 doi:10.20944/preprints202207.0030.v1

Research Article

Meta-Analysis Identifies Novel Common Genes Differently Altered in Cross-Species Models of Rett Syndrome

Florencia Haase [1,2,3], Rachna Singh [4], Brian Gloss [5], Patrick Tam [6,7], Wendy Gold [1,2,3]*

[1] School of Medical Science, Faculty of Medicine and Health, The University of Sydney, NSW, Australia
2 Kids Neuroscience Centre, Kids Research, Children's Hospital at Westmead, Westmead, New South Wales, Australia.
3 Molecular Neurobiology Research Laboratory, Kids Research, Children's Hospital at Westmead, Westmead, NSW, Australia
4 School of Medicine Sydney, The University of Notre Dame, NSW, Australia
5 Westmead Research Hub, Westmead Institute for Medical Research, Westmead, Sydney, New South Wales, Australia.
6 Embryology Research Unit, Children's Medical Research Institute, The University of Sydney, NSW, Australia.
* Correspondence: wendy.gold@sydney.edu.au

Abstract: Rett syndrome (RTT) is a rare disease and one of the most abundant causes for intellectual disabilities in females. Single mutations in the gene coding for methyl-CpG-binding protein 2 (MECP2), are responsible for the disease. MeCP2 regulates gene expression as a transcriptional regulator as well as through epigenetic imprinting and chromatin condensation. Consequently, numerous biological pathways on multiple levels are influenced however, the exact molecular pathways from genotype to phenotype are currently not fully elucidated. Treatment of RTT is purely symptomatic where no curative options for RTT have yet to reach the clinic. The paucity of this is mainly due to an incomplete understanding of the underlying pathophysiology of the disorder with no clinically useful common disease drivers, biomarkers or therapeutic targets being identified. With the premise of identifying universal and robust disease drivers and therapeutic targets, here we interrogated a range of RTT transcriptomic studies spanning different species, models and MECP2 mutations. A meta-analysis using RNA sequencing data from brains of RTT mouse models human post-mortem brain tissue and patient-derived induced pluripotent stem cells (iPSC) neurons was performed using Weighted Gene Correlation Network Analysis (WGCNA). This study identified a module of genes common to all datasets with the following ten hub genes driving the expression ATRX, ADCY7, ADCY9, SOD1, CACNA1A, PLCG1, CCT5, RPS9, BDNF and MECP2. Here we discuss the potential benefits of these genes as therapeutic targets.

Keywords: Rett syndrome; WGCNA; *MECP2*

1. Introduction

Rett syndrome (RTT) is one of the most common genetic causes of intellectual disabilities in females and affects 1 in 10,000 births [1]. RTT is an X-linked dominant disorder caused by mutations in the *MECP2* gene, which encodes the Methyl-CpG Binding Protein 2 (MeCP2) protein. The molecular pathogenesis of RTT remains poorly understood, with patients presenting with numerous complex disabilities, which are likely due to the pleiotropic molecular functions of MeCP2 and its ubiquitous expression. MeCP2 was initially characterised as a transcriptional repressor [2], however, further functions have been recently uncovered, including binding to additional motifs across the genome [2-4]; activation and repression of gene transcription [5, 6] and regulation of alternative splicing [7].

 doi:10.20944/preprints202207.0030.v1

Mice harbouring mutations in the *MECP2* gene are one of the most clinically relevant models for RTT as they recapitulate many of the features observed in RTT patients, such as seizures and motor and cognitive dysfunction, which has assisted in our understanding of the underlying pathophysiology [8]. However, despite the vast majority of RTT patients being female, most gene therapy and other preclinical studies in animal models of RTT have used male mice which is not truly representative of the patient population.

Despite the robust phenotype of the RTT mouse models, there are many differences in brain development and structure between humans and mice that may confound findings in translational preclinical studies [9] [10]. For example, the origin of cortical neurons in brain development differs in humans and mice with the subventricular zone where human neurogenesis mostly occurs, is significantly reduced in mice [10]. Thus, regardless of the significant insight gained from these models, inconsistencies between mouse models and human disease may affect the validity of preclinical findings.

Immortalised cell lines and postpost-mortem brain tissue have also been used extensively to study the pathophysiology of RTT. However, the use of post-mortem brain tissue is limited, in that cell lines do not represent the complex organisation of the brain and only reflect end-stage disease, and immortalised cell lines are not fully translatable to other models. Furthermore, post-mortem tissue from the brain cannot be used in live-cell testing studies such as electrophysiology [11].

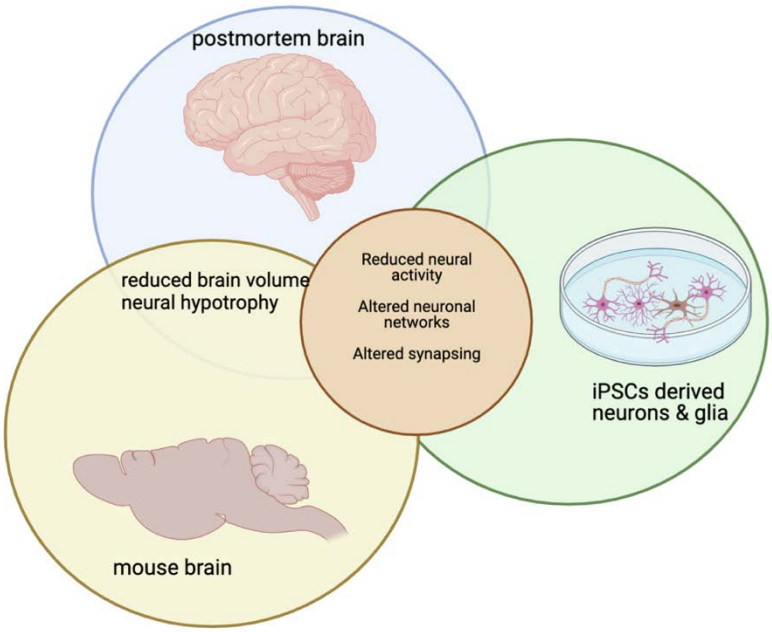
More recently, stem cells, including human embryonic stem cells (hECSs) and induced pluripotent stem cells (iPSCs), have come to play an important role in in vitro disease modelling. hECs are generated from early-stage human embryos and have the potential to differentiate into various cell types, whereas iPSCs are derived from patients and can be differentiated into any cell type [12]. Reprogramming of somatic cells to iPSCs through the overexpression of transcription factors was demonstrated over a decade ago [13] and this technology has now strengthened the utility of stem-cell-based disease models [14]. Over the past few years, several studies have successfully generated iPSC lines from RTT fibroblasts and have differentiated these lines into neural progenitor cells (NPCs), neurons and glial cells [11, 12, 15-18]. Stem-cell-based modelling has been demonstrated to be effective for RTT research, because iPSC lines can harbour pathogenic *MECP2* mutations and thus can demonstrate neuronal morphological defects, such as reduced dendritic branching, spine density and smaller soma size [19]. Several studies have reported differentiated neuronal cells from RTT-iPSCs in two-dimensional (2D) cultures, with a smaller soma size compared to that of controls [12, 20, 21]. Additionally, the dysregulation in cellular maturation and morphological complexities in RTT-iPSC neurons have recapitulated the findings of mouse studies and in human post-mortem brain tissues [22].

The complexities of RTT at a clinical level and MeCP2 function have resulted in significant challenges for developing safe and effective therapies [23]. It is unclear whether novel therapies that have shown promising preclinical efficacy would effectively mitigate systemic manifestations of the disease when administered in the clinic; this is due in part to the lack of models that cover all aspects of disease. Thus, well-characterised, disease-relevant models are critical to uncovering the underlying molecular, cellular, and physiological intermediate phenotypes in the pathophysiology of RTT that may provide insights into potential therapies. Therefore, we hypothesise that by taking advantage of all the existing models, both old and new, (Figure 1), useful insights into the pathophysiology of RTT may be gleaned, that will drive the discovery of novel therapeutic targets. To do this, we resorted to a meta-analysis of the transcriptomic data from three different

 doi:10.20944/preprints202207.0030.v1

RTT models: mouse brain, postpost-mortem human brain tissue and iPSC-derived neurons. Weighted Gene Correlation Network Analysis (WGCNA) offers a powerful method to untangle novel disease pathways compared to approaches such as differential gene expression. In this study we used WGCNA to examine three previously published transcriptomic datasets of human post-mortem brain tissue, iPSC-derived neurons and mouse brain samples. After identifying a consensus module between the three datasets, we analysed the genes in that module against another two datasets that could not be included in the WGCNA analysis, using differential gene expression.

Figure 1. Molecular phenotype shared by mouse brain, post-mortem brain and iPSC-derived neurons and glia.



2. Results

2.1. Data Pre-Processing and Identification of Common Genes

Publicly available genome-wide transcriptomic datasets of iPSC-derived neurons, post-mortem human brain tissue, and mouse brains were retrieved from the NCBI Gene Expression Omnibus database. These included: GSE75303 (post-mortem), GSE123753 (iPSC-derived neurons) [24] and GSE96684 (mouse brain) [25] (Table 1). The post-mortem and mouse datasets included RTT and wild type samples, whereas the iPSC-derived neurons included RTT and isogenic controls. The post-mortem dataset included sequencing results from both the temporal and frontal cortex and the age of the patients ranged from 17 to 20 years, and all subjects were female, harbouring three different mutations: c.378-2A>G, c.763C>T and c.451G>T. The mouse samples were from the brain cortex and were all were MECP2 knockout males [19]. The iPSC-derived neurons were females harbouring a deletion between exons 3 and 4 of MECP2. All samples were included in this study.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 July 2022 doi:10.20944/preprints202207.0030.v1

Table 1. Summary of samples used in WGCNA. WT refers to wild type and MT refers to mutant.

Study	Sample	Age	Gender	Tissue	Disease State	Mutation
Post-mortem human brain GSE75303	GSM1949097	19y 231d	F	frontal cortex	WT	NA
	GSM1949098	17y 28d	F	frontal cortex	WT	NA
	GSM1949099	17y 28d	F	temporal cortex	WT	NA
	GSM1949100	20y 228d	F	temporal cortex	WT	NA
	GSM1949101	18y 138d	F	frontal cortex	WT	NA
	GSM1949102	18y 138d	F	temporal cortex	WT	NA
	GSM1949103	18y 130d	F	frontal cortex	RTT	c.378-2A>G
	GSM1949104	18y 130d	F	temporal cortex	RTT	c.378-2A>G
	GSM1949105	20y 356d	F	frontal cortex	RTT	c.763C>T
	GSM1949106	20y 356d	F	temporal cortex	RTT	c.763C>T
	GSM1949107	19y 280d	F	frontal cortex	RTT	c.451G>T
	GSM1949108	19y 280d	F	temporal cortex	RTT	c.451G>T
iPSC-derived neurons GSE123753	GSM3510829	NA	F	neurons	WT	isogenic
	GSM3510835	NA	F	neurons	MT	Exon 3-4 deletion
	GSM3510857	NA	F	neurons	WT	isogenic
	GSM3510863	NA	F	neurons	MT	Exon 3-4 deletion
	GSM3510877	NA	F	neurons	WT	isogenic
	GSM3510883	NA	F	neurons	MT	exon 3-4 deletion
Mouse brain GSE96684	GSM2538276	P60	M	cortex	WT	NA
	GSM2538277	P60	M	cortex	WT	NA
	GSM2538278	P60	M	cortex	WT	NA
	GSM2538279	P60	M	cortex	WT	NA
	GSM2538280	P60	M	cortex	MT	R168X
	GSM2538281	P60	M	cortex	MT	R168X
	GSM2538282	P60	M	cortex	MT	R168X
	GSM2538283	P60	M	cortex	MT	R168X

The data was normalised and filtered prior to WGCNA analysis (Figure 2). To do this, abnormal samples were first filtered through hierarchical clustering where any missing count data was eliminated. The genes in the mouse dataset were homologated to the human genome, with only the common genes being included in the study. Overall, there were a total of 9864 genes included in this analysis (Figure 2).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 July 2022 doi:10.20944/preprints202207.0030.v1

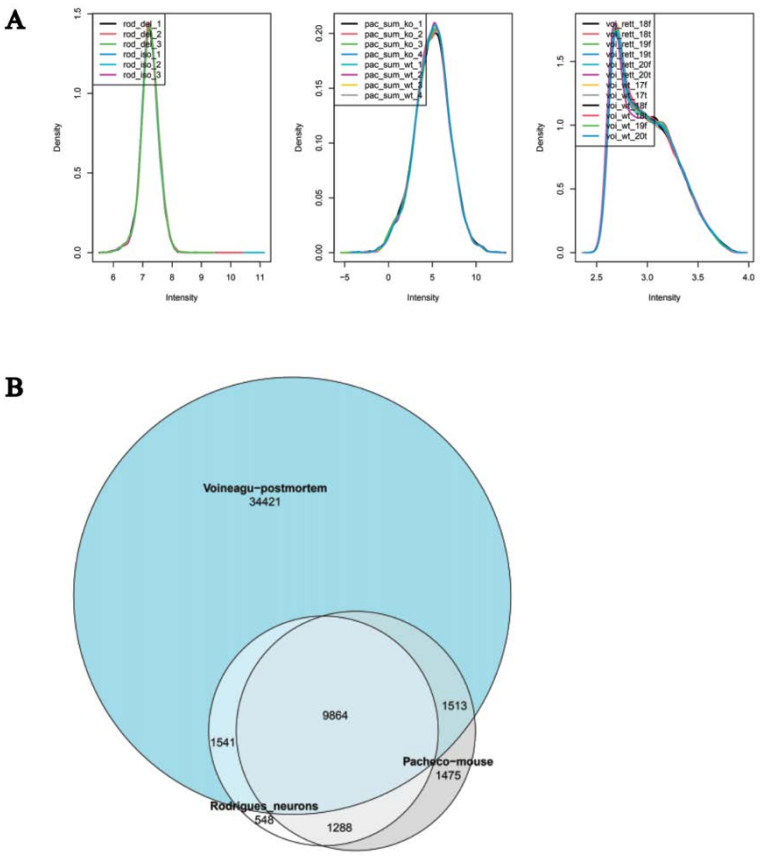


Figure 2. Data pre-processing. (**a**) Normalisation of datasets. Each panel represents a study in the following order (left to right): Rodrigues, Pacheco and Voineagu, each color line represents a sample. (**b**) Venn diagram showing genes in common between the three datasets.

2.2. Weighted Gene Co-Expression Networks

Weighted gene co-expression networks were constructed based on the identified genes following the soft threshold analysis using all three datasets combined. An optimal soft-thresholding power is needed to calculate co-expression similarity. Hence, to assess the similarity between genes at the expression and network topology levels we created a topological overlap matrix (TOM) which was achieved by calculating the adjacency and correlation matrices of the gene expression profile. As shown in Figure 3A in the scale free topology plot, power 8 was the lowest power where all three datasets reached a topology fit index of 0.9, hence it was chosen to produce the hierarchical clustering tree (dendrogram). Using the hierarchical average linkage clustering method in combination with the TOM, we proceeded to identify gene modules of each gene network. The dynamic tree cut algorithm highlighted all gene modules and each was identified by a colour (Figure 3B), each tree branch constitutes a module and each leaf in the branch is one gene.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 July 2022 doi:10.20944/preprints202207.0030.v1

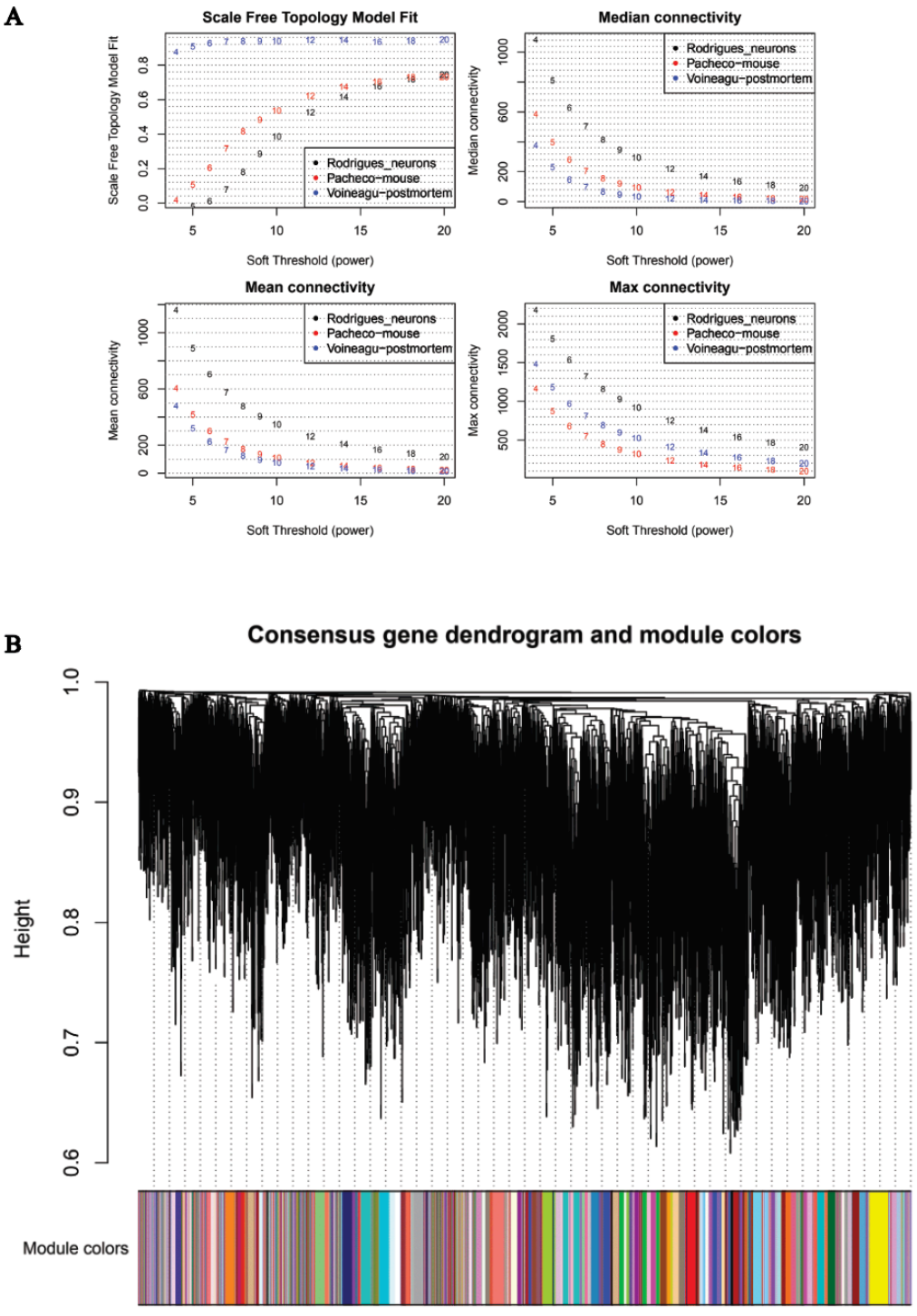
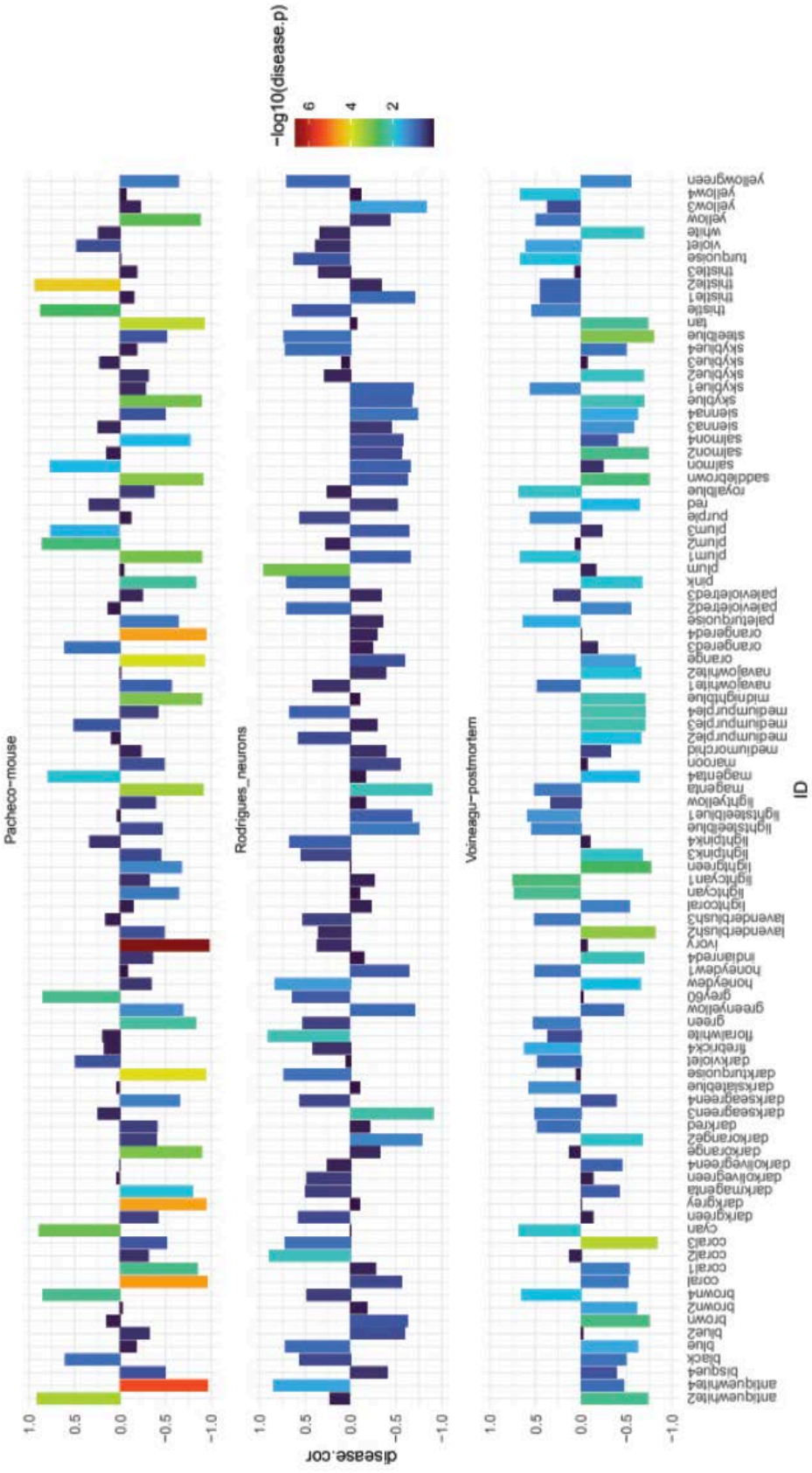


Figure 3.WGCNA analysis of the three datasets. (**a**) Analysis of network topology as a function of the soft-thresholding power for all three datasets. The panels show the scale-free fit index (top left), median connectivity (in degrees, top right), mean connectivity (in degrees, bottom left) and the maximum connectivity (in degrees, bottom right). (**b**) Clustering dendrogram of genes. Gene clustering tree (dendrogram) obtained by hierarchical clustering of adjacency-based dissimilarity. The coloured row below the dendrogram indicates module membership identified by the dynamic tree cut method, together with assigned merged module colours and the original module colours.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 July 2022 doi:10.20944/preprints202207.0030.v1

2.3. Correlation Between Modules and Clinical Traits

The module-trait associations were analysed by correlating module-sample eigengenes with clinical traits to identify significant associations. The colours of all the modules were selected at random to distinguish between modules. Correlation coefficients were assigned to each module and the disease status trait (RTT vs wild type (WT). Subsequently, only modules with a significant correlation to the disease trait (p-value < 0.0.5) were identified for all three datasets (Figure 4).



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 July 2022 doi:10.20944/preprints202207.0030.v1

Figure 4. Module-feature associations. Each row corresponds to a module eigengene and its correlation with the clinical phenotype (disease status). Correlation coefficient is represented in log 10 scale, where blue corresponds to a negative one correlation coefficient, and red corresponds to positive one correlation coefficient.

2.4. Different Brain Tissues/Models Generate Modules of Dysregulated Genes

As we were interested in common modules across all models and tissues with a significant correlation coefficient to the disease trait in the three datasets, we found four modules to be significantly dysregulated across all three datasets: brown4, blue, magenta and skyblue (Figure 4).

2.5. Module Analysis

To better understand the biological functions of the genes in the four modules, each module was subjected to KEGG pathway enrichment analysis (Figure 5). The level of significance of each pathway enrichment was calculated and expressed in adjusted p-values using the Bonferroni correction method. We then focussed on those pathways that had higher adjusted p-values (depicted in yellow in Figure 5).

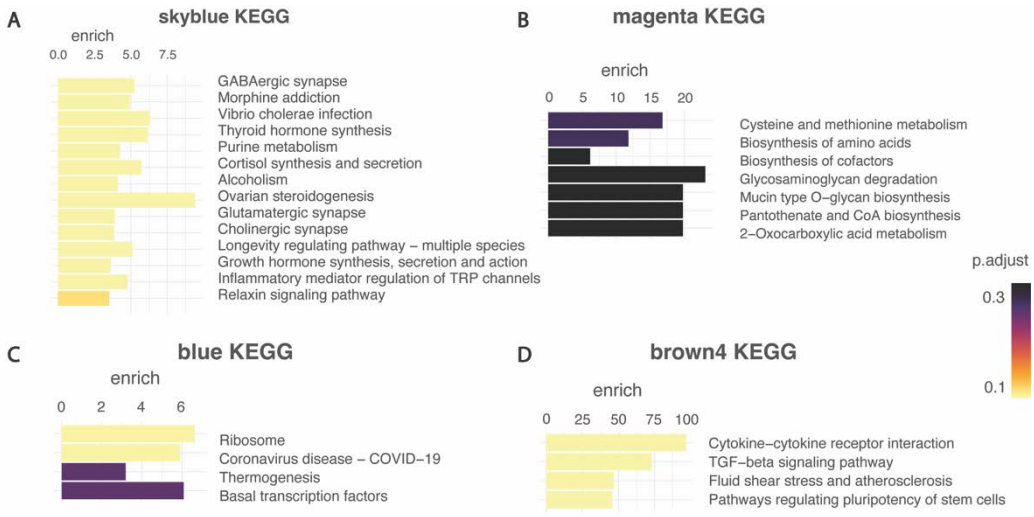


Figure 5. Enrichment analysis in interesting modules. (**a**) Skyblue module, (**b**) magenta module, (**c**) blue modules (**d**) brown4 module. Results include level of significance of each pathway enrichment using The Kyoto Encyclopedia for Genes and Genomes (KEGG) calculated and expressed in adjusted p-value, yellow represents more significant and purple least significant.

The brown4 module was highly enriched in cytokine-cytokine receptor interaction, the TGF-beta signalling pathway, fluid shear stress and atherosclerosis, and signalling pathways regulating pluripotency of stem cells. The blue module was highly enriched in pathways including ribosomes, Covid-19 disease, thermogenesis and basal transcription factors. The magenta module was enriched in cysteines and methionine metabolism, biosynthesis of amino acids, biosynthesis of cofactors, glycosaminoglycan degradation, mucin-type 0-glycan biosynthesis, pantothenate and CoA biosynthesis, and 2-oxocarboxylic acid metabolism. Finally, the skyblue module was enriched in GABAergic synapses, morphine addiction, vibrio cholerate infection, thyroid hormone synthesis, purine metabolism, cortisol synthesis and secretion, alcoholism, ovarian steroidogenesis, glutamatergic synapsis, cholinergic synapse, the longevity regulating pathway, growth hormone synthesis, secretion, and action, inflammatory mediator regulation of transient receptor potential (TRP) channels and relaxin signalling.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 July 2022 doi:10.20944/preprints202207.0030.v1

2.6 Key Cellular Pathways Involved in Synapses Dysregulated in Rett Models

Given the relevance to the known pathophysiology of RTT of the pathways identified in the skyblue module, we investigated this module further. Furthermore, focussing on the disease trait (WT vs RTT), the skyblue module exhibited the highest correlation (Figure 4) and more disease-relevant enrichment (Figure 5); therefore, this module was identified as a key module in RTT and was subjected to further analysis.

Interestingly, we found that the skyblue module was driven by hub genes: MECP2, BDNF, SOD1, PLCG1, CCT5, RPS9, ADCY9, ADCY7, ATRX and CACNA1A (Figure 6 and Table 2). Hub genes are defined as genes with connectivity (degree) greater than 10 in the genetic interaction network. All genes were shown to be downregulated in RTT except for CCT5 and ADCY9.

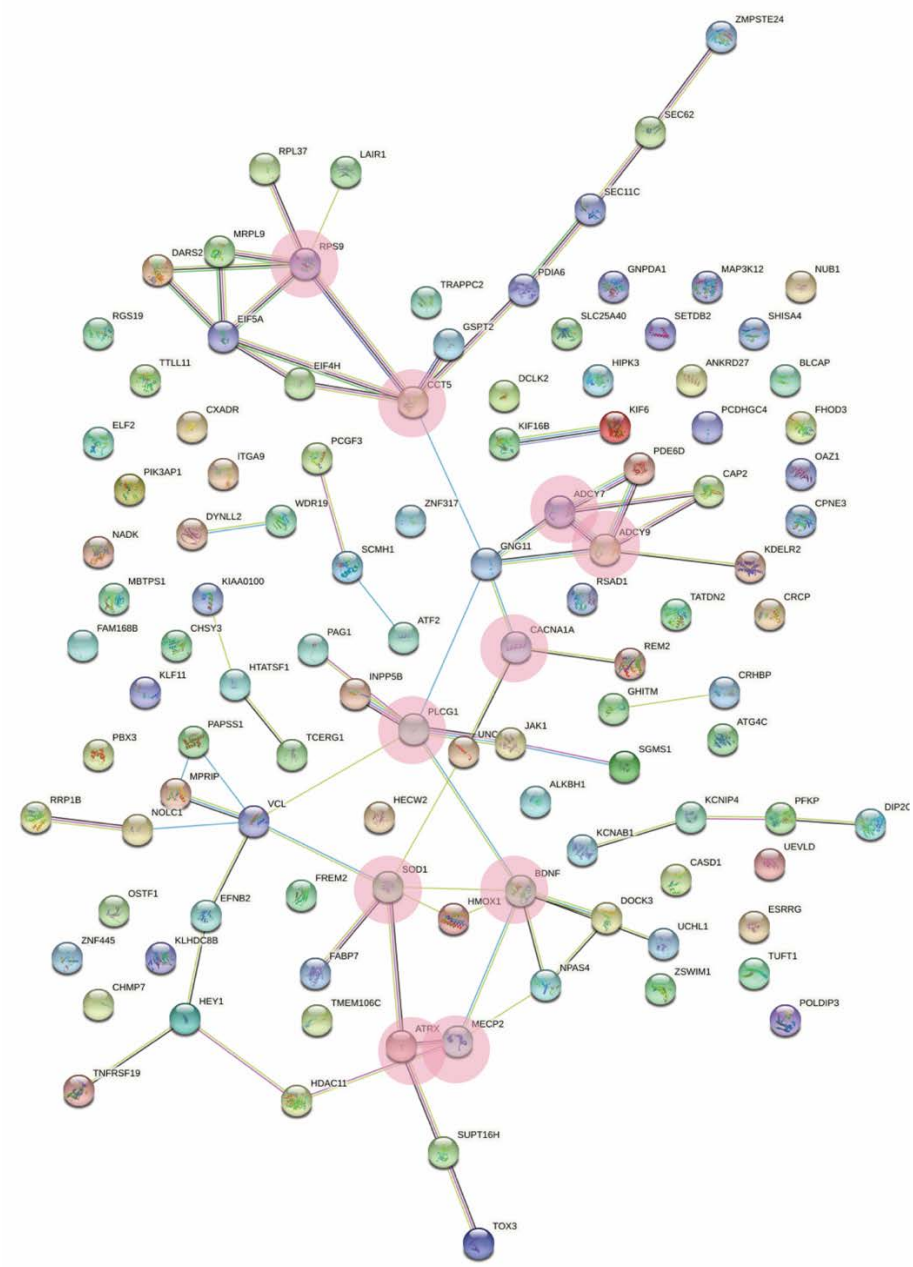


Figure 6. String diagram depicting the gene network of the skyblue module. Pink circles depict the identified hub genes.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 July 2022 doi:10.20944/preprints202207.0030.v1

Table 2. Hub genes and functions. Information adapted from GeneCards

Gene symbol	Gene Function
MECP2	Methyl-CpG-binding protein 2; a chromatin-associated protein that can both activate and repress transcription. It is required for maturation of neurons and is developmentally regulated.
BDNF	Brain-derived neurotrophic factor; during development, promotes the survival and differentiation of selected neuronal populations of the peripheral and central nervous systems. Participates in axonal growth, pathfinding and in the modulation of dendritic growth and morphology. Major regulator of synaptic transmission and plasticity at adult synapses in many regions of the central nervous system (CNS). The versatility of BDNF is emphasised by its contribution to a range of adaptive neuronal responses including long-term potentiation (LTP), long-term depression (LTD), certain forms of short-term synaptic plasticity.
CCT5	T-complex protein 1 subunit epsilon; a molecular chaperone that assists the folding of proteins upon ATP hydrolysis. As part of the BBS/CCT protein complex it may play a role in the assembly of BBSome, a complex involved in ciliogenesis, regulating transport vesicles to the cilia. Known to play a role in vitro in the folding of actin and tubulin.
CACNA1A	Voltage-dependent P/Q-type calcium channel subunit alpha-1A; voltage-sensitive calcium channels (VSCC) mediate the entry of calcium ions into excitable cells and are also involved in a variety of calcium-dependent processes, including muscle contraction, hormone or neurotransmitter release, gene expression, cell motility, cell division and cell death. The isoform alpha-1A gives rise to P and/or Q-type calcium currents. P/Q-type calcium channels belong to the 'high-voltage activated' (HVA) group and are blocked by the funnel toxin (Ftx) and by omega-agatoxin- IVA (omega-Aga-IVA).
ADCY9	Adenylate cyclase type 9; an adenylyl cyclase that catalyses the formation of the signalling molecule cAMP in response to activation of G-protein-coupled receptors. Contributes to signalling cascades activated by CRH (corticotropin-releasing factor), corticosteroids and beta-adrenergic receptors.
ADCY7	Adenylate cyclase type 7; a membrane-bound, calcium-inhibitable adenylyl cyclase.
ATRX	Transcriptional regulator ATRX; involved in transcriptional regulation and chromatin remodelling. Facilitates DNA replication in multiple cellular environments and is required for efficient replication of a subset of genomic loci. Binds to DNA tandem repeat sequences in both telomeres and euchromatin, and in vitro binds DNA quadruplex structures. May helpin stabilising G-rich regions into regular chromatin structures by remodelling G4 DNA and incorporating H3.3-containing nucleosomes. Catalytic component of the chromatin remodelling complex ATRX:DAXX, which has ATP-dependent DNA translocase activity.
RPS9	Small subunit ribosomal protein s9e; ribosomal protein S9.
SOD1	Superoxide dismutase [Cu-Zn]; destroys radicals that are normally produced within the cells and toxic to biological systems.
PLCG1	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase gamma-1; mediates the production of the second messenger molecules diacylglycerol (DAG) and inositol 1,4,5-trisphosphate (IP3). Plays an important role in the regulation of intracellular signalling cascades. Becomes activated in response to ligand-mediated activation of receptor-type tyrosine kinases, such as PDGFRA, PDGFRB, FGFR1, FGFR2, FGFR3 and FGFR4. Plays a role in actin reorganisation and cell migration.

2.7 Differential gene expression in the OH and MT datasets

Differences in gene expression associated with RTT were also explored by analysing expression profiles obtained from the MT (GSE6955) and OH (GSE107399) datasets

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 July 2022 doi:10.20944/preprints202207.0030.v1

(Figure 7). Among the 12,625 genes identified in the MT dataset, 156 were significantly upregulated and 61 significantly downregulated ($p<0.05$). Conversely, analysis of the OH data demonstrated 20,055 differentially expressed genes. Of these, 655 were significantly upregulated, and 992 were significantly downregulated ($p<0.05$).

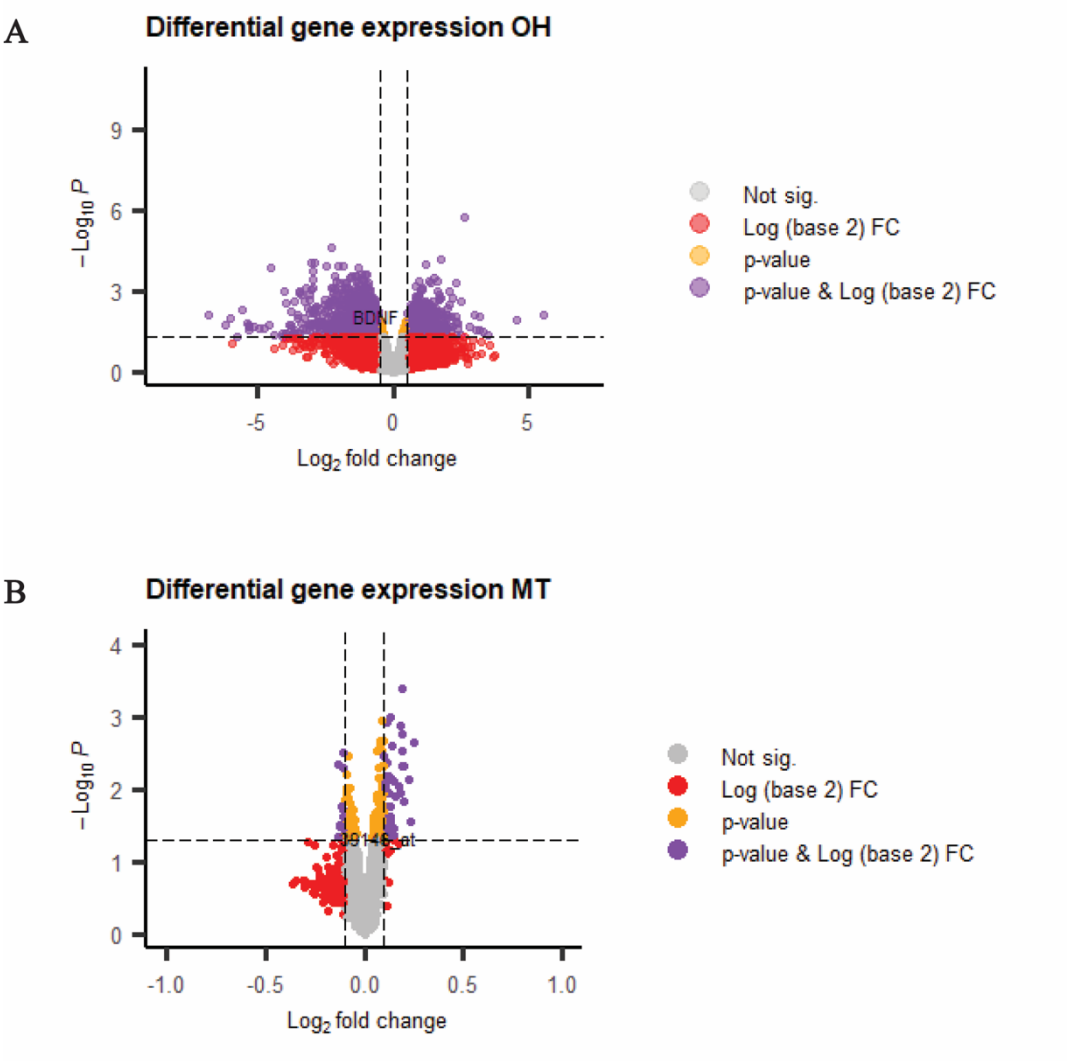


Figure 7. Volcano plot showing differential gene expression between RTT samples and controls from OH and MT dataset. (**a**) Volcano plot of differentially expressed genes in the MT dataset comparing brain tissue from patients with RTT and normal age-matched controls. The log2FC (x-axis) of each gene is plotted against the -log10P (y-axis). Expression difference is considered significant for log2FC of 0.1, and p-value < 0.05, as indicated by the purple-coloured points. The grey points represent genes with non-significant differences in expression. (**b**) Volcano plot of differentially expressed genes between patient-derived neuronal samples and controls in the OH dataset. The log2FC (x-axis) of each gene is plotted against the -log10P (y-axis). Significant differences in gene expression, where log2FC is 0.5, and p-value < 0.05, are depicted by the purple points. The grey points represent genes with non-significant differences in expression.

Next the gene expression profiles of MT and OH were then cross-referenced with those identified in the skyblue module to investigate any similarities. Overall, there were 71 genes shared between the skyblue module and the MT dataset. TOX3, FABP7, ATRX and SGMS1 had the largest positive log-fold changes of 0.13, 0.07, 0.07, 0.06 and 0.06, respectively. BDNF, GNG11, FAM168B, HMOX1 and VCL were identified as having the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 July 2022 doi:10.20944/preprints202207.0030.v1

largest negative log-fold changes with values of -0.07, -0.05, -0.04, -0.03 and -0.03, respectively. Additionally, 107 genes were commonly expressed in the skyblue module and OH data. The top five upregulated genes with the greatest positive log-fold change were NPAS4, FABP7, HECW2, TOX3 and CACNA1A with values of 2.16, 1.01, 0.92, 0.89 and 0.71, respectively. Conversely, FREM2, BDNF, HMOX1, KDELR2 and CXADR had the highest negative log-fold changes of -0.80, -0.62, -0.61, -0.59 and -0.46, respectively.

The expression of the hub genes identified in the skyblue module was also cross-examined in the MT and OH datasets (Figure 8). From the meta-analysis, CCT5 and ADCY7 were upregulated in the skyblue module, whilst the remaining eight (MECP2, BDNF, CACNA1A, ADCY9, ATRX, RPS9, SOD1 and PLCG1) were downregulated. When compared with the differential gene expression data from MT, ATRX was significantly upregulated ($p<0.05$) with a log-fold change of 0.07 (Figure 8A). Meanwhile, in the OH and MT data, BDNF was significantly downregulated ($p<0.05$) with a -0.62 log-fold change (Figure 8B). CACN1A was shown to be upregulated in OH. The remaining genes from skyblue showed non-significant changes in MT and OH.

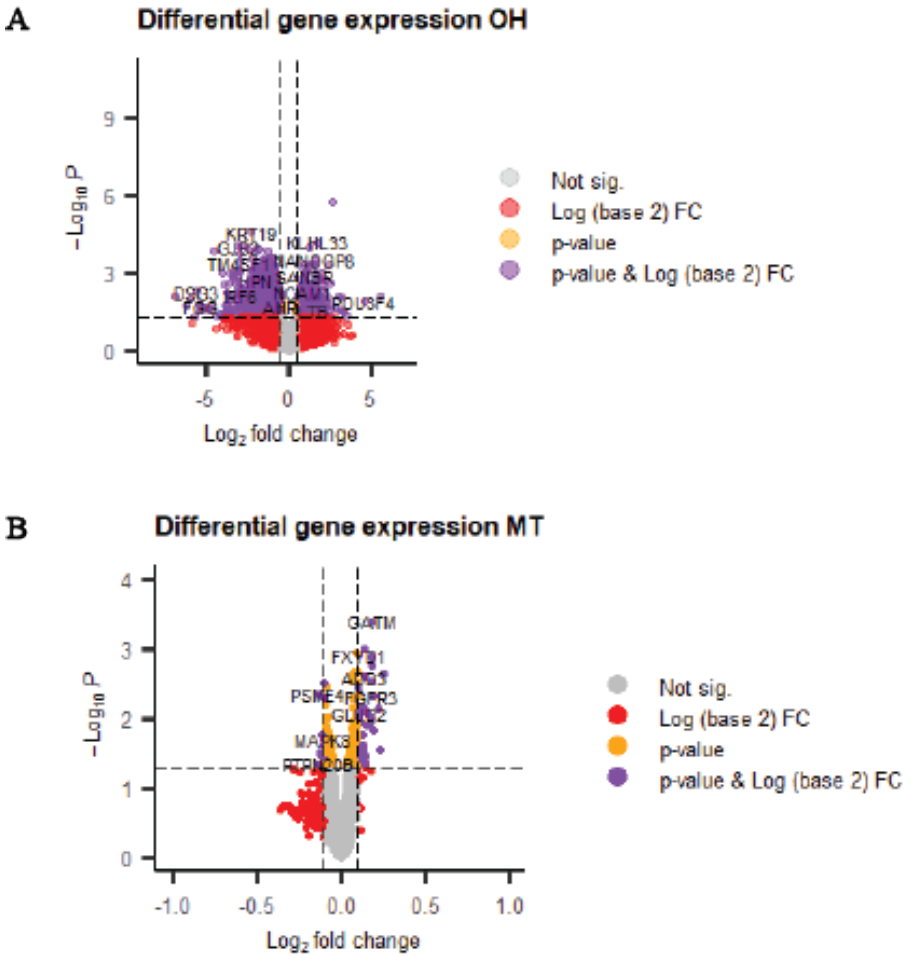


Figure 8. Expression of hub genes from skyblue module in MT and OH datasets. (**a**) Volcano plot of differential gene expression profiles in the MT data highlighting hub genes with significant changes. ATRX, BDNF, ADCY7, PLCG1, RPS9 and ADCY9 are depicted. (**b**) Volcano plot of differentially expressed genes in the OH dataset, highlighting the hub genes BDNF, CACNA1A, MECP2 and PLCG1. Genes with significant changes in logFC ($p<0.05$) are denoted by purple points. Genes with no significant changes in expression are represented by grey points.

 doi:10.20944/preprints202207.0030.v1

2.8 Hub gene PCR validation

To validate the expression of the genes identified through the in silico analysis, we performed RT-qPCR in iPSC-derived neurons from a male RTT patient and a paedritic control as well as a male post mortem forebrain sample and a control (Figure 9). From the ten hub genes identified in this study, five were chosen given the known relevance to RTT. MECP2 and BDNF were excluded given their known and well reported association with RTT. Interestingly, ADCY7 was found to be upregulated in the RTT brain sample compared to the iNeurons and ADCY9 was downregulated in the brain samples compared to the iNeurons. In the meta-analysis ADCY9 was also found to be upregulated in RTT but not ADCY7. CACNA1 was found to be upregulated in the RTT iNeurons like the results found in the OH dataset from iPSC-derived neurons. ATRX and SOD1 were upregulated in both models whereas in the meta-analysis it was shown to be downregulated.

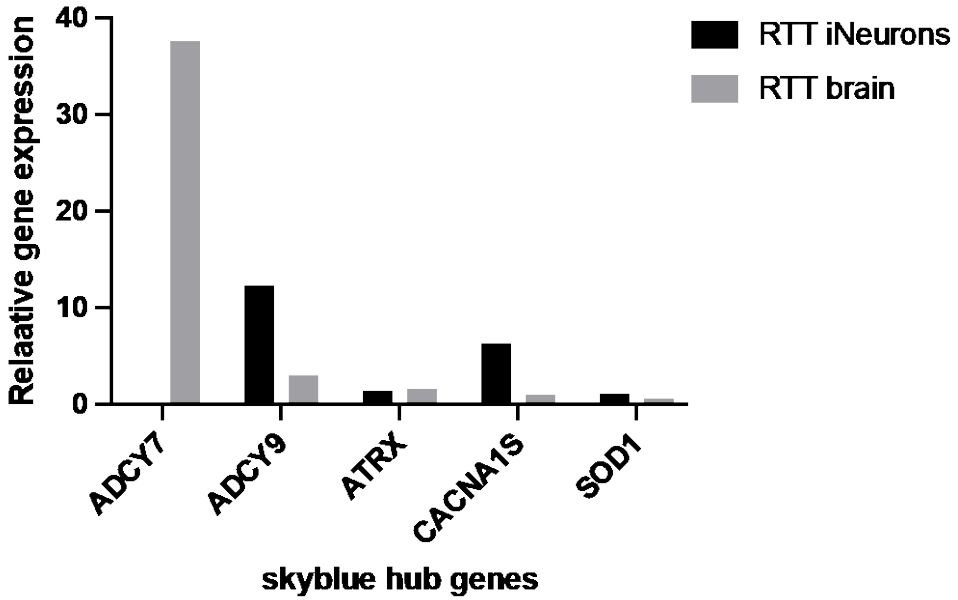


Figure 9. Relative gene expression levels (Delta Ct value assessed by RT qPCR) of five hub genes identified through the WGCNA meta-analysis in RTT iNeurons and RTT postmortem brain tissue normalised to the control cell lines and GAPDH (housekeeping gene). Data is presented as n=2.

3. Discussion

The overarching aim of this study was to identify common pathways and genes that intersect RTT transcriptomic studies spanning different species and models with the premise of identifying universal and robust disease drivers and therapeutic targets. To do so, a meta-analysis and bioinformatics approach consisting of the identification of gene modules rather than differential gene expression was employed to interrogate the transcriptomic landscape of RTT using human post-mortem brain tissue, mouse models and patient-derived iNeurons.

After identifying the statistically significant dysregulated modules between all the RTT samples and the controls, the module that had the highest correlation to disease status was interrogated as well as the genes that had the highest connectivity within the module to identify the main genetic drivers across all tissue samples and models. Reassuringly, the identified hub genes included MECP2 and BDNF, where the correlation

 doi:10.20944/preprints202207.0030.v1

between the two genes in RTT is well recognised, with BNDF being a well-established target gene of MeCP2 [26, 27].

3.1 Meta-analyses produced four significant modules correlated to disease status

Through this meta-analysis four modules of genes that were significantly dysregulated in the RTT transcriptome relative to the controls was identified. The pathways that were enriched in each of the four modules were investigated and it was identified that the brown4 module was mostly enriched in pathways related to immunological aberrations, which is consistent with previously published studies in RTT including one of our own [28, 29], [30]. Next, the magenta module was enriched for pathways mostly involving the metabolic system, which also aligns with previously reported literature [31, 32] . On the other hand, enrichment of the blue module did not produce any pathways known to be of relevance to RTT. The fourth module, skyblue consisted of enriched modules including glutamatergic, GABAergic and cholinergic synaptic pathways as well as protein export, and was identified to have the most enriched pathways relevant to the neuropathology of RTT supporting our focus into this module further.

3.2 Meta-analysis Hub genes within the skyblue module are relevant to RTT pathology

The ten hub genes that were identified as the main drivers of the skyblue module were ATRX, ADCY7, ADCY9, SOD1, CACNA1A, PLCG1, CCT5, RPS9, BDNF and MECP2 and therefore surmised to play key roles in the pathology of RTT and may assist in the understanding of the underlaying pathophysiology as well as in the identification of disease drivers and drug targets.

ATRX (ATRX Chromatin Remodeler) has recently been implicated in RTT as a binding partner of MeCP2 where together they modulate pericentric heterochromatin (PCH) organization in neurons [33]. Mutations in ATRX cause ATR-X syndrome, implicated in abnormal brain development and associated with severe intellectual disability [34]. The downregulation of ATRX in this meta-analysis, supports previous reports of an interaction with MeCP2, where MeCP2 recruits the helicase domain of ATRX to heterochromatic foci in a DNA methylation dependent manner, as shown in living mouse cells [35]. Furthermore, it has been shown that the heterochromatin location of ATRX is disrupted in Mecp2-null mice neurons These data together suggest that a MeCP2-ATRX interaction leads to pathological changes that contribute to the mental retardation phenotype. Interestingly, as an epigenetic modifier, ATRX has been implicated in cancer and has received a level of attention in the identification of expression modifying drugs [36]. ATRX loss leads to increased DNA damage, and general genomic instability [37], and thus drugs or small molecules aimed at increasing the stability of the genome may be potential therapeutic options for RTT.

Adenylate Cyclases 7 and 9 (ADCY7 and ADCY9) are membrane-bound enzymes that catalyse the formation of cyclic AMP from ATP and are highly expressed in the brain. De novo mutations in ADCY7 have been reported in Autism Spectrum Disorders (ASD) where the gene has been proposed to be a risk factor [38]. ASD and RTT share some commonalities with RTT individuals showing some ASD-like behaviours [39, 40]. ADCY7 mRNA is highly expressed in microglia and plays an important role in presynaptic GABA release and evidence suggests that ADCY7 is involved in mood regulation and plays an essential role in the immune response [41]. Conversely, despite ADCY9 being highly expressed in the brain, its function in the CNS remains largely unknown, however some findings have suggested that ADCY9 may regulate cognitive function

 doi:10.20944/preprints202207.0030.v1

and learning and memory [41]. Interestingly, ADCY9 has been shown to be downregulated in Mecp2 null embryonic cortexes, suggesting ADCY9 as a target of MeCP2 [42]. This effect is lost postnatally suggesting ADCY9's crucial role in embryogenesis [35, 41] . Interestingly, there are two common pathways that both genes are involved in, namely the GPER1 signalling and Integrin Pathway and these pathways could provide potential drug targets in which to target downstream pathways in RTT.

SOD1 plays a crucial role in the oxidative stress response and systemic redox alterations and the related oxidative stress is well reported in RTT [28, 43]. It is therefore not surprising to find the free radical scavenger SOD1 enzyme downregulated in this meta-analysis. Loss of SOD1 has been hypothesised to result in an accumulation of mitochondrial reactive oxygen species, leading to oxidative damage and mitochondrial dysfunction [44]. Animal studies have suggested a possible direct correlation between Mecp2 mutations and increased ROS levels, and the debate continues whether oxidative stress is a cause or consequence of RTT.

The Voltage-dependent P/Q-type calcium channel subunit alpha-1A (CACNA1A) gene has been implicated in epileptic encephalopathy, familial hemiplegic migraine, episodic ataxia, and spinocerebellar ataxia [45, 46] and has recently been reported in a small number of atypical Rett patients previously lacking known genetic mutations [47]. Voltage-sensitive calcium channels mediate the entry of calcium ions into excitatory neurons and are also involved in a variety of calcium-dependent processes, neurotransmitter release. Our findings suggest that the downregulation of CACNA1A in this meta-analysis may be contributing to the epileptic encephalopathy of RTT. Among its related pathways are the CREB and Integrin pathways.

PLCG1 (Phospholipase C, gamma 1), also known as, is a protein involved in cell growth, migration, apoptosis, and proliferation. Among its related pathways are the CREB pathways (like CACNA1A), GDNF-Family Ligands and Receptor Interactions. Even though no direct link to MECP2 has been reported in the literature, it is known that activation of the neurotrophin receptor TRKB by BDNF triggers downstream PLCG1 signalling [48, 49].

No direct relation has been reported between CCT5 and RPS9 and MECP2. However, CCT5 is implicated in the cellular pathways related to trafficking to the periciliary membrane and cell cycle and has also been linked to intellectual disabilities and early onset motor neuropathies [50, 51]. On the other hand, RPS9 is linked to RNA binding and structural constituent of ribosome, and as ribosomal dysfunction has been previously reported in RTT iNeurons by Rodrigues et al. 2021 [24], the dysregulation of RPS9 in this study supports these findings and provides further evidence of ribosomal dysfunction in RTT.

In addition, common cellular pathways such as the CREB and integrin signalling pathways are common amongst the hub genes. The CREB pathway has previously been reported to be implicated in RTT where overexpression of CREB signalling in RTT forebrain neurons rescued the phenotype of neurite growth, dendritic complexity, and mitochondrial function [52]. Furthermore, pharmacological activation of CREB in female RTT mice rescued several behavioral phenotypes [52]. These findings support the motion to investigate the CREB pathway as a potential therapeutic target [52]. In addition, while the integrin pathway has not been reported in RTT, it has been previously implicated in dendritic development, autism spectrum disorder and intellectual disabilities [53, 54] suggesting that pathway too could also be a potential target for future RTT therapeutics.

 doi:10.20944/preprints202207.0030.v1

3. 4 Meta-analysis shows commonly dysregulated synaptic pathways

Through this study, three synaptic pathways enriched in the skyblue module were identified, namely the cholinergic, glutamatergic, and GABAergic pathways. A loss of excitation/inhibition (E/I) balance in the neural circuit is a major hallmark of RTT pathology, causing many neurological symptoms such as loss of purposeful hand movements, impaired motor coordination, breathing irregularities and seizures amongst others [55]. This loss of E/I balance is caused by MeCP2 deficiency leading to dysregulation of the glutamatergic and GABAergic pathways, furthermore downstream genes affected in RTT such as BDNF play an important role influencing neurotransmission activity. Many drugs have been tested to improve the E/I balance in RTT including glutamatergic modulators such as AMPAkines to increase excitatory synapsis and enhance BDNF expression, ketamine, and NMDAR antagonist to enhance neuronal activity [56-58]. GABAergic modulators have also shown potential in aiding with behavioral dysfunction in RTT patients and mice, however, while respiratory alterations were ameliorated by treatment using benzodiazepines and Midazolan in mice, the phenotype was not fully rescued [59].

3. 5 Expression of overlapping genes in MT and OH in comparison to skyblue

A comparison of differentially expressed genes in the MT and the OH datasets with the skyBlue module identified 71 and 107 commonly expressed genes, respectively. Interestingly, TOX3 was upregulated in both datasets and the skyblue module. TOX3 plays a role in shaping DNA and altering chromatin structure and while the protein has been shown to be a neuron survival factor [60] it is yet to be linked with neurodevelopmental disorders and specifically to RTT. BDNF and HMOX1 were also commonly dysregulated in all the datasets where they were observed to be significantly downregulated.

HMOX1 is a heme oxygenase and responsible for the degradation of heme to biliverdin/bilirubin, free iron and heavily implicated in aging and disease. The expression of HMOX1 is confined to small populations of neurons and glia and is upregulated by a wide range of pro-oxidant and other stressors [61]. While there have been no reports linking HMOX1 to RTT pathology, its downregulation confirms the role of oxidative stress in the pathology of RTT. Conversely, BDNF, a previously identified hub gene for skyblue, has been widely associated with RTT. Additionally, two hub genes, ATRX and CACNA1A were identified to be dysregulated in the skyblue module as well in either the MT or OH datasets where the expression of ATRX was upregulated in the MT dataset and CACNA1A was upregulated in OH.

3. 6 Hub gene expression comparison across studies

Of the ten identified hub genes in the meta-analysis, 8 were downregulated suggesting that wild type MeCP2 transcriptionally activates these genes and two (ADCY9 and CCT5) were upregulated, suggesting that MeCP2 transcriptionally represses these genes.

From the differential gene expression analysis performed on the OH and MT datasets, we showed that BDNF was downregulated in both studies, ARTX was upregulated in MT and CACN1A was downregulated. Using RT-qPCR five of the identified hub genes relevant to RTT pathology were validated in RTT and control iPSC-derived neurons and postmortem brain. We showed an overall trend of upregulation in the five

 doi:10.20944/preprints202207.0030.v1

tested genes ATRX, ADCY7, ADCY9, CACNA1A and SOD1. These results were different to that found in the meta-analysis as only ADCY9 was upregulated. This disparity in expression between the meta-analysis, differential gene expression and RT-qPCR validation demonstrates the complexity of RTT and the context dependant expression of MECP2

The identification of BDNF as the only consistent gene to be downregulated in the meta-analysis and differential gene expression studies, comes as no surprise given the reported association with RTT. BDNF has been vastly explored as therapeutic target for RTT, however as BDNF has a low brain blood barrier permeability, this limits its bioavailability of peripheral administration as a therapy [62]. Three clinical trials aimed at augmenting BDNF expression, trialing Copaxone (glatiramer acetate) [63] and Fingolimod have been conducted[64]. However, as yet no therapies have entered and clinic with the glatiramer acetate trial being withdrawn due to reported potential life-threatening reactions in one of the clinical trials [65]. Additional compounds have been described to increase BDNF levels and improve RTT-like symptoms in mice, however none have reached human clinical trials, showing the difficulty of this approach [62].

The fact that BDNF was identified to be consistently downregulated in all studies consisting of two different analysis approaches (WGCNA and DGE) across different species and models shows the strength of the analysis tool. Importantly, these findings support the power of this bioinformatic approach and supports the significance of the identified hub genes through WGCNA.

4. Materials and Methods

4.1 Dataset Selection

The three datasets included in the WGCNA analysis were obtained from the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/): GSE75303 (Post-mortem), GSE123753 (iPSC-derived neurons) and GSE96684 (Mouse Brain). The GSE75303 dataset contained 12 samples in total, including three female RTT patient frontal and temporal cortexes harbouring mutations at c.378-2A>G, c.763C>T and c.451G>T and three female age-matched controls. The GSE123753 dataset consisted of six female samples: three patients involving rearrangements that removed exons 3 and 4, creating a functionally null mutation, and their three corresponding isogenic controls. The GSE96684 dataset consisted of eight male mouse samples: four MECP2 knockout and four wild type mice. The characteristics of the samples in each dataset are summarised in Table 1.

4.2 Dataset Pre-Processing

Since the three datasets were from different sequencing platforms, we performed pre-processing according to a previously published WGCNA pipeline. Briefly, raw counts and probe intensity data were pre-processed using the Limma package [66] in the R environment. Counts data were transformed by mean-variance modelling at the observational level (voom) [67] before all studies were subjected to quantile normalisation and data quality control as recommended for WGCNA. Finally, for differential gene expression analysis, raw count data were voom transformed and array data were log-transformed in the Limma package [66] in the R environment, followed by quantile normalisation and data quality control.

 doi:10.20944/preprints202207.0030.v1

4.3 Weighted Gene Correlation Analysis

Unsigned co-expression networks were built using the WGCNA 1.63 package in R software [68]. Clusters of genes that behaved similarly were grouped together into different colour modules. These modules were related to specific traits. In heatmaps, red represents genes upregulated within that dataset and green represents genes downregulated within that dataset. The top 1,000 connections within a gene network were determined by WGCNA. For the multiple array consensus analysis, WGCNA was performed on the individual datasets first as suggested by Langfelder and Horvath's tutorial [69], using "1 step function for network construction and detection of consensus modules". The default WGCNA soft thresholding power β in which co-expression was raised was chosen to calculate the adjacency of each data set. The soft thresholding power β was used to allow us to compare each data set by approximate scale-free topology, thus compensating for scale differences between data sets.

4.4 Module Selection

The correlation between module eigengenes and clinical traits was analysed to identify modules of interest that were significantly associated with clinical traits. For the purpose of this study, we identified the modules that were significantly correlated with disease status in all three datasets. The correlation values were then displayed within a heatmap. Gene significance (GS) was defined as the correlation between gene expression and each trait. In addition, module membership (MM) was defined as the association between gene expression and each module eigengene. Subsequently, the correlations between GS and MM were examined to verify certain module-trait associations. The correlation analyses in this study were performed using Pearson correlation as described in the WGCNA package [68].

4.5 Module Enrichment

The genes in each module of interest were extracted from the network and enrichment analysis was performed to further explore the functions of the respective modules. The R package 'clusterProfiler' was used to perform Kyoto Encyclopedia of Genes and Genomes (KEGG) [70, 71] pathway enrichment analysis. A statistical p value of <0.05 was set as the significance threshold, and the enrichment results of KEGG pathways in each module of interest module were obtained.

4.6 Module Visualisation and Identification of Hub Genes

The intramodular connectivity of genes in the corresponding modules of interest was measured using module eigengene-based connectivity (kME). The top 30 genes of each module of interest, which represent the central status in the module gene network, were selected to visualise the subordinate module using String software [72]. Subsequently, one key module was chosen that exhibited the highest levels of positive or negative correlation with RTT to search for hub genes for RTT in the modules. The top ten genes with the highest kME were selected as the hub genes in the corresponding module [68] and their gene significance (GS) for RTT (disease status) and intramodular connectivity kME were determined to confirm the reliability of these hub genes.

 doi:10.20944/preprints202207.0030.v1

4.7 Differential gene expression

Differential gene expression analysis was performed on six patient-derived datasets from the MT study (GSE6955) and six iPSC-derived neuronal samples from the OH study (GSE107399). Samples in the MT study were taken from the superior frontal gyrus of patients with RTT and age-matched controls. From the OH study, eight samples were utilised for analysis. Of these, four were RTT mutants, three isogenic controls, and one wild-type control. The datasets were analysed in R using the EdgeR package (R Bioconductor). Firstly, genes with low expression with a CPM value ≤1 were filtered. Then, the remaining counts were used to generate linear models. Statistical analysis was conducted using eBayes() and topTable() arguments. To identify overlapping differentially expressed genes from the OH and MT datasets corresponding to the skyblue module, the log fold change was noted for genes that overlapped using the function VLOOKUP() in Microsoft Excel.

4.8 Hub genes PCR validation NGN2 neuron differentiation

To validate the hub genes identified in silico, RT-qPCR was performed on iPSC-derived neurons and postmortem human brain. Firstly, excitatory neurons were generated as previously described [73]. Briefly, lentiviruses expressing the transcription factor Neurogenin 2 (NGN2) and reverse tetracycline transactivator (rtTA) genes were produced using the FUW-TetO-Ngn2-P2A-EGFP-T2A-puromycin and FUW-rtTA plasmids kindly gifted by Dr Simon Maksour and A/Prof Mirella Dottori (Illawarra Health and Medical Research Institute, University of Wollongong).

To package the lentivirus, HEK293T cells (7.5×106) were seeded in a T-75 flask with DMEM:F12 containing 10% FBS media. Two lentiviral preps were produced in HEK293T cells using three plasmids: pMDG.2, pRSV-Rev and pMDLg/pRRE, and co-transfected with either FUW-TetO-Ngn2-P2A-EGFP-T2A-puromycin or FUW-rtTA plasmids. Viral particles in the supernatant were collected 48 h post-transfection, filtered through a 0.45 μm filter, and concentrated at 30,000 × g centrifugation for 2 h at 4 °C. The supernatant was discarded, and the pellet resuspended in 50 μl PBS and left overnight at 4 °C. The next day, the solution was triturated by manual pipetting, aliquoted and frozen at −80 °C.

4.9 Cell Culture and RNA extraction

To generate excitatory neurons, a RTT male iPScs line harbouring a mutation at c.806delG and a neurotypical paedritic control, were dissociated into single cells with 50% Accutase (Cat# A1110501, ThermoFisher Scientific) in DPBS -/- (Cat# 14190086, ThermoFisher Scientific) and seeded at 5×105 cells per well in a 6-well plate, previously coated with 15 μg/ml rhLaminin-521 (Cat# A29248, ThermoFisher Scientific) diluted in DPBS +/+ (Cat# 14040133, ThermoFisher Scientific), in 2 ml of StemFlex (Cat# A3349401; ThermoFisher Scientific) supplemented with 10 μM of RevitaCell (Cat# 39 A2644501, ThermoFisher). On day one, the medium was changed and replaced with fresh media supplemented with10 μM RevitaCell. For lentiviral transduction, the medium was aspirated and replaced with fresh StemFlex media supplemented with RevitaCell, containing the NGN2 and rtTA lentiviruses at the appropriate titre (Table 2.5). After 20 h, the medium was changed and for NGN2 induction, 2 μg/ml doxycycline (DOX) (Cat# 39 9891 Sigma) was added to neural media (BrainPhys media 1× N-2 supplement, 1× B-27 with vitamin A supplement), with 10 μM RevitaCell. On day three, the medium was changed and freshly supplemented medium with 2 μg/ml DOX and 2 μg/mlPuromycin

 doi:10.20944/preprints202207.0030.v1

[74 A1113802, ThermoFisher Scientific] was added to initiate the puromycin selection process. The media were changed every day and this feeding regime continued for four days. On day seven, the puromycin selection was discontinued and the media replaced with neuronal media (BrainPhys media 1× N-2 supplement, 1× B-27 with vitamin A supplement) supplemented with 2 µg/ml DOX and 10 ng/µl BDNF. For inhibition of non-neuronal cell growth 10 µM cytarabine Ara-C (Cat# C1768, Sigma Aldrich) was added to the media. Finally, cells were detached using 0.5 mM EDTA on day 8 and replated into new 6-well plates fand harvested on day 16 for RNA extraction. RNA was isolated using a Qiagen RNeasy Mini Kit (Cat # 74104; Qiagen) as per the manufacturer's instructions for cultured cells. RT-qPCR was performed on cDNA extracted from the iNeurons and post-mortem human brain using the primers below (Table 3).

Table 3. Primers used to validate hub genes.

Gene Name	Forward (5′-3′)	Reverse (3′-5′)
CACNA1A	CTGGTAGCCTTTGCCTTCACTG	CTGCCAAAGCTCAAGGCTGTGT
ADCY9	CTCAAAACGGCTGCCAAGACGA	GCTCACCAACAGTCAGACTTCTC
ADCY7	GACGAGATGCTGTCAGCCATTG	CTTCCTGGAGAAGGGCTTTGAG
ATRX	ACGGCGTTAGTGGTTTGTCCTC	CAGGAGAGAAGCTACATGCTGC
SOD1	CTCACTCTCAGGAGACCATTGC	CTGGAAGTCGTTTGGCTTG

5. Conclusions

In conclusion, through this meta-analysis and sub-analysis using murine, post-mortem and iPSC-derived neurons, shared genes that drive the RTT pathology, were identified. Some genes have previously been reported to be linked to RTT such as BDNF, ADCY9, ATRX and CACNA1A while others including CCT5, RPSP and PLCG1 are novel findings Validating BDNF as a known target of MeCP2, demonstrates the power of this bioinformatic approach and highlights this novel approach in identifying genes targets for further therapeutic development. Interestingly, none of the hub genes in sky-blue were identified in the multi-transcriptomic study by Ehrhart et al who only analysed human samples to identify dysregulated genes and networks, suggesting the limitation of this study [75]. Further exploration of these known and novel genes may unravel molecular mechanisms of RTT and pave the way for novel therapies.

Author Contributions: FH developed the methods, analysed the data and wrote the manuscript, RS performed the differential expression analysis, BG provided bioinformatic direction and assisted with manuscript prepa- ration, PPLT and WG supervised the project and revised the manuscripts. All authors reviewed and approved the final manuscript.

Funding: This Project was funded by the trustee for Neil & Norma Hill Foundation

Institutional Review Board Statement: All procedures were in accordance with the ethical standards of the Sydney Children's Hospitals Network Human Research Ethics committee on human experimentation (institutional and national).

Informed Consent Statement: Informed consent was obtained from all legal guardians of each participant included in the study.

Data Availability Statement: Code provided on request.

Acknowledgments: We thank Sarah Alshamery for her contribution formatting the manuscript and overall continuous support.

Conflicts of Interest: The authors FH, BG, PPLT and WG declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Neul, J.L., et al., *Rett syndrome: revised diagnostic criteria and nomenclature.* Annals of neurology, 2010. **68**(6): p. 944-950.
2. Meehan, R., J.D. Lewis, and A.P. Bird, *Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA.* Nucleic acids research, 1992. **20**(19): p. 5085-5092.
3. Gabel, H.W., et al., Disruption of DNA-methylation-dependent long gene repression in Rett syndrome. Nature, 2015. **522**(7554): p. 89-93.
4. Guo, J.U., et al., Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. Nature neuroscience, 2014. **17**(2): p. 215-222.
5. Chahrour, M., et al., MeCP2, a key contributor to neurological disease, activates and represses transcription. Science, 2008. **320**(5880): p. 1224-1229.
6. Boxer, L.D., et al., MeCP2 Represses the Rate of Transcriptional Initiation of Highly Methylated Long Genes. Molecular Cell, 2020. **77**(2): p. 294-309.e9.
7. Young, J.I., et al., Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. Proc Natl Acad Sci U S A, 2005. **102**(49): p. 17551-8.
8. Krishnaraj, R., et al., Genome-wide transcriptomic and proteomic studies of Rett syndrome mouse models identify common signaling pathways and cellular functions as potential therapeutic targets. Human Mutation, 2019. **40**(12): p. 2184-2196.
9. Clowry, G., Z. Molnár, and P. Rakic, *Renewed focus on the developing human neocortex.* J Anat, 2010. **217**(4): p. 276-88.
10. Dolmetsch, R. and D.H. Geschwind, *The human brain in a dish: the promise of iPSC-derived neurons.* Cell, 2011. **145**(6): p. 831-4.
11. Hinz, L., et al., Generation of isogenic controls for in vitro disease modelling of X-chromosomal disorders. Stem cell reviews and reports, 2019. **15**(2): p. 276-285.
12. Marchetto, M.C., et al., A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. Cell, 2010. **143**(4): p. 527-539.
13. Takahashi, K. and S. Yamanaka, Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell, 2006. **126**(4): p. 663-76.
14. Yamanaka, S., *Induced Pluripotent Stem Cells: Past, Present, and Future.* Cell Stem Cell, 2012. **10**(6): p. 678-684.
15. Tang, X., et al., *KCC2 rescues functional deficits in human neurons derived from patients with Rett syndrome.* Proceedings of the National Academy of Sciences of the United States of America, 2016. **113**(3): p. 751-756.
16. Kim, J.J., et al., Proteomic analyses reveal misregulation of LIN28 expression and delayed timing of glial differentiation in human iPS cells with MECP2 loss-of-function. PloS one, 2019. **14**(2).
17. de Souza, J.S., et al., *IGF1 neuronal response in the absence of MECP2 is dependent on TRalpha 3.* Human molecular genetics, 2017. **26**(2): p. 270-281.
18. Rodrigues, D.C., et al., MECP2 Is Post-transcriptionally Regulated during Human Neurodevelopment by Combinatorial Action of RNA-Binding Proteins and miRNAs. Cell Rep, 2016. **17**(3): p. 720-734.
19. Belichenko, P.V., et al., Widespread changes in dendritic and axonal morphology in Mecp2-mutant mouse models of Rett syndrome: evidence for disruption of neuronal networks. J Comp Neurol, 2009. **514**(3): p. 240-58.

20. Cheung, A.Y.L., et al., X-chromosome inactivation in rett syndrome human induced pluripotent stem cells. Frontiers in psychiatry, 2012. **3**: p. 24-24.

21. Trujillo, C., et al., Nested oscillatory dynamics in cortical organoids model early human brain network development. BioRxiv, 2018.

22. Chen, R.Z., et al., Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. Nat Genet, 2001. **27**(3): p. 327-31.

23. Matagne, V., et al., Severe offtarget effects following intravenous delivery of AAV9-MECP2 in a female mouse model of Rett syndrome. Neurobiol Dis, 2021. **149**: p. 105235.

24. Rodrigues, D.C., M. Mufteev, and J. Ellis, *Regulation, diversity and function of MECP2 exon and 3'UTR isoforms.* Hum Mol Genet, 2020. **29**(R1): p. R89-r99.

25. Pacheco, N.L., 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**: p. 56.

26. Kim, H.J., et al., Brain-Derived Neurotrophic Factor Secreting Human Mesenchymal Stem Cells Improve Outcomes in Rett Syndrome Mouse Models. Frontiers in Neuroscience, 2021. **15**.

27. Sun, Y.E. and H. Wu, *The Ups and Downs of BDNF in Rett Syndrome.* Neuron, 2006. **49**(3): p. 321-323.

28. De Felice, C., et al., *Rett syndrome: An autoimmune disease?* Autoimmun Rev, 2016. **15**(4): p. 411-6.

29. Ehrhart, F., et al., Integrated analysis of human transcriptome data for Rett syndrome finds a network of involved genes. The World Journal of Biological Psychiatry, 2019.

30. Krishnaraj, R., et al., Genome-wide transcriptomic and proteomic studies of Rett syndrome mouse models identify common signaling pathways and cellular functions as potential therapeutic targets. Hum Mutat, 2019. **40**(12): p. 2184-2196.

31. Gold, W.A., et al., Mitochondrial dysfunction in the skeletal muscle of a mouse model of Rett syndrome (RTT): implications for the disease phenotype. Mitochondrion, 2014. **15**: p. 10-7.

32. Kyle, S.M., N. Vashi, and M.J. Justice, *Rett syndrome: a neurological disorder with metabolic components.* Open Biol, 2018. **8**(2).

33. Marano, D., et al., ATRX Contributes to MeCP2-Mediated Pericentric Heterochromatin Organization during Neural Differentiation. International Journal of Molecular Sciences, 2019. **20**(21): p. 5371.

34. Gibbons, R.J., et al., Mutations in a putative global transcriptional regulator cause X-linked mental retardation with alpha-thalassemia (ATR-X syndrome). Cell, 1995. **80**(6): p. 837-45.

35. Nan, X., et al., Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation. Proc Natl Acad Sci U S A, 2007. **104**(8): p. 2709-14.

36. Valenzuela, M., et al., *The Multiple Facets of ATRX Protein.* Cancers, 2021. **13**(9): p. 2211.

37. Ramamoorthy, M. and S. Smith, Loss of ATRX Suppresses Resolution of Telomere Cohesion to Control Recombination in ALT Cancer Cells. Cancer Cell, 2015. **28**(3): p. 357-69.

38. Kim, N., et al., Whole Exome Sequencing Identifies Novel De Novo Variants Interacting with Six Gene Networks in Autism Spectrum Disorder. Genes (Basel), 2020. **12**(1).

39. Neul, J.L., *The relationship of Rett syndrome and MECP2 disorders to autism.* Dialogues in clinical neuroscience, 2012. **14**(3): p. 253-262.

40. Percy, A.K., *Rett syndrome: exploring the autism link.* Archives of neurology, 2011. **68**(8): p. 985-989.

41. Devasani, K. and Y. Yao, *Expression and functions of adenylyl cyclases in the CNS.* Fluids and Barriers of the CNS, 2022. **19**(1): p. 23.

42. Bedogni, F., et al., Defects During Mecp2 Null Embryonic Cortex Development Precede the Onset of Overt Neurological Symptoms. Cerebral Cortex, 2015. **26**(6): p. 2517-2529.

43. Carrì, M.T., et al., *Oxidative stress and mitochondrial damage: importance in non-SOD1 ALS.* Frontiers in cellular neuroscience, 2015. **9**: p. 41-41.

44. Fischer, L.R., et al., SOD1 targeted to the mitochondrial intermembrane space prevents motor neuropathy in the Sod1 knockout mouse. Brain, 2010. **134**(1): p. 196-209.

45. Damaj, L., et al., CACNA1A haploinsufficiency causes cognitive impairment, autism and epileptic encephalopathy with mild cerebellar symptoms. European journal of human genetics : EJHG, 2015. **23**(11): p. 1505-1512.

46. Martínez-Monseny, A.F., et al., CACNA1A Mutations Causing Early Onset Ataxia: Profiling Clinical, Dysmorphic and Structural-Functional Findings. International journal of molecular sciences, 2021. **22**(10): p. 5180.

47. Epperson, M.V., et al., An Atypical Rett Syndrome Phenotype Due to a Novel Missense Mutation in CACNA1A. J Child Neurol, 2018. **33**(4): p. 286-289.

48. Minichiello, L., et al., Mechanism of TrkB-mediated hippocampal long-term potentiation. Neuron, 2002. **36**(1): p. 121-37.

49. Yamada, M., et al., Analysis of tyrosine phosphorylation-dependent protein-protein interactions in TrkB-mediated intracellular signaling using modified yeast two-hybrid system. J Biochem, 2001. **130**(1): p. 157-65.

50. Corrêa, T., et al., Candidate Genes Associated With Neurological Findings in a Patient With Trisomy 4p16.3 and Monosomy 5p15.2. Frontiers in Genetics, 2020. **11**.

51. Antona, V., et al., *A Novel CCT5 Missense Variant Associated with Early Onset Motor Neuropathy.* International journal of molecular sciences, 2020. **21**(20): p. 7631.

52. Bu, Q., et al., *CREB Signaling Is Involved in Rett Syndrome Pathogenesis.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2017. **37**(13): p. 3671-3685.

53. Swinehart, B.D., et al., Integrin β3 organizes dendritic complexity of cerebral cortical pyramidal neurons along a tangential gradient. Molecular Brain, 2020. **13**(1): p. 168.

54. Wu, X. and D.S. Reddy, *Integrins as receptor targets for neurological disorders.* Pharmacology & therapeutics, 2012. **134**(1): p. 68-81.

55. Li, W., Excitation and Inhibition Imbalance in Rett Syndrome. Front Neurosci, 2022. **16**: p. 825063.

56. Lynch, G. and C.M. Gall, *Ampakines and the threefold path to cognitive enhancement.* Trends in neurosciences, 2006. **29**(10): p. 554-562.

57. Jackson, M.E., H. Homayoun, and B. Moghaddam, *NMDA receptor hypofunction produces concomitant firing rate potentiation and burst activity reduction in the prefrontal cortex.* Proceedings of the National Academy of Sciences, 2004. **101**(22): p. 8467-8472.

58. Kron, M., et al., Brain activity mapping in Mecp2 mutant mice reveals functional deficits in forebrain circuits, including key nodes in the default mode network, that are reversed with ketamine treatment. Journal of Neuroscience, 2012. **32**(40): p. 13860-13872.

59. Voituron, N. and G. Hilaire, *The benzodiazepine Midazolam mitigates the breathing defects of Mecp2-deficient mice.* Respiratory Physiology & Neurobiology, 2011. **177**(1): p. 56-60.

60. Dittmer, S., et al., TOX3 is a neuronal survival factor that induces transcription depending on the presence of CITED1 or phosphorylated CREB in the transcriptionally active complex. J Cell Sci, 2011. **124**(Pt 2): p. 252-60.

61. Schipper, H.M., et al., Heme oxygenase-1 and neurodegeneration: expanding frontiers of engagement. J Neurochem, 2009. **110**(2): p. 469-85.

62. Pozzo-Miller, L., S. Pati, and A.K. Percy, *Rett Syndrome: Reaching for Clinical Trials.* Neurotherapeutics, 2015. **12**(3): p. 631-640.

63. Djukic, A., et al., Pharmacologic Treatment of Rett Syndrome With Glatiramer Acetate. Pediatr Neurol, 2016. **61**: p. 51-7.

64. Naegelin, Y., et al., *Fingolimod in children with Rett syndrome: the FINGORETT study.* Orphanet Journal of Rare Diseases, 2021. **16**(1): p. 19.

65. Nissenkorn, A., M. Kidon, and B. Ben-Zeev, *A Potential Life-Threatening Reaction to Glatiramer Acetate in Rett Syndrome.* Pediatr Neurol, 2017. **68**: p. 40-43.

66. Ritchie, M.E., et al., limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic acids research, 2015. **43**(7): p. e47-e47.

67. Law, C.W., et al., voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biology, 2014. **15**(2): p. R29.

68. Langfelder, P. and S. Horvath, *WGCNA: an R package for weighted correlation network analysis.* BMC Bioinformatics, 2008. **9**(1): p. 559.

69. Langfelder, P. and S. Horvath, *WGCNA: an R package for weighted correlation network analysis.* BMC bioinformatics, 2008. **9**(1): p. 1-13.

70. Kanehisa, M. and S. Goto, *KEGG: kyoto encyclopedia of genes and genomes.* Nucleic Acids Res, 2000. **28**(1): p. 27-30.

71. Kanehisa, M., et al., *New approach for understanding genome variations in KEGG.* Nucleic acids research, 2019. **47**(D1): p. D590-D595.

72. Szklarczyk, D., et al., The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res, 2021. **49**(D1): p. D605-d612.

73. Zhang, Y., et al., Rapid single-step induction of functional neurons from human pluripotent stem cells. Neuron, 2013. **78**(5): p. 785-98.

74. Fumagalli, F., et al., Lentiviral haematopoietic stem-cell gene therapy for early-onset metachromatic leukodystrophy: long-term results from a non-randomised, open-label, phase 1/2 trial and expanded access. The Lancet, 2022. **399**(10322): p. 372-383.

75. Ehrhart, F., et al., New insights in Rett syndrome using pathway analysis for transcriptomics data. Wien Med Wochenschr, 2016. **166**(11-12): p. 346-52.