table S4c). TF analysis with ChEA3 for the 100 DEGs revealed that *CREB1* (BH<0.05 in DE analysis) is upregulated and *SRF* is downregulated, and that they regulate 39% and 22% of the shared 100 DEGs, respectively. Moreover, the following TFs related to neuronal function are also enriched in the ChEA3 analysis: *HEYL*, *GLIS2*, *NFATC4* and *JUN* (Fig. 4C, Additional file 8: Table S5c).

Among the shared 12 DEPs, three, *APPL2*, *CNPY4* and *CTSC*, regulate immune response and are downregulated in RTT and upregulated in MDS (Fig. 4E, Additional file 9: Table S6c). Two other DEPs are related to cytoskeleton functions: *REPS1* and *CNN1*. *REPS1* (OMIM\*614825) is a signalling adaptor protein that mediates cytoskeletal changes as endocytosis, and the protein is upregulated in both syndromes. *CNN1* (OMIM\*600806) can bind to the cytoskeleton and produce smooth muscle contractions and is upregulated in RTT and downregulated in MDS.



**Fig. 3** Summary of the enriched biological processes found in RTT-MECP2 versus healthy control analysis. **A** Significant DEGs and DEPs related to cytoskeleton (coloured in the dot plot) and the enriched terms found for those DEGs (upper bar plot) and DEPs (lower bar plot). **B** Significant DEGs and DEPs related to vesicular activity and the enriched terms found for them. **C** Significant DEGs and DEPs related to RNA processing and the enriched terms found for them. **D** Significant DEGs and DEPs related to metabolism and the enriched terms found for them

**Patients with RTT versus patients with RTT-like phenotypes**

Our RTT-like cohort was recruited considering their resemblance to the RTT phenotype. It encompassed nine patients with mutations in five different genes plus three patients without an established molecular diagnosis. The greater heterogeneity of this group complicated the identification of DEGs, as well as the interpretation of the differential expression results. Therefore, we established a significance threshold of  $BH < 0.1$  for transcriptomics to be able to call DEGs despite the data heterogeneity. We interpreted these results in comparison with those obtained in typical patients with RTT, searching for shared molecular alterations that could constitute common grounds in the pathogenesis of overlapping disorders of diverse genetic nature.

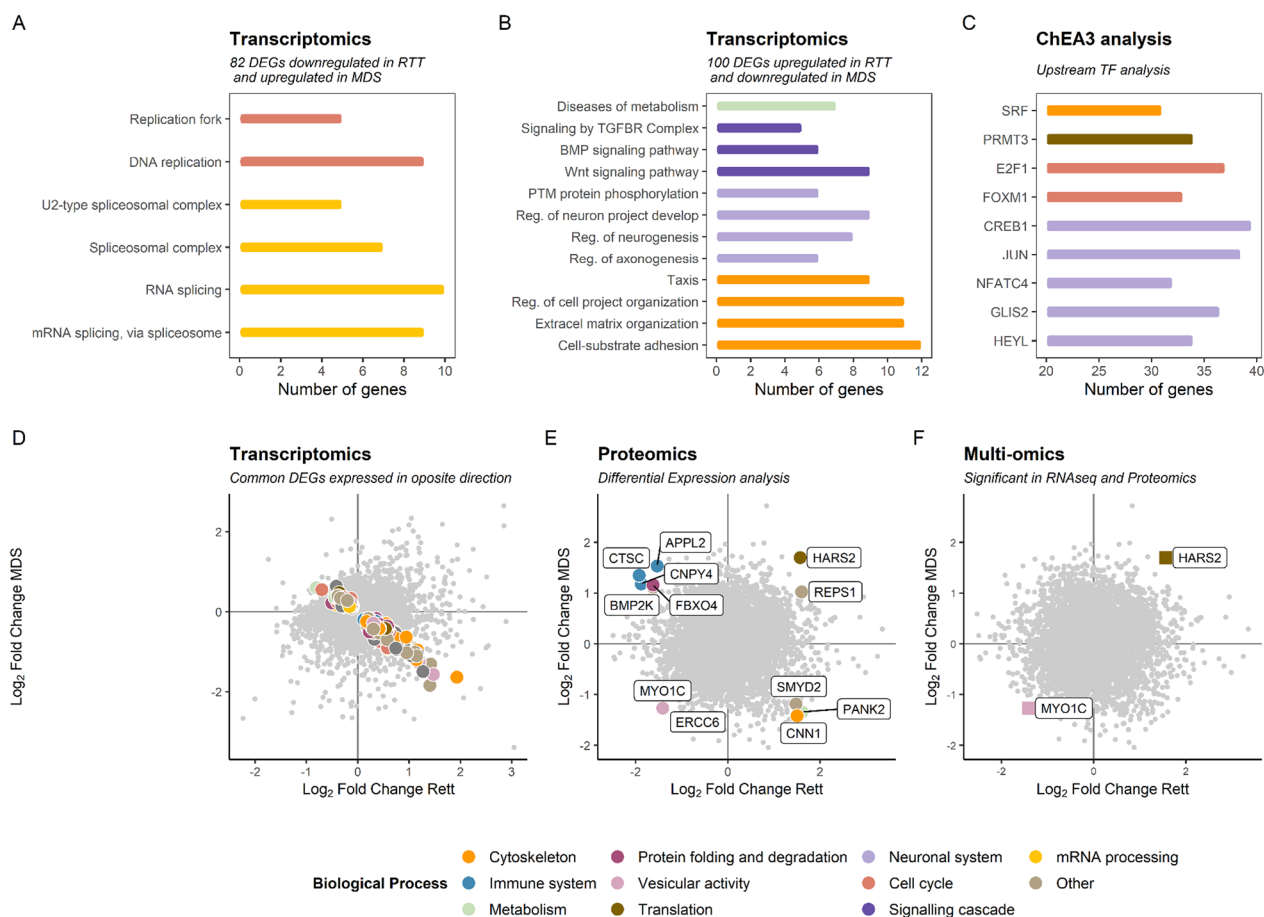
DE analysis of transcriptomics data revealed 63 genes consistently altered in patients with RTT and RTT-like phenotypes (25 upregulated and 38 downregulated) (Fig. 5A, Additional file 9: Table S6d). SRF targets were significantly overrepresented in these common DEGs, with 31 putative targets out of 63 common DEGs. This could implicate SRF transcriptional regulation as a

common mechanism linking the molecular phenotypes in RTT-spectrum disorders.

Proteomics data showed 81 proteins consistently dysregulated (39 upregulated and 42 downregulated) (Fig. 5C; Additional file 9: Table S6e), but no gene was altered in both transcriptomics and proteomics reaching significance. Nevertheless, some of the candidate genes identified in the multi-omics approach in RTT patients maintained a consistent dysregulation at the protein level in patients with RTT-like phenotypes [Table 3].

Transcriptomic and proteomic profiles of patients with RTT-like phenotypes are significantly correlated with those of patients with typical RTT ( $r = 0.69$ ,  $adj-R^2 = 0.47$ ,  $p < 0.001$  in transcriptomics;  $r = 0.75$ ,  $adj-R^2 = 0.56$ ,  $p < 0.001$  in proteomics) (Fig. 5A and C). Enrichment analysis of common DEGs and DEPs revealed terms related to cytoskeletal organisation, RNA processing, vesicular activity and metabolism, which constitute shared molecular alterations shared in patients with typical RTT and RTT-like phenotypes and could explain phenotypic overlap to some extent (Fig. 5B and D; Additional file 7: Table S4d-e).





**Fig. 4** Summary of the common findings between the analysis of female RTT-MECP2 patients versus female controls and male MDS patients versus male controls. **A** Enrichment analysis results for the shared 82 DEGs downregulated in RTT and upregulated in MDS, coloured by Biological Process (BP). **B** Enrichment analysis results for the shared 100 DEGs upregulated in RTT and downregulated in MDS coloured by BP. **C** Relevant TF from ChEA3 analysis for the 82 and 100 DEGs dysregulated at transcriptomic level coloured by BP. **D** Transcriptomics DE analysis results. The common 182 DEGs expressed in opposite directions are coloured by BP. **E** Common 12 DEPs coloured by BP. **F** Common DEGs and DEPs obtained from a multi-omics approach

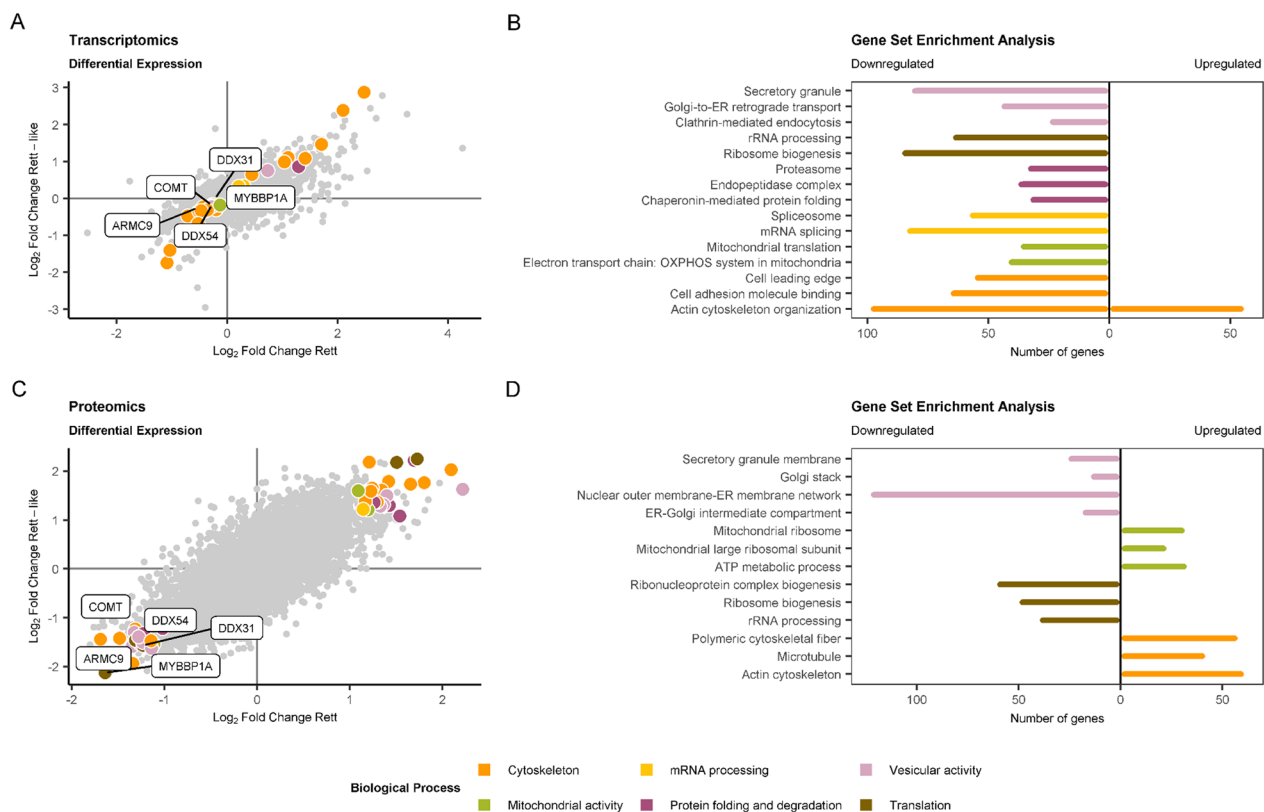
## Discussion

### Multi-omics expression in RTT patients with *MECP2* mutations

**Cytoskeletal** actin filament-based processes play a crucial role in neuronal development, and their dysregulation is associated with cognitive disorders like RTT [34–37]. We discovered a significant enrichment of cytoskeleton-related DEGs and DEPs in RTT patients compared to healthy controls.

Our study found that *ARMC9* (OMIM\*617612), a gene involved in cilium assembly, signalling and transport, was significantly downregulated in both mRNA and protein levels in patients with RTT. Its implication in cytoskeletal dynamics and the cytoskeletal abnormalities found in patients with RTT suggest a potential link between *ARMC9* and RTT pathogenesis

[38]. Scaffolding proteins, actin monomers and regulatory proteins were upregulated in RTT patients. We found an upregulation of p21-activated kinase 1 (*PAK1*, OMIM\*602590), essential for regulation of the actin cytoskeleton and controls dendritic spine morphogenesis and excitatory synapse formation [39]. Moreover, Roux et al. [40] found an upregulation of proteins related to cytoskeletal motor activities, such as tubulin monomers and kinesins, that could be implicated in axonal transport to the neuronal growth cone. Our study also showed a downregulation of protein levels of Ca<sup>2+</sup>/calmodulin-activated Ser-Thr kinase (*CASK*, OMIM\*300172), a scaffolding protein that is involved in synaptic transmembrane protein anchoring in the brain [41]. *CASK* dysfunction is a promising route towards understanding some of the pathomechanisms behind



**Fig. 5** Summary of the common findings between the analysis of RTT-*MECP2* patients versus healthy controls and RTT-like patients versus healthy controls. **A** Common 63 DEGs (25 upregulated and 38 downregulated), coloured by Biological Process (BP). **B** Gene set enrichment analysis results for the shared 63 DEGs coloured by BP. **C** Common 81 DEPs (39 upregulated and 42 downregulated), coloured by BP. **D** Gene set enrichment analysis results for the shared 81 DEGs coloured by BP

RTT since it has been linked to neurodevelopmental disorders with overlapping phenotypes with RTT [42].

Another consistently downregulated mRNA and protein was *COMT* (OMIM+116790), a methyltransferase required for the metabolism and degradation of catecholamine neurotransmitters, including epinephrine, norepinephrine and dopamine [43]. Patients with RTT and RTT mouse models have shown low levels of these biogenic amines, and alteration in dopaminergic

metabolism has been associated with the characteristic motor deficits of RTT [44].

We also found a significant enrichment in genes and proteins related to **vesicular activity** located in the Golgi apparatus and the nuclear outer membrane–endoplasmic reticulum membrane network, as well as secretory vesicles. We found a significant upregulation of vesicular proteins located in neuronal axons. Prolyl endopeptidase like (*PREPL*, OMIM\*609557) is a cytoplasmic protein with high expression in neuronal tissues. *PREPL* interacts with adaptor complex 1 (AP-1), a protein complex that plays an essential role in vesicular trafficking [45]. NCALD is a neuronal calcium sensor protein that is involved in calcium signalling. It interacts with clathrin and actin and is involved in the modulation of endocytosis and synaptic vesicle recycling. NCALD was found to bind clathrin only at low calcium levels, resulting in inhibition and modulation of synaptic vesicle recycling [46]. Our study also found a significant downregulation of ZFPL1, a cis-Golgi membrane protein that regulates trafficking from the endoplasmic reticulum to the Golgi apparatus and

**Table 3** Candidate genes identified in multi-omics analysis of patients with RTT and with a concordant alteration at the protein level in patients with RTT-like phenotypes

Gene	Direction	Biological process	Potential TF
<i>ARMC9</i>	Upregulated	Cytoskeletal processes	SRF
<i>DDX31</i>	Downregulated	RNA processing	CREB1, SRF
<i>DDX54</i>	Downregulated	RNA processing	SRF
<i>MYBBP1A</i>	Downregulated	RNA processing	SRF
<i>COMT</i>	Downregulated	Metabolism	CREB1

maintains cis-Golgi structural and functional integrity [47].

We identified a downregulation of genes and proteins involved in **rRNA processing** and ribosome biogenesis in patients with RTT [48–50]. This could affect general protein translation in affected cells, possibly due to a reduction in mTORC1 activity [48, 49]. The downregulation of three proteins that interact with MeCP2, *MYBBP1A*, *DDX31* and *DDX54*, could explain alterations in rRNA processing and mRNA splicing [51]. The exact nature of the interaction between MeCP2 and these proteins is still unknown. We also observed downregulation of DEGs involved in **mRNA splicing** and spliceosomal complexes in patients with RTT. *MECP2* is known to interact with splicing factors [37, 52, 53], but a recent publication questions its role as a global regulator of splicing [54]. Additional studies are needed to clarify *MECP2*'s role in splicing since many genes involved in mRNA splicing are repeatedly dysregulated in different transcriptomics experiments.

*CREB1* (OMIM\*123810), which is a known MeCP2 interactor, regulates transcription in processes of relevance for neuronal survival and memory consolidation, among others [55, 56]. In astrocytes, it even regulates genes related to mitochondrial function, vesicle dynamics and the cytoskeleton [57]. Besides, one-third of our DEGs are regulated by *CREB1* and *CREB1* itself was significantly upregulated in our cohort at the mRNA level. *SRF* (OMIM\*600589), which is an integrator of mitogen-activated protein kinase (MAPK) and Rho-GTPase-mediated signalling, regulates cytoskeletal dynamics. *SRF* binds to the serum response element (SRE) sequence, present in a subset of cytoskeletal genes such as *ACTB* and immediate early genes (IEGs) [58]. Besides, *SRF* regulates neuronal morphology and activity-dependent transcription [59] and suppression of *SRF*-mediated transcriptional responses has been found to produce a reduction in dendritic complexity in cortical neurons, which could contribute to the neuronal spine dysgenesis phenotype observed in patients with an RTT-spectrum phenotype [60].

None of the genes regulated in opposite directions in transcriptomics and proteomics were known MeCP2 partners. We analysed the functional relationships between them, but no clear biological processes were identified. We hypothesise that the discordance in transcriptomics and proteomics may be due to cellular compensatory processes.

#### RTT and MDS: *MECP2* gene, two syndromes

Our study found that there are common genes between RTT and MDS. We found two shared genes, *MYOIC* and *HARS2*, which are a cytoskeletal component and a

tRNA synthetase, respectively, in common significant DEGs and DEPs. The AKT/mTOR signalling pathway is downregulated in *Mecp2* null models, indicating a deregulation of transcription followed by a limited ability to generate functional proteins [48]. Both syndromes seem to have a deregulation of the correct protein synthesis.

In total, 82 DEGs were downregulated in patients with RTT and upregulated in patients with MDS, indicating that *MECP2* is important for mRNA processing. Moran-Salvador et al. found a downregulated group of genes involved in DNA replication and cell proliferation in hepatic stellate cells of *Mecp2*-null mice and suggested inhibition of *Mecp2* phosphorylation as a liver fibrosis treatment [61].

The 100 DEGs upregulated in RTT and downregulated in MDS revealed processes related to neurogenesis regulation, cytoskeleton, and Wnt, BMP and TGF $\beta$  signalling cascades. The Wnt, BMP and TGF $\beta$  signalling pathways are also involved in osteoblast activity and maintenance of cartilage [62–64]. Patients with RTT suffer from scoliosis, low bone mass density and a higher bone fracture rate than the general population [65, 66]. Scoliosis is the most commonly reported orthopaedic issue in patients with MDS, and osteopenia, contractures of joints and fractures have also been reported [67].

Our results detected four transcription factors related to neuronal function, *HEYL*, *GLIS2*, *NFATC4* and *JUN*. *HEYL* and *GLIS2* promote neuronal differentiation [68, 69], whilst *NFATC4* regulates adult hippocampal neurogenesis and shares a common signalling process with *BDNF* for neuron maturation [70, 71]. *BDNF* modulates many aspects of neuronal development, synaptic transmission and plasticity, and its dysregulation is found in RTT [72]. *JUN* plays a role in neuronal migration and axon–dendritic architecture, and its inhibition reduces breathing abnormalities in RTT mice and induced pluripotent stem cell neuronal models, and rescues the dendritic spine alterations [73].

Our findings indicate a resemblance of both syndromes at a molecular level, with several TFs involved in neural processes and dendritic complexity. Therapeutic strategies that seem promising for one syndrome could also benefit the other if the correct gene dosages are reached.

#### TT-spectrum: one common clinical presentation, different mutated genes

The results of our study found that patients with RTT-spectrum disorders share common molecular alterations that could impact neuronal phenotypes. Almost one-third of the common DEGs are involved in cytoskeleton organisation and regulation, and some of these have important functions in neurons. The malfunctioning of cytoskeletal genes with prominent functions in neurite

development could lead to neuronal spine dysgenesis and, consequently, to the emergence of disorders with common traits derived from this structural synaptic dysfunction [74, 75]. The enrichment in putative SRF targets among shared DEGs highlights the potential implication of SRF transcriptional regulation in RTT-spectrum common molecular alterations leading to overlapping phenotypes.

We also detected an overrepresentation of several terms related to nervous system development and structure, supporting that common molecular alterations found in patients with RTT-spectrum phenotypes can impact neuronal phenotypes. The downregulation of *ARMC9* observed in patients with typical RTT can also be observed in patients with RTT-like phenotypes, constituting a link between RTT-spectrum disorders and the overlapping phenotype caused by loss-of-function variants in this gene.

The patients with RTT-spectrum phenotypes in our study shared a downregulation of *SNRPC* expression at the RNA level that was not replicated in proteomics. This transcriptional alteration was also previously found in post-mortem brain tissue and embryonic stem cells of patients with RTT [49, 76]. *SNRPC* is a spliceosome component involved in 5' splice-site recognition, so it may affect the splicing of many different targets and could constitute a shared mechanism of splicing dysregulation of patients with RTT-spectrum phenotypes. The dysregulation of splicing factors and regulators has been described in RTT as well as in other monogenic intellectual disabilities and in autism spectrum disorders (ASD) [77].

Protein translation may be affected in all patients with RTT-spectrum phenotypes. Several rRNA processing and ribosome biogenesis-related proteins found altered in patients with RTT were also consistently dysregulated in patients with RTT-like phenotypes, indicating this commonality. *DDX54*, *DDX31* and *MYBBP1A* are MeCP2 partners and are linked to rRNA expression and preprocessing and could explain, at least to some extent, the shared dysregulation of ribosome biogenesis.

## Conclusions

Numerous studies have investigated the transcriptomes of individuals with RTT, resulting in over 60 published articles. Our study found that studying other human tissues, such as fibroblasts, can reflect the same dysregulations caused by loss of function of *MECP2*. However, integrating all knowledge is complicated by the heterogeneity in experiments and tissue-specific effects of *MECP2*. Dysregulation of various cellular functions was identified, including cytoskeletal organisation, vesicular

activity, translation and mRNA processing, which are altered in patients with RTT, RTT-like phenotypes and MDS. *ARMC9* could be a potential biomarker for RTT and RTT-spectrum disorders. TF analysis supports *CREB1* and *SRF* as potential therapeutic targets. Shared dysregulated biological processes and cellular functions were found between patients with RTT, MDS and RTT-like phenotypes, with RTT and RTT-like being more similar than MDS. Further studies are necessary to validate these findings.

## Abbreviations

ASD	Autism spectrum disorder
BH	Benjamini–Hochberg-corrected p-value
bp	Base pair
<i>CDKL5</i>	Cyclin-dependent kinase-like 5
cDNA	Double-stranded complementary DNA
CEIC	Comitè d'Ètica d'Investigació Clínica-Fundació Sant Joan de Déu ChIP-X Enrichment Analysis 3
CPM	Counts per million mapped reads
DEG	Differential expression genes
DEP	Differential expression proteins
EEGs	Electroencephalograms
<i>FOXP1</i>	Forkhead box G1
GRCh37 (hg19)	Homo sapiens (human) genome assembly GRCh37 (hg19) from Genome Reference Consortium
GSEA	GO: Gene Ontology
HSJD	Hospital Sant Joan de Déu
ID	Intellectual disability
IEGs	Immediate early genes
iPSCs	Induced pluripotent stem cells
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC–MS	Liquid chromatography–mass spectrometry
MDS	MECP2-duplication syndrome
<i>MECP2</i>	Methyl-CpG-binding protein 2
NGS	Next-generation sequencing
NF-κB	Nuclear factor kappa B
ORA	Overrepresentation analysis
PCA	Principal component analysis
RNAseq	RNA sequencing
RP	Reactome pathway
RT-qPCR	Reverse transcription-quantitative polymerase chain reaction
RTT	Rett syndrome
RTT-like	Rett-like
TF	Transcription factors
ChEA3: WP	WikiPathways
XCI	X-chromosome inactivation

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40246-023-00532-1>.

**Additional file 1: Fig. S1.** Detection of fibroblast samples under oxidative stress conditions.

**Additional file 2: Fig. S2.** Comparison of the gene expression results between RNAseq and RT-qPCR.

**Additional file 3: Fig. S3.** Comparison of genes from Genotype-Tissue Expression (GTEx) project and mean TPM (Transcripts per Kilobase Million) in fibroblast cultured cells in RNAseq and RT-qPCR.

**Additional file 4: Table S1.** Detailed description of the study cohort.

**Additional file 5: Table S2.** Sequence of the designed primers.



**Additional file 6: Table S3.** RNAseq and proteomics differential expression results for RTT - MECP2 versus healthy controls.

**Additional file 7: Table S4.** Enrichment analysis results for the RTT-MECP2 versus healthy controls, RTT versus MDS, RTT versus RTT-like, transcriptomic and proteomic data.

**Additional file 8: Table S5.** ChEA3 upstream analysis results for RTT patients versus healthy controls, RTT versus MDS, RTT versus RTT-like.

**Additional file 9: Table S6.** Summary for the DEG and DEP from RTT-MECP2 vs MDS and RTT-MECP2 versus RTT-like.

### Acknowledgements

We thank all patients and their families for their collaboration and support in this study. We would like to thank the Cytogenetics Department of the Hospital Sant Joan de Déu, especially Lluïsa, Teresa, Carlos, Rosa and Esther for starting some of the fibroblast primary cell cultures. We would like to thank the Bioinformatics department of the Hospital Sant Joan de Déu for their help in optimising the RNAseq data obtention protocol. We would like to thank Riccardo from the Institute of Human Genetics of Munich for the raw RNAseq data preparation and the Proteomics Core Facility at TUM for performing the proteomics experiments. We thank the 'Biobanc de l'Hospital Infantil Sant Joan de Déu per a la Investigació' integrated in the Spanish Biobank Network of Instituto de Salud Carlos III (ISCIII) for sample storage.

### Author contributions

APA, CX and JA conceptualised the study, interpreted the data and wrote the original draft. JA designed the work. CX and APA assisted with the acquisition. APA, CX, DS and RK analysed the data. JA and HP contributed to supervision. All authors took part in writing—reviewing and editing. All authors read and approved the final manuscript.

### Funding

This research was funded with a grant from the Spanish Ministry of Health (Instituto de Salud Carlos III/FEDER, PI20/00389), the parent association 'Síndrome duplicación MECP2: Miradas que hablan' (PFNR0085), funding from Muévete por los que no Pueden (PCP/00282), the doctoral grant FI from the Government of Catalonia (Secretaria d'Universitats i Recerca) and the European Social Fund (2020 FI-B 00888) to Ainhoa Pascual-Alonso, and the FPU (Formación del Profesorado Universitario) doctoral grant from the Spanish Ministry of Science, Innovation and Universities (FPU18/02152) to Clara Xiol.

### Availability of data and materials

Our ethics approval and consent agreements allow us to share non-identifiable patient data and analysis data only, as such, we cannot provide BAM or VCF files. The analysis data provided are the gene expression count matrices, as well as the privacy-preserving count matrices of split and unsplit reads overlapping annotated splice sites from RNAseq. They will be available for download without restriction when this article will be published.

### Declarations

#### Ethics approval and consent to participate

All individuals included or their legal guardians provided written informed consent before evaluation. The study was approved by the ethical committee of the Fundació Sant Joan de Déu (#PIC-219–20). The research conformed to the principles of the Declaration of Helsinki.

#### Consent for publication

All individuals included or their legal guardians provided written consent to share pseudonymised patient data and analysis data.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Fundació Per La Recerca Sant Joan de Déu, Esplugues de Llobregat, Spain. <sup>2</sup>Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Spain. <sup>3</sup>Institute of Human Genetics, Technical University of Munich, Munich, Germany.

<sup>4</sup>Institute of Neurogenomics, Helmholtz Zentrum München, Munich, Germany. <sup>5</sup>CIBER-ER (Biomedical Network Research Center for Rare Diseases), Instituto de Salud Carlos III (ISCIII), Madrid, Spain. <sup>6</sup>Genomic Unit, Molecular and Genetic Medicine Section, Hospital Sant Joan de Déu, Barcelona, Spain.

Received: 18 January 2023 Accepted: 3 September 2023

Published online: 15 September 2023

### References

- Ehrhart F, Coort SLM, Cirillo E, Smeets E, Evelo CT, Curfs LMG. Rett syndrome - Biological pathways leading from MECP2 to disorder phenotypes. *Orphanet J Rare Dis*. 2016;11(1):158.
- Neul JL, Kaufmann WE, Glaze DG, Christodoulou J, Clarke AJ, Bahi-Buisson N, et al. Rett syndrome: revised diagnostic criteria and nomenclature. *Ann Neurol*. 2010;68(6):944–50.
- Neul JL, Fang P, Barrish J, Lane J, Caeg EB, Smith EO, et al. Specific mutations in methyl-CpG-binding protein 2 confer different severity in Rett syndrome. *Neurology*. 2008;70(16):1313–21.
- Bahi-Buisson N, Nectoux J, Rosas-Vargas H, Milh M, Boddaert N, Girard B, et al. Key clinical features to identify girls with CDKL5 mutations. *Brain*. 2008;131(10):2647–61.
- Ariani F, Hayek G, Rondinella D, Artuso R, Mencarelli MA, Spanhol-Rosseto A, et al. FOXP1 is responsible for the congenital variant of Rett syndrome. *Am J Hum Genet*. 2008;83(1):89–93.
- Vidal S, Xiol C, Pascual-Alonso A, O'Callaghan M, Pineda M, Armstrong J. Genetic landscape of Rett syndrome spectrum: improvements and challenges. *Int J Mol Sci*. 2019;20(16):3925.
- Schönewolf-Greulich B, Bisgaard AM, Møller RS, Dunø M, Brøndum-Nielsen K, Kaur S, et al. Clinician's guide to genes associated with Rett-like phenotypes—Investigation of a Danish cohort and review of the literature. *Clin Genet*. 2019;1(95):221–30.
- Ehrhart F, Sangani NB, Curfs LMG. Current developments in the genetics of Rett and Rett-like syndrome. *Curr Opin Psychiatry*. 2018;31(2):103–8.
- Van Esch H. MECP2 duplication syndrome. *Mol Syndromol*. 2011;2:128–36.
- Peters SU, Fu C, Suter B, Marsh E, Benke TA, Skinner SA, et al. Characterizing the phenotypic effect of Xq28 duplication size in MECP2 duplication syndrome. *Clin Genet*. 2019;95:575–81.
- Murdock DR, Dai H, Burrage LC, Rosenfeld JA, Ketkar S, Müller MF, et al. Transcriptome-directed analysis for Mendelian disease diagnosis overcomes limitations of conventional genomic testing. *J Clin Invest*. 2021;131(11):e141500.
- Yépez VA, Gusic M, Kopajtich R, Mertens C, Smith NH, Alston CL, et al. Clinical implementation of RNA sequencing for Mendelian disease diagnostics. *Genome Med*. 2022;1(14):38.
- Gomathi M, Padmapriya S, Balachandrar V. Drug studies on Rett syndrome: from bench to bedside. *J Autism Dev Disord*. 2020;50(8):2740–64.
- Pacheco NL, Heaven MR, Holt LM, Crossman DK, Boggio KJ, Shaffer SA, et al. RNA sequencing and proteomics approaches reveal novel deficits in the cortex of Mecp2-deficient mice, a model for Rett syndrome. *Mol Autism*. 2017;8(56):1–24.
- Vidal S, Brandi N, Pacheco P, Gerotina E, Blasco L, Trotta JR, et al. The utility of Next Generation Sequencing for molecular diagnostics in Rett syndrome. *Sci Rep*. 2017;7:12288.
- Pascual-Alonso A, Blasco L, Vidal S, Gean E, Rubio P, O'Callaghan M, et al. Molecular characterization of Spanish patients with MECP2 duplication syndrome. *Clin Genet*. 2020;1(97):610–20.
- Monrós E, Armstrong J, Aibar E, Poo P, Canós I, Pineda M. Rett syndrome in Spain: mutation analysis and clinical correlations. *Brain Dev*. 2001;23:S251–3.
- Allen RC, Zoghbi HY, Moseley AB, Rosenblatt HM, Belmont JW. Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *Am J Hum Genet*. 1992;51:1229–39.
- Al-Kafaji G, Sabry MA, Bakhiet M. Increased expression of mitochondrial DNA-encoded genes in human renal mesangial cells in response

- to high glucose-induced reactive oxygen species. *Mol Med Rep.* 2016;13(2):1774–80.
20. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3-new capabilities and interfaces. *Nucleic Acids Res.* 2012;40(15):e115.
  21. Anders S, Pyl PT, Huber W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics.* 2015;31(2):166–9.
  22. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;5(15):550.
  23. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innov.* 2021;2(3):100141.
  24. Yu G, He QY. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol Biosyst.* 2016;12(2):477–9.
  25. Carbon S, Douglass E, Good BM, Unni DR, Harris NL, Mungall CJ, et al. The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res.* 2021;49(D1):D325–34.
  26. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res.* 2021;49(D1):D545–51.
  27. Martens M, Ammar A, Riutta A, Waagmeester A, Slenter DN, Hanspers K, et al. WikiPathways: connecting communities. *Nucleic Acids Res.* 2021;49(D1):D613–21.
  28. Gillespie M, Jassal B, Stephan R, Milacic M, Rothfels K, Senff-Ribeiro A, et al. The reactome pathway knowledgebase 2022. *Nucleic Acids Res.* 2022;50(D1):D687–92.
  29. Keenan AB, Torre D, Lachmann A, Leong AK, Wojciechowicz ML, Utti V, et al. ChEA3: transcription factor enrichment analysis by orthogonal omics integration. *Nucleic Acids Res.* 2019;47(W1):W212–24.
  30. Zecha J, Satpathy S, Kanashova T, Avanesian SC, Kane MH, Clauser KR, et al. TMT labeling for the masses: a robust and cost-efficient, in-solution labeling approach. *Mol Cell Proteomics.* 2019;18(7):1468–78.
  31. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc.* 2016;11(12):2301–19.
  32. Kopajtich R, Smirnov D, Stenton SL, Loipfinger S, Meng C, Scheller IF, et al. Integration of proteomics with genomics and transcriptomics increases the diagnostic rate of Mendelian disorders. *medRxiv*
  33. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7): e47.
  34. Pecorelli A, Cervellati C, Cortelazzo A, Cervellati F, Sticozzi C, Mirasole C, et al. Proteomic analysis of 4-hydroxynonenal and nitrotyrosine modified proteins in RTT fibroblasts. *Int J Biochem Cell Biol.* 2016;1(81):236–45.
  35. Cortelazzo A, De Felice C, Pecorelli A, Belmonte G, Signorini C, Leoncini S, et al. Beta-actin deficiency with oxidative posttranslational modifications in Rett syndrome erythrocytes: insights into an altered cytoskeletal organization. *PLoS ONE.* 2014;9(3):e93181.
  36. Varderdou-Minasian S, Hinz L, Hagemans D, Posthuma D, Altelaar M, Heine VM. Quantitative proteomic analysis of Rett iPSC-derived neuronal progenitors. *Mol Autism.* 2020;27(11):38.
  37. Ehrhart F, Coort SL, Eijssen L, Cirillo E, Smeets EE, Bahram Sangani N, et al. Integrated analysis of human transcriptome data for Rett syndrome finds a network of involved genes. *World J Biol Psychiatry.* 2019;21(10):712–25.
  38. Latour BL, van de Weghe JC, Rusterholz TDS, Letteboer SJF, Gomez A, Shaheen R, et al. Dysfunction of the ciliary ARMC9/TOGARAM1 protein module causes Joubert syndrome. *J Clin Invest.* 2020;140(8):4423–39.
  39. Zhang H, Webb DJ, Asmussen H, Niu S, Horwitz AF. A GIT1/PIX/Rac/PAK signaling module regulates spine morphogenesis and synapse formation through MLC. *J Neurosci.* 2005;25(13):3379–88.
  40. Roux JC, Zala D, Panayotis N, Borges-Correia A, Saudou F, Villard L. Modification of Mecp2 dosage alters axonal transport through the Huntingtin/Hap1 pathway. *Neurobiol Dis.* 2012;45(2):786–95.
  41. Setou M, Nakagawa T, Seog D-H, Hirokawa N. Kinesin superfamily motor protein KIF17 and mLin-10 in NMDA receptor-containing vesicle transport. *Science.* 2000;288:1796–802.
  42. Patel PA, Hegert JV, Cristian I, Kerr A, LaConte LEW, Fox MA, et al. Complete loss of the X-linked gene CASK causes severe cerebellar degeneration. *J Med Genet.* 2022;59(11):1044–57.
  43. Chen J, Song J, Yuan P, Tian Q, Ji Y, Ren-Patterson R, et al. Orientation and cellular distribution of membrane-bound catechol-O-methyltransferase in cortical neurons: implications for drug development. *J Biol Chem.* 2011;286(40):34752–60.
  44. Szczesna K, De La Caridad O, Petazzi P, Soler M, Roa L, Saez MA, et al. Improvement of the rett syndrome phenotype in a mecp2 mouse model upon treatment with levodopa and a dopa-decarboxylase inhibitor. *Neuropsychopharmacology.* 2014;39(12):2846–56.
  45. Radhakrishnan K, Baltus J, Creemers JWM, Schu P. Trans-Golgi network morphology and sorting is regulated by prolyl-oligopeptidase-like protein PREPL and the AP-1 complex subunit  $\mu$ 1A. *J Cell Sci.* 2013;126(5):1155–63.
  46. Riessland M, Kaczmarek A, Schneider S, Swoboda KJ, Löhr H, Bradler C, et al. Neurocalcin delta suppression protects against spinal muscular atrophy in humans and across species by restoring impaired endocytosis. *Am J Hum Genet.* 2017;100(2):297–315.
  47. Chiu CF, Ghanekar Y, Frost L, Diao A, Morrison D, McKenzie E, et al. ZFPL1, a novel ring finger protein required for cis-Golgi integrity and efficient ER-to-Golgi transport. *EMBO J.* 2008;27(7):934–47.
  48. Ricciardi S, Boggio EM, Grosso S, Lonetti G, Forlani G, Stefanelli G, et al. Reduced AKT/mTOR signaling and protein synthesis dysregulation in a Rett syndrome animal model. *Hum Mol Genet.* 2011;20(6):1182–96.
  49. Li Y, Wang H, Muffat J, Cheng AW, Orlando DA, Lovén J, et al. Global transcriptional and translational repression in human-embryonic-stem-cell-derived rett syndrome neurons. *Cell Stem Cell.* 2013;3(13):446–58.
  50. Rodrigues DC, Mufteev M, Weatheritt RJ, Djuric U, Ha KCH, Ross PJ, et al. Shifts in ribosome engagement impact key gene sets in neurodevelopment and ubiquitination in Rett syndrome. *Cell Rep.* 2020;30(12):4179–96.
  51. Huttlin EL, Bruckner RJ, Paulo JA, Cannon JR, Ting L, Baltier K, et al. Architecture of the human interactome defines protein communities and disease networks. *Nature.* 2017;545(7655):505–9.
  52. Cheng TL, Chen J, Wan H, Tang B, Tian W, Liao L, et al. Regulation of mRNA splicing by MeCP2 via epigenetic modifications in the brain. *Sci Rep.* 2017;7:1–12.
  53. Long SW, Ooi JYY, Yau PM, Jones PL. A brain-derived MeCP2 complex supports a role for MeCP2 in RNA processing. *Biosci Rep.* 2011;31(5):333–43.
  54. Chhatbar K, Cholewa-Waclaw J, Shah R, Bird A, Sanguinetti G. Quantitative analysis questions the role of MeCP2 as a global regulator of alternative splicing. *PLoS Genet.* 2020;16(10):1–14.
  55. Chen Y, Shin BC, Thamotharan S, Devaskar SU. Creb1-Mecp2-mCpG complex transactivates postnatal murine neuronal glucose transporter isoform 3 expression. *Endocrinology.* 2013;154(4):1598–611.
  56. Kaldun JC, Sprecher SG. Initiated by CREB: resolving gene regulatory programs in learning and memory. *BioEssays.* 2019;41(8):1900045.
  57. Pardo L, Valor LM, Eraso-Pichot A, Barco A, Golbano A, Hardingham GE, et al. CREB regulates distinct adaptive transcriptional programs in astrocytes and neurons. *Sci Rep.* 2017;1(7):6390.
  58. Tabuchi A, Ihara D. Regulation of dendritic synaptic morphology and transcription by the SRF cofactor MKL/MRTF. *Front Mol Neurosci.* 2021;2(14):7678.
  59. Knöll B, Nordheim A. Functional versatility of transcription factors in the nervous system: the SRF paradigm. *Trends Neurosci.* 2009;32(8):432–42.
  60. Shiota J, Ishikawa M, Sakagami H, Tsuda M, Baraban JM, Tabuchi A. Developmental expression of the SRF co-activator MAL in brain: role in regulating dendritic morphology. *J Neurochem.* 2006;98(6):1778–88.
  61. Moran-Salvador E, Garcia-Macia M, Sivaharan A, Sabater L, Zaki MYW, Oakley F, et al. Fibrogenic activity of MECP2 is regulated by phosphorylation in hepatic stellate cells. *Gastroenterology.* 2019;1(157):1398–1412.e9.
  62. Liu J, Xiao Q, Xiao J, Niu C, Li Y, Zhang X, et al. Wnt/ $\beta$ -catenin signalling: function, biological mechanisms, and therapeutic opportunities. *Signal Transduct Target Ther.* 2022;7(1):3.
  63. Lowery JW, Rosen V. The BMP pathway and its inhibitors in the skeleton. *Physiol Rev.* 2018;98:2431–52.
  64. Finnson KW, Chi Y, Bou-Gharios G, Leask A, Philip A. TGF-beta signaling in cartilage homeostasis and osteoarthritis. *Front Biosci.* 2012;54:251–68.
  65. Downs J, Bebbington A, Woodhead H, Jacoby P, Jian L, Jefferson A, et al. Early determinants of fractures in Rett syndrome. *Pediatrics.* 2008;121(3):540–6.
  66. Pecorelli A, Cordone V, Schiavone ML, Caffarelli C, Cervellati C, Cerbone G, et al. Altered bone status in Rett syndrome. *Life.* 2021;11(11):521.

67. Ta D, Downs J, Baynam G, Wilson A, Richmond P, Leonard H. A brief history of MECP2 duplication syndrome: 20-years of clinical understanding. *Orphanet J Rare Dis.* 2022;1(17):131.
68. Jalali A, Bassuk AG, Kan L, Israsena N, Mukhopadhyay A, McGuire T, et al. HeyL promotes neuronal differentiation of neural progenitor cells. *J Neurosci Res.* 2011;89(3):299–309.
69. Lamar E, Kintner C, Goulding M. Identification of NKL, a novel Gli-Kruppel zinc-finger protein that promotes neuronal differentiation. *Development.* 2001;128:1335–46.
70. Li L, Ke K, Tan X, Xu W, Shen J, Zhai T, et al. Up-regulation of NFATc4 involves in neuronal apoptosis following intracerebral hemorrhage. *Cell Mol Neurobiol.* 2013;33(7):893–905.
71. Ding B, Dobner PR, Mullikin-Kilpatrick D, Wang W, Zhu H, Chow CW, et al. BDNF activates an NFI-dependent neurodevelopmental timing program by sequestering NFATc4. *Mol Biol Cell.* 2018;29(8):975–87.
72. Li W, Pozzo-Miller L. BDNF deregulation in Rett syndrome. *Neuropharmacology.* 2014;76:737–46.
73. Musi CA, Castaldo AM, Valsecchi AE, Cimini S, Morello N, Pizzo R, et al. JNK signaling provides a novel therapeutic target for Rett syndrome. *BMC Biol.* 2021;1(19):256.
74. Xu X, Miller EC, Pozzo-Miller L. Dendritic spine dysgenesis in Rett syndrome. *Front Neuroanat.* 2014;10(8):1–8.
75. Phillips M, Pozzo-Miller L. Dendritic spine dysgenesis in autism related disorders. *Neurosci Lett.* 2015;3(601):30–40.
76. Colantuoni C, Jeon OH, Hyder K, Chenchik A, Khimani AH, Narayanan V, et al. Gene expression profiling in postmortem Rett syndrome brain: differential gene expression and patient classification. *Neurobiol Dis.* 2001;8:847–65.
77. Shah S, Richter JD. Do fragile X syndrome and other intellectual disorders converge at aberrant pre-mRNA splicing? *Front Psychiatry.* 2021;10(12):715346.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

