

ORIGINAL ARTICLE

Investigation of modifier genes within copy number variations in Rett syndrome

Rosangela Artuso^{1,5}, Filomena T Papa^{1,5}, Elisa Grillo¹, Mafalda Mucciolo¹, Dag H Yasui², Keith W Dunaway², Vittoria Disciglio¹, Maria A Mencarelli¹, Marzia Pollazzon¹, Michele Zappella³, Giuseppe Hayek⁴, Francesca Mari¹, Alessandra Renieri¹, Janine M LaSalle² and Francesca Ariani¹

MECP2 mutations are responsible for two different phenotypes in females, classical Rett syndrome and the milder Zappella variant (Z-RTT). We investigated whether copy number variants (CNVs) may modulate the phenotype by comparison of array-CGH data from two discordant pairs of sisters and four additional discordant pairs of unrelated girls matched by mutation type. We also searched for potential MeCP2 targets within CNVs by chromatin immunoprecipitation microarray (ChIP–chip) analysis. We did not identify one major common gene/region, suggesting that modifiers may be complex and variable between cases. However, we detected CNVs correlating with disease severity that contain candidate modifiers. *CROCC* (1p36.13) is a potential MeCP2 target, in which a duplication in a Z-RTT and a deletion in a classic patient were observed. *CROCC* encodes a structural component of ciliary motility that is required for correct brain development. *CFHR1* and *CFHR3*, on 1q31.3, may be involved in the regulation of complement during synapse elimination, and were found to be deleted in a Z-RTT but duplicated in two classic patients. The duplication of 10q11.22, present in two Z-RTT patients, includes *GPRIN2*, a regulator of neurite outgrowth and *PPYR1*, involved in energy homeostasis. Functional analyses are necessary to confirm candidates and to define targets for future therapies.

Journal of Human Genetics (2011) 56, 508–515; doi:10.1038/jhg.2011.50; published online 19 May 2011

Keywords: copy number variants; modifier genes; Rett syndrome

INTRODUCTION

Rett syndrome (RTT, OMIM#312750) is an X-linked neurodevelopmental disorder predominantly affecting females. In the classic form, after a period of normal development (6–18 months), patients show growth retardation and regression of speech and purposeful hand movements, with appearance of stereotyped hand movements, microcephaly, autism and seizures.^{1,2} RTT syndrome has a wide spectrum of clinical phenotypes including: the Zappella variant (Z-RTT), the early onset seizure variant and the congenital variant.³ Z-RTT, first described by M Zappella in 1992, represents the most common RTT variant. Z-RTT is characterized by a recovery of the ability to speak in single words or third person phrases and by an improvement of purposeful hand movements.^{4,5} Z-RTT patients also show milder intellectual disabilities (up to IQ of 50) and often normal head circumference, weight and height respect to classic RTT.⁵

De novo mutations in the *MECP2* gene (Xq28) account for the majority of girls with classic RTT (95–97%) and for about half of cases with Z-RTT.⁵ The other two variants have been associated with different loci, with mutations in *CDKL5* (Xp22) found in the early

onset seizure variant and mutations in *FOXG1* (14q13) found in the congenital variant.^{6–8}

Only a few *MECP2*-mutated familial cases have been reported so far. Some cases have been explained by skewing of X-inactivation towards the wild-type allele in an asymptomatic carrier.^{9–11} In others cases, germline mosaicism has been a possible explanation.^{12–14}

X-chromosome inactivation (XCI) is one important candidate factor modulating RTT phenotype. However, studies performed on blood yielded conflicting results. In 2007, Archer *et al.*¹⁵ performed the first systematic study of XCI in a large cohort of patients and found a correlation between the degree and direction of XCI in leukocytes and RTT severity. However, it has been shown that XCI may vary remarkably between tissues.^{16,17} Thus, the extrapolations of results based on sampling peripheral tissues, such as lymphocytes, to other tissues, such as brain, may be misleading. The few studies carried out on human RTT brain tissues suggest that balanced XCI patterns are prevalent.^{16,18–21} However, XCI has been investigated in a limited number of brain regions and no definitive conclusions can be drawn. In addition, previous studies demonstrated that other factors

¹Biotechnology Department, Medical Genetics Section, University of Siena, Siena, Italy; ²Medical Microbiology and Immunology, Genome Center, School of Medicine, University of California, Davis, CA, USA; ³Child Neuropsychiatry, Versilia Hospital, Viareggio, Italy and ⁴Infantile Neuropsychiatry, Siena General Hospital, Italy

⁵These authors contributed equally to the work.

Correspondence: Professor A Renieri, Biotechnology Department, Medical Genetics Section, University of Siena, Policlinico 'S. Maria alle Scotte', viale Bracci 2, 53100 Siena, Italy. E-mail: renieri@unisi.it

Received 5 January 2011; revised 7 April 2011; accepted 11 April 2011; published online 19 May 2011

such as *MECP2* mutation type and environment can influence RTT phenotype.^{5,22,23} As available data cannot fully explain RTT variability, it is likely that a combination of different factors cooperate in a complex manner to modulate the phenotype. In favor of this hypothesis, there are cases of RTT sisters with identical *MECP2* mutation, balanced X-inactivation, similar environments and discordant phenotype (one classic and one Z-RTT sister).^{9,12}

Copy number variations (CNVs) are segments of DNA ranging from kilobases (Kb) to multiple megabases (Mb) in length that contain a variable number of copies compared with the reference genome sequence. It has been demonstrated that CNVs are associated with detectable differences in transcript levels for genes within the CNV breakpoints that are predicted to have causative, functional effects in some cases. CNVs have been reported to be associated with human diseases such as neurological and autoimmune disorders and cancer.^{24–33} CNVs, to a greater extent than single nucleotide polymorphisms, represent an important source of variability in both phenotypically normal subjects and individuals with diseases.^{34,35} It is therefore reasonable to hypothesize that CNVs can modulate the phenotypic expression of RTT syndrome.

To test this hypothesis, we analyzed by array comparative genomic hybridization (array-CGH) two pairs of RTT sisters and four additional pairs of unrelated RTT girls matched by mutation type showing discordant phenotype (classic and Z-RTT). Complementary analysis of chromatin immunoprecipitation microarray (ChIP–chip) data was also carried out to identify hypothetical *MeCP2* targets included in the identified CNVs.

PATIENTS AND METHODS

Patients

From the Italian RTT database and biobank (<http://www.biobank.unisi.it>), we recruited two rare familial cases with two RTT sisters with discordant phenotype: one classic (#897 and #138) and one Z-RTT (#896 and #139).³⁶ Blood DNA from these cases were screened by both denaturing high-performance liquid chromatography and multiplex ligation-dependent probe amplification techniques to identify *MECP2* mutations. The first pair carries a large *MECP2* deletion in exon 3 and exon 4, whereas the second pair has a late truncating *MECP2* mutation: c.1157del32. Clinical descriptions of these patients have been reported in previous manuscripts.^{9,12} Furthermore, we selected four additional pairs (#565/601, #185/119, #421/109 and #402/368) of unrelated RTT patients with discordant severity of RTT phenotype (classic and Z-RTT) and the same *MECP2* mutation (c.1163del26, p.R306C, c.1159del44 and p.R133C) (Tables 1 and 2). XCI tested using the assay as modified from Pegoraro *et al.*,³⁷ revealed that all patients show balanced XCI except for case #421 displaying a skewed XCI. All cases included in the bank have been clinically evaluated by the Medical Genetics Unit of Siena. Patients were classified in classic and RTT variant according to the international criteria.^{2,38}

Genomic DNA isolation

Blood samples were obtained after informed consent. Genomic DNA of the patients was isolated from an EDTA-preserved peripheral blood sample using the QIAamp DNA Blood Kit according to the manufacturer's protocol (Qiagen SPA, Milano, Italy). Genomic DNA from normal male and female controls was obtained from Promega (Promega Italia SRL, Milano, Italy). A measure of 10 µg of genomic DNA from the patient (test sample) and the control (reference sample) were sonicated. Test and reference DNA samples were subsequently purified using affinity column purification (DNA Clean and Concentrator, Zymo Research, Irvine, CA, USA) and the appropriate DNA concentrations were determined by a DyNA Quant 200 Fluorometer (GE Healthcare, Piscataway, NJ, USA).

Array comparative genomic hybridization

Array CGH analysis was carried out using commercially available oligonucleotide microarrays containing ~99 000 60-mer probes with an estimated average resolution of 65 Kb. Probe locations are assigned according to position on the human reference genome as shown in UCSC genome browser—NCBI build 36/hg18, March 2006 (<http://genome.ucsc.edu>).

DNA labeling was performed according to the Agilent Genomic DNA Labeling Kit Plus using the Oligonucleotide Array-Based CGH for Genomic DNA Analysis 2.0v protocol (Agilent Technologies Italia SpA, Milano, Italy). Genomic DNA (3.5 µg) from patients with classical RTT and Z-RTT was mixed with Cy5-dNTP, whereas 3.5 µg of genomic DNA from a control sample with known CNVs was mixed with Cy3-dNTP, as previously reported.³⁹ The array was disassembled and washed according to the manufacturer protocol with wash buffers supplied with the Agilent 105A kit. The slides were dried and scanned using an Agilent G2565BA DNA microarray scanner (Agilent Technologies).

Array-CGH image and data analysis

Image analysis was carried out using the CGH Analytics software v 5.0.14 using the default settings (Agilent Technologies). The software automatically first determines the fluorescence intensities of the spots for both fluorochromes performing background subtraction and data normalization, then compiles the data into a spreadsheet that links the fluorescent signal of every oligo on the array to the oligo name, its position on the array and its position in the genome. The linear order of the oligos is reconstituted in the ratio plots consistent with an ideogram. The ratio plot is arbitrarily assigned such that gains and losses in DNA copy number at a particular locus are observed as a deviation of the ratio plot from a modal value of 1.0.

Analysis of *MeCP2* bound promoters within defined CNVs

ChIP–chip analysis of genome-wide promoters was carried out in a previous study.⁴⁰ Briefly, *MeCP2* ChIP was performed on two replicate human SH-SY5Y neuroblastoma cultures differentiated by 48 h treatment with phorbol 12-myristate 13-acetate (PMA) and hybridized to a commercial genome-wide promoter microarray (Nimblegen, Roche, Madison, WI, USA). In this 1.5 kb promoter array, tiled oligonucleotide probes extend 1.3 kb upstream and 0.2 kb downstream of the transcriptional start sites of 24 275 human transcripts. Statistical analysis of promoter ChIP–chip data indicated that 2600–4300 promoters were bound by *MeCP2*, with 1524 promoters common to two replicate hybridizations. Promoters were ranked according to *MeCP2* binding 'hits' based on ChIP–chip log₂ values for the two arrays (*MeCP2_B* and *MeCP2_C*). In this way, 1 represents the strongest *MeCP2* bound promoter out of 24 275 annotated genes. The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, <http://www.ncbi.nlm.nih.gov/geo> (accession no. GSE9568).

Analyses of phenotypically discordant RTT pairs resulted in 29 CNVs that included 67 candidate genes, which could potentially modify RTT phenotype. The *MeCP2* promoter rankings were compared for the list of 67 candidate genes using all gene aliases. *MeCP2* promoter levels could not be identified for 24 of the 67 CNV genes because these genes were not annotated on the NimbleGen promoter array.

RESULTS

Overall, we identified 29 CNVs, 28 of them corresponding to known polymorphic regions and one on 3q13.12 corresponding to an apparently private rearrangement duplicated in only one Z-RTT patient (#119) (Tables 1 and 2). Among the 29 CNVs, we considered 14 of them as 'unlikely modifiers' as they were apparently not associated with phenotypic severity (Table 2). These include regions containing olfactory receptors and class-II HLA molecules that are not expected to directly correlate with the phenotypic variability related to classic/Z-RTT phenotype. The remaining 15 CNVs were considered as 'likely modifiers' (Table 1). In three cases, the copy number change was consistent with severity differences in at least two pairs of RTT patients (Table 1) (Figure 1). Genes included in these potential modifier regions are listed and described in Table 3.

Table 1 CNVs classified as 'likely modifiers' as they correlate with phenotypic RTT severity

Polymorphic CNVs	Breakpoints (bp)	Gene content	MeCP2_B promoter hits rank	MeCP2_C promoter hits rank	897 C/896 Z (Ex 3/4 del)	138 C/139 Z (c.1157del32)	565 C/601 Z (c.1163del26)	185 C/119 Z (p.R306C)	421 C/109 Z (c.1159del44)	402 C/368 Z (p.R133C)
1p36.13(426 kb)	16 698 906–17 124 554	ESPNP MSTP9	13 822 —	22 690 —	Dup Z					Del C
1q31.3(55 kb)	195 011 34–195 065 867	CROCC CFHR1 CFHR3	658 23 144 20 253	16 300 7651 6994				Dup C		Amp C/Del Z
1q42.12(139 kb)	223 731 55–223 870 819	EMAH	18 604	13 553	Dup Z					
2p25.2(400 kb)	3 060 975–3 460 506	TSSC1 TTC15	941 20 740	3 174 21 641	Del Z					
2q37.3(141 kb)	242 514 59–242 655 973	/	—	—						Del Z
3q13.12(281 kb)	110 116 09–110 397 433	GUCA1C MORC1	19 293 18 317	6 167 20 394				Dup Z		
5p15.33(85 kb)	763 944–848 744	C3orf66 ZDHC11	12 136 4349	8 147 13 284	Dup Z					
6q27(210 kb)	168 114 26–168 324 002	MLL74 C6orf54 KIF25	— 2389 8 778	— 6 671 3 159						Dup Z
7p21.3(89 kb)	11 720 901–11 809 763	FRMD1 THSD7A	15 800 8 520	10 530 8 160						Del Z
8q21.3(87 kb)	87 136 222–87 222 795	PSKH2 ATP6VOD2	19 413 22 858	9 491 4 087						Dup C
10q11.22(172 kb)	46 396 163–46 568 496	GPRIN2 PPYR1	17 343 16 722	23 312 9 812		Dup Z				Dup Z
14q32.33(125 kb)	105 708 20–105 833 372	SLK COL17A1	14 769 4 292	10 236 2 579						Del Z
15q14(49 kb)	32 523 241–32 572 315	/	—	—					Del Z	
16p11(200 kb)	34 399 543–34 539 890	/	—	—						Dup Z
22q13.2(49 kb)	41 237 731–41 287 060	SERHL SERHL2	— —	— —					Dup Z	

Abbreviations: Amp, amplification; CNVs, copy number variants; C, classic; Del, deletion; Dup, duplication; Z, Z-RTT. Bold numbers are in the top 10% of MeCP2 promoter hits.

Table 2 CNVs classified as 'unlikely modifiers' as they were apparently not associated with phenotypic severity

Polymorphic CNVs	Breakpoints (bp)	Gene content	MeCP2_B promoter hits rank	MeCP2_C promoter hits rank	897 C/896 Z (Ex 3/4 del)	138 C/139 Z (c.1157del32)	565 C/601 Z (c.1163del26)	185 C/119 Z (p.R306C)	421 C/109 Z (c.1159del44)	402 C/368 Z (p.R133C)
1q44(58kb)	246 794 522–246 852 126	OR2T34	22 567	14 244	Dup Z	Del Z	Del C	Del C	Del Z	Del Z
2p11(494 kb)	89 401 838–89 895 566	OR2T10	22 348	8566		Del Z				
3q26(104 kb)	163 997 228–164 101 776	IGKV1-16D ^a	—	—		Del C				Del C/Del Z
3q29(36 kb)	196 905 767–196 942 158	/	—	—		Del C				Dup C
4q13.2(108 kb)	69 057 735–69 165 814	MUC20	17 727	12 873		Dup C				Del Z
6p21.32(65 kb)	29 939 288–30 004 636	HCG4P6	21 719	18 336		Dup C				Del C
6p21.33(77 kb)	32 595 402–32 672 983	HLA-DRB5			Del C					Del Z
		HLA-DRB1	11 824	21 879		Dup Z				Amp C
8p11.23(143 kb)	39 356 595–39 499 752	ADAM5P	14 070	29 16	Dup C	Dup C	Amp Z ^a	Amp C/Amp Z		Amp C/Amp Z
10q11.22(144 kb)	47 017 598–47 161 232	/	—	—		Dup C	Dup Z			Dup Z
14q11(860 kb)	18 624 383–19 484 013	OR11H13P				Del C				Dup C
		OR4Q3	23 539	330						
		OR4M1	24 054	4686						
		OR4N2	23 383	7030						
		OR4K2	21 957	3567						
		OR4K5	21 814	7944						
		OR4K1	23 684	162						
15q11.2(727 kb)	18 810 004–19 537 035	/	—	—	Del C					Del C
16p11.2(220 kb)	28 732 295–28 952 218	ATXN2L	—	—	Dup C					Dup Z
		TUFM	7097	3848						
		SH2B1	12 675	12 680						
		ATP2A1	15 182	23 566						
		RABEP2	10 298	16 794						
		CD19	18 469	15 604						
		NFATC2IP	23 326	1627						
		SPNS1	1393	20 317						
		LAT	15 145	14 271						
17q21.31(163 kb)	41 544 224–41 706 870	KIAA1267	—	—			Amp C/Dup Z	Dup Z		Dup Z
22q11.23(30 kb)	22 681 995–22 712 211	GSTT1	9984	14 237			Dup Z	Dup C		Dup Z

Abbreviations: Amp, amplification; CNVs, copy number variants; C, classic; Del, deletion; Dup, duplication; Z, Z-RTT. Bold numbers are in the top 10% of MeCP2 promoter hits.
^a16 isoforms.

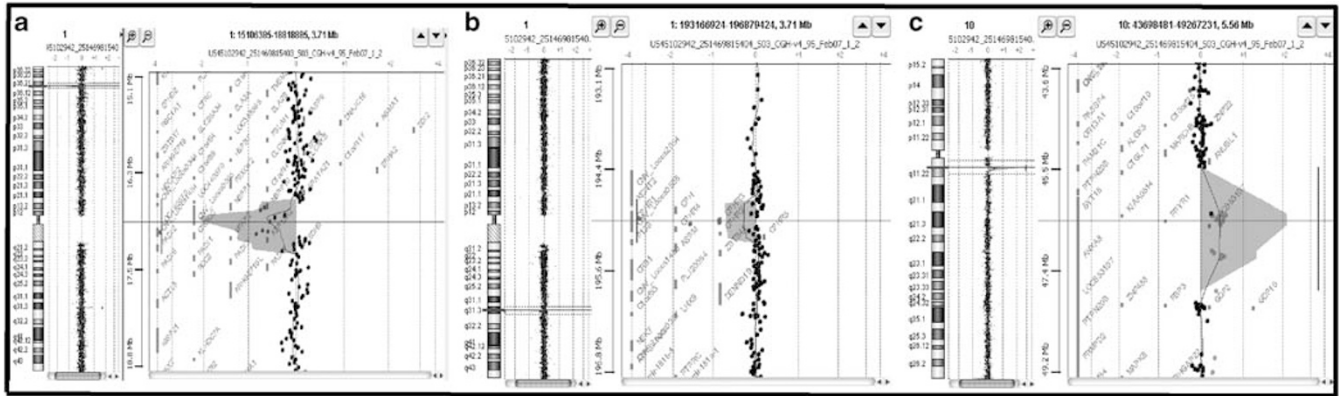


Figure 1 Array-CGH ratio profiles. (a) Array-CGH ratio profiles of CNV on 1p36.13 of #402 classic RTT patient. On the left, the chromosome 1 ideogram. On the right, the log₂ ratio of the chromosome 1 probes plotted as a function of chromosomal position. Copy-number loss shifts the ratio to the left. (b) Array-CGH ratio profiles of CNV on 1q31.3 of #368 Z-RTT patient. On the left, the chromosome 1 ideogram. On the right, the log₂ ratio of the chromosome 1 probes plotted as a function of chromosomal position. Copy-number loss shifts the ratio to the left. (c) Array-CGH ratio profiles of CNV on 10q11.22 of #139 Z-RTT patient. On the left, the chromosome 10 ideogram. On the right, the log₂ ratio of the chromosome 10 probes plotted as a function of chromosomal position. Copy-number gain shifts the ratio to the right. A full color version of this figure is available at the *Journal of Human Genetics* journal online.

To determine whether the CNVs found in phenotypically discordant RTT pairs contained possible MeCP2 target genes, we compared promoter rankings of MeCP2 binding using promoter-wide ChIP-chip analysis.⁴⁰ The ranking from total number of genes from 1 to 24 134 is shown for two replicate MeCP2–ChIP microarrays (MeCP2 B and MeCP2 C promoter hits rank, Tables 1 and 2). Genes with promoters in the top 10% of MeCP2 promoter hits for at least one replicate are indicated in bold. Among CNVs classified as ‘likely modifiers’, ChIP–chip analysis identified potential MeCP2 target genes within the 1p36.13 (*CROCC* gene, the duplication of which was found in the Z-RTT #896 and deletion in the classic form #402) and the 2p25.2 (*TSSCI* gene, the deletion of which was found in the Z-RTT #896) regions. Among CNVs classified as ‘unlikely modifiers’, ChIP–chip analysis identified potential MeCP2 target genes on 14q11 (*OR4Q3* and *OR4Q1*, deleted in a classic patient #138 and duplicated in another classic patient #421) and on 16p11.2 (*NEAT-C2IP* and *SPNS1*, duplicated in both a classic #897 and a Z-RTT patient #368).

DISCUSSION

To test the hypothesis that genes contained within common CNVs may modulate the RTT phenotype, we analyzed by array-CGH two pairs of RTT sisters and four additional pairs of unrelated RTT girls matched by *MECP2* mutation type showing discordant phenotype: classic and Z-RTT. Our study did not identify a single major common modifier gene/region, suggesting that genetic modifiers may be complex and variable between cases (Tables 1 and 2). In total we found 29 CNVs that were divided into two groups: ‘likely modifiers’ and ‘unlikely modifiers’ (Tables 1 and 2).

Among the first group, the rearrangement on 1p36.13 includes *CROCC* (ciliary rootlet coiled-coil) that represents an interesting potential modifier gene. This gene is duplicated in the Z-RTT patient #896 and deleted in the classical patient #402, suggesting that change in its expression may modulate RTT outcome. Moreover, according to ChIP–chip analysis, *CROCC* could be a potential MeCP2 target gene (Table 1). *CROCC* encodes for a major structural component (Rootletin) of the ciliary rootlet, a cytoskeletal-like structure in ciliated cells, which originates from the basal body at the proximal end of a cilium and extends proximally toward the cell nucleus.⁴¹

In non-ciliated cells, a miniature ciliary rootlet is located at the centrosome and does not project a fibrous network into the cytoplasm.⁴¹ Rootletin is expressed in retina, brain, trachea and kidney.⁴¹ Cilia generate specialized structures that perform critical functions of several broad types: sensation, development, fluid movement, sperm motility and cell signaling. Their functional significance in tissues is reflected in the severity and diversity of pathologies caused by defects in cilia. These include anosmia, retinitis pigmentosa and retinal degeneration, polycystic kidney disease, diabetes, neural tube defects and neural patterning defects, chronic sinusitis and bronchiectasis, obesity, heterotaxias, polydactyly and infertility.⁴² Defects in cilia are therefore underlying causes of several diseases with pleiotropic symptoms.⁴³ Several pleiotropic disorders (Bardet-Biedl syndrome, Alstrom syndrome, Meckel-Gruber syndrome and Joubert syndrome) caused by disruption of the function of cilia present, mental retardation or other cognitive defects as part of their phenotypic spectrum.⁴⁴ The presence of cilia in different types of neurons supports the notion that dysfunction in specific neuronal populations might explain, at least in part, such defects.^{42,45} If MeCP2 exerts an effect as a positive regulator of *CROCC*, it can be hypothesized that higher protein levels because of the presence of three copies of the gene may counteract the *MECP2* mutation, whereas lower protein level because of single gene copy may worsen the phenotype.

The *CFHR* gene family members (*CFHR1* and *CFHR3*) located on 1q31.3 are duplicated in classic girls (#185 and #402) and deleted in Z-RTT (#368), suggesting that the phenotype may benefit from the reduced expression of these proteins involved in complement regulation.⁴⁶ The complement system is a tightly controlled component of the host innate immune defence. Imbalances in regulation of this system contribute to tissue injury and can result in autoimmune diseases. In particular, *CFHR1* and *CFHR3* was previously associated with hemolytic uremic syndrome and age-related macular degeneration.^{47–49} It is well known that the immune system participates in the development and functioning of the CNS, and an immune etiology for RTT and autism has been recently hypothesized.⁵⁰ Interestingly, complement proteins have been demonstrated to be fundamental for CNS synapse elimination.⁵¹ Morphological studies in postmortem brain samples from RTT individuals described a characteristic neuropathology, which included decreased dendritic arborization, a

Table 3 Genes included in potential modifier regions

Position	Gene	Description	Function
Chr1:16 890 300–16 919 239	<i>ESPNP</i>	Espin pseudo gene, non-coding RNA	Unknown
Chr1:16 954 395–16 959 139	<i>MSTP9</i>	Macrophage stimulating, pseudogene 9, non-coding RNA	Unknown
Chr1:17 121 032–17 172 061	<i>CROCC</i>	Ciliary rootlet coiled-coil, rootletin	Ciliary rootlet formation
Chr1:195 055 484–195 067 942	<i>CFHR1</i>	Complement factor H-related 1	Complement regulation
Chr1:195 010 553–195 029 496	<i>CFHR3</i>	Complement factor H-related 3	Complement regulation
Chr1:223 741 157–223 907 468	<i>EMAH</i>	Enabled homolog (Drosophila)	Involvement in a range of processes dependent on cytoskeleton remodeling and cell polarity, such as axon guidance and lamellipodial and filopodial dynamics in migrating cells
Chr2:3 171 750–3 360 605	<i>TSSC1</i>	Tumor suppressing subtransferable candidate 1	Possible involvement in tumor suppression
Chr2:3 362 453–3 462 349	<i>TTC15</i>	Tetratricopeptide repeat domain 15	Possible involvement in autophagy
Chr3:110 109 340–110 155 310	<i>GUCA1C</i>	Guanylate cyclase activator 1C	Ca(2+)-sensitive regulation of guanylyl cyclase
Chr3:110 159 777–110 319 658	<i>MORC1</i>	MORC family CW-type zinc-finger 1	Possible role in spermatogenesis
Chr3:110 379 702–110 386 794	<i>C3orf66</i>	Chromosome 3 open reading frame 66, non-coding RNA	Unknown
Chr5:848 722–904 101	<i>ZDHHC11</i>	Zinc finger, DHHC-type containing 11	Probable palmitoyltransferase activity
Chr6:167 970 520–168 115 552	<i>MLLT4</i>	Myeloid/lymphoid or mixed-lineage leukemia	Involvement in adhesion system, probably together with the E-cadherin–catenin system, which has a role in the organization of homotypic, interneuronal and heterotypic cell–cell adherens junctions
Chr6:168 136 351–168 140 606	<i>C6orf54</i>	Chromosome 6 open reading frame 54	Unknown
Chr6:168 161 402–168 188 618	<i>KIF25</i>	Kinesin family member 25	Possible involvement in microtubule-dependent molecular transport of organelles within cells and movement of chromosomes during cell division.
Chr6:168 199 313–168 222 688	<i>FRMD1</i>	FERM domain containing 1	Unknown
Chr7:11 380 696–11 838 349	<i>THSD7A</i>	Thrombospondin, type I, domain containing 7A	Interaction with alpha (V) beta (3) integrin and paxillin to inhibit endothelial cell migration and tube formation.
Chr8:87 129 807–87 150 967	<i>PSKH2</i>	Protein serine kinase H2	CAMK Ser/Thr protein kinase
Chr8:87 180 318–87 234 433	<i>ATP6V0D2</i>	ATPase, H+ transporting, lysosomal 38 kDa, V0 subunit d2	Subunit of the integral membrane VO complex of vacuolar ATPase, which is responsible for acidifying a variety of intracellular compartments in eukaryotic cells, thus providing most of the energy required for transport processes in the vacuolar system. It may have a role in coupling of proton transport and ATP hydrolysis.
Chr10:46 413 552–46 420 574	<i>GPRIN2</i>	G protein-regulated inducer of neurite outgrowth 2	Possible role in the control growth of neurites
Chr10:46 503 540–46 508 326	<i>PPYR1</i>	Pancreatic polypeptide receptor 1	It belongs to a family of receptors for neuropeptide Y involved in a diverse range of biological functions including, stimulation of food intake, anxiolysis, modulation of circadian rhythm, pain transmission and control of pituitary hormone release.
Chr10:105 717 460–105 777 332	<i>SLK</i>	STE20-like kinase (yeast)	Possible mediation of apoptosis and actin stress fiber dissolution
Chr10:105 781 036–105 835 628	<i>COL17A1</i>	Collagen, type XVII, alpha 1	It encodes the alpha chain of type XVII collagen, which is a structural component of hemidesmosomes, multiprotein complexes at the dermal–epidermal basement membrane zone mediating adhesion of keratinocytes to the underlying membrane.
Chr22:41 226 540–41 238 510	<i>SERHL</i>	Serine hydrolase-like	Possible role in normal peroxisome function and skeletal muscle growth in response to mechanical stimuli.
Chr22:41 279 869–41 300 332	<i>SERHL2</i>	Serine hydrolase-like 2	Probable serine hydrolase. May be related to muscle hypertrophy.

reduction in dendritic spines and increased packing density.⁵² It is therefore possible that the protein product of *CFHR* could be involved in the regulation of synaptic connections and that these genes could influence RTT severity.

The duplication on 10q11.22, present in two Z-RTT patients (#139 and #368), includes two interesting candidate modifier genes: *GPRIN2* and *PPYR1*. *GPRIN2* is highly expressed in the cerebellum and interacts with activated members of the Gi subfamily of G protein α subunits and functions together with *GPRIN1* to regulate neurite outgrowth.⁵³ *PPYR1*, also named as neuropeptide Y receptor or pancreatic polypeptide 1, is a key regulator of energy homeostasis and directly involved in the regulation of food intake. Previous studies have reinforced the potential influence of *PPYR1* on body weight in humans.⁵⁴ Moreover, it has been demonstrated that *PPYR1* knockout mice display lower body weight and reduced white adipose tissue.⁵⁵ Thus, a higher level of *PPYR1* expression because of gene duplication may correlate with the higher body weight, characterizing Z-RTT patients in respect to classic RTT.⁵ In contrast, a recent study demonstrated that 10q11.22 gain is associated with lower body mass index value in the Chinese population.⁵⁶ However this CNV is much larger with respect to the one reported here and includes two additional genes.⁵⁶

The 3q13.12 duplication found in a Z-RTT patient (#119) encompasses about 280 Kb and does not contain interesting candidate RTT modifier genes. *GUCA1C* encodes for a granulate cyclase activating protein expressed in retina and *MORC1* encodes for a testis-specific protein with a putative role in spermatogenesis. However, it is known that CNVs can also induce altered expression of genes that lie near the boundaries of the CNV and that this effect can be as far as 2–7 Mb away from the breakpoints.⁵⁷ Therefore we cannot totally exclude a role for this CNV in modulating RTT phenotype.

The 1q42.12 region, duplicated in one Z-RTT patient (#896), includes *ENAH*. This gene was identified as a mammalian homolog of *Drosophila* Ena and initially named Mena (Mammalian enabled).⁵⁸ It localizes to cell-substrate adhesion sites and sites of dynamic actins assembly and disassembly. It is a member of the Ena/VASP family that also includes VASP and EVL in vertebrates. Work carried out in *Drosophila*, *Caenorhabditis elegans* and mice showed that these proteins participate in axonal outgrowth, dendrite morphology, synapse formation and also function downstream of attractive and repulsive axon guidance pathways.^{59–61} Previous evidence shows that knocking out the three murine genes encoding ENA/VASP proteins results in a blockade of axon fibre tract formation in the cortex *in vivo*, and that failure in neuritis initiation is the underlying cause of the axonal defects.^{62,63} *ENAH* therefore represents an interesting potential gene modifier in RTT. Further investigations are necessary to test whether the duplication of *ENAH* gene in Z-RTT #896 effectively corresponds with increased mRNA levels in brain, and whether this mechanism is confined to one pair of discordant girls or is a common mechanism in Z-RTT possibly throughout single nucleotide polymorphism modulation.

The intersection of CNV and MeCP2 promoter binding analyses was useful in identifying potential modifier genes for further investigation. However, genes with MeCP2 bound promoters were not apparently enriched within the CNVs in the 'likely' versus 'unlikely' modifier categories. MeCP2 binding is found more frequently in non-promoter regions when analyzed by genomic tiling microarray to selected regions, so the analysis of promoters only in identifying potential MeCP2 target genes was a limitation of this study.⁴⁰ Further studies to detect MeCP2 binding genome wide in human neurons by Chip sequencing may reveal additional insights.

A second limitation of this study is that the number of patients is too low to perform a statistically significant analysis of CNVs in classic and Z-RTT, and this is principally due to the difficulty in recruiting Z-RTT cases. Furthermore, mRNA expression analysis of genes within CNVs has not been conducted because of a lack of sufficient blood RNA samples. However, an analysis of transcript levels in blood would not be conclusive because the genes within likely modifier CNVs exhibit tissue-specific expression in tissues other than blood cells. Our studies do suggest genes for further studies in animal models or in new cellular models such as neurons derived from human-induced pluripotent stem cells (iPS).

Moreover this study indicates possible candidate genes to test for functional single nucleotide polymorphisms in array-CGH negative cases. In fact this study is focused on CNVs but single nucleotide polymorphisms could also have an important role in determining RTT phenotypic variability. By candidate gene approach, this has been already demonstrated for the p.Val66Met polymorphism in *BDNF*, even if with contrasting results.^{64,65} The recent feasibility of exome sequencing will allow to yield important results that will further improve the understanding of RTT phenotypic variability.

In conclusion, we present a novel approach for investigating genetic modifiers for RTT severity by identifying CNVs different between pairs with discordant phenotype: classic and Z-RTT. Further investigation using gene expression and/or statistical analysis in a larger number of patients will be necessary to confirm these data and to define targets for future therapeutic intervention.

ACKNOWLEDGEMENTS

We would first like to thank Rett patients and their families. This work was supported by 'Cell Lines and DNA Bank of Rett syndrome and X mental retardation' (Medical Genetics-Siena)-Telethon Genetic Biobank Network (Project No. GTB07001C to AR) and NIH R01HD041462 to JML.

- 1 Trevathan, E. & Moser, H. W. Diagnostic criteria for Rett syndrome. *Ann. Neurol.* **23**, 425–428 (1988).
- 2 Neul, J., Kaufmann, W., Glaze, D., Christodoulou, J., Clarke, A., Bahi-Buisson, N. et al. Rett syndrome: revised diagnostic criteria and nomenclature. *Annals of Neurology* **68**, 944–950 (2010).
- 3 Mencarelli, M. A., Spanhol-Rosseto, A., Artuso, R., Rondinella, D., De Filippis, R., Bahi-Buisson, N. et al. Novel FOXP1 mutations associated with the congenital variant of Rett syndrome. *J. Med. Genet.* **47**, 49–53 (2010).
- 4 Zappella, M. The Rett girls with preserved speech. *Brain Dev.* **14**, 98–101 (1992).
- 5 Renieri, A., Mari, F., Mencarelli, M. A., Scala, E., Ariani, F., Longo, I. et al. Diagnostic criteria for the Zappella variant of Rett syndrome (the preserved speech variant). *Brain Dev.* **31**, 208–216 (2009).
- 6 Weaving, L., Christodoulou, J., Williamson, S., Friend, K., McKenzie, O., Archer, H. et al. Mutations of CDKL5 cause a severe neurodevelopmental disorder with infantile spasms and mental retardation. *Am. J. Hum. Genet.* **75**, 1079–1093 (2004).
- 7 Scala, E., Ariani, F., Mari, F., Caselli, R., Pescucci, C., Longo, I. et al. CDKL5/STK9 is mutated in Rett syndrome variant with infantile spasms. *J. Med. Genet.* **42**, 103–107 (2005).
- 8 Ariani, F., Hayek, G., Rondinella, D., Artuso, R., Mencarelli, M. A., Spanhol-Rosseto, A. et al. FOXP1 is responsible for the congenital variant of Rett syndrome. *Am. J. Hum. Genet.* **83**, 89–93 (2008).
- 9 Zappella, M., Meloni, I., Longo, I., Hayek, G. & Renieri, A. Preserved speech variants of the Rett Syndrome: molecular and clinical analysis. *Am. J. Med. Genet.* **104**, 14–22 (2001).
- 10 Hardwick, S. A., Reuter, K., Williamson, S. L., Vasudevan, V., Donald, J., Slater, K. et al. Delineation of large deletions of the *MECP2* gene in Rett syndrome patients, including a familial case with a male proband. *Eur. J. Hum. Genet.* **15**, 1218–1229 (2007).
- 11 Amir, R. E., Van den Veyver, I. B., Schultz, R., Malicki, D. M., Tran, C. Q., Dahle, E. J. et al. Influence of mutation type and X chromosome inactivation on Rett syndrome phenotypes. *Ann. Neurol.* **47**, 670–679 (2000).
- 12 Scala, E., Longo, I., Ottimo, F., Speciale, C., Sampieri, K., Katzaki, E. et al. MECP2 deletions and genotype-phenotype correlation in Rett syndrome. *Am. J. Med. Genet. A* **143**, 2775–2784 (2007).

- 13 Evans, J. C., Archer, H. L., Whatley, S. D. & Clarke, A. Germline mosaicism for a MECP2 mutation in a man with two Rett daughters. *Clin. Genet.* **70**, 336–338 (2006).
- 14 Mari, F., Azimonti, S., Bertani, I., Bolognese, F., Colombo, E., Caselli, R. *et al.* CDKL5 belongs to the same molecular pathway of MeCP2 and it is responsible for the early-onset seizure variant of Rett syndrome. *Hum. Mol. Genet.* **14**, 1935–1946 (2005).
- 15 Archer, H., Evans, J., Leonard, H., Colvin, L., Ravine, D., Christodoulou, J. *et al.* Correlation between clinical severity in patients with Rett syndrome with a p.R168X or p.T158M MECP2 mutation, and the direction and degree of skewing of X-chromosome inactivation. *J. Med. Genet.* **44**, 148–152 (2007).
- 16 Zoghbi, H. Y., Percy, A. K., Schultz, R. J. & Fill, C. Patterns of X chromosome inactivation in the Rett syndrome. *Brain Dev.* **12**, 131–135 (1990).
- 17 Sharp, A., Robinson, D. & Jacobs, P. Age- and tissue-specific variation of X chromosome inactivation ratios in normal women. *Hum. Genet.* **107**, 343–349 (2000).
- 18 Anvret, M. & Wahlstrom, J. Rett syndrome: random X chromosome inactivation. *Clin. Genet.* **45**, 274–275 (1994).
- 19 LaSalle, J. M., Goldstine, J., Balmer, D. & Greco, C. M. Quantitative localization of heterogeneous methyl-CpG-binding protein 2 (MeCP2) expression phenotypes in normal and Rett syndrome brain by laser scanning cytometry. *Hum. Mol. Genet.* **10**, 1729–1740 (2001).
- 20 Shahbazian, M. D., Sun, Y. & Zoghbi, H. Y. Balanced X chromosome inactivation patterns in the Rett syndrome brain. *Am. J. Med. Genet.* **111**, 164–168 (2002).
- 21 Gibson, J. H., Williamson, S. L., Arbuckle, S. & Christodoulou, J. X chromosome inactivation patterns in brain in Rett syndrome: implications for the disease phenotype. *Brain Dev.* **27**, 266–270 (2005).
- 22 Nag, N., Mellott, T. J. & Berger-Sweeney, J. E. Effects of postnatal dietary choline supplementation on motor regional volume and growth factor expression in a mouse model of Rett syndrome. *Brain Res.* **1237**, 101–109 (2008).
- 23 Lonetti, G., Angelucci, A., Morando, L., Boggio, E. M., Giustetto, M. & Pizzorusso, T. Early environmental enrichment moderates the behavioral and synaptic phenotype of MeCP2 null mice. *Biol. Psychiatry.* **67**, 657–665.
- 24 Aitman, T. J., Dong, R., Vyse, T. J., Norsworthy, P. J., Johnson, M. D., Smith, J. *et al.* Copy number polymorphism in Fcgr3 predisposes to glomerulonephritis in rats and humans. *Nature* **439**, 851–855 (2006).
- 25 Mamtani, M., Rovin, B., Brey, R., Camargo, J. F., Kulkarni, H., Herrera, M. *et al.* CCL3L1 gene-containing segmental duplications and polymorphisms in CCR5 affect risk of systemic lupus erythematosus. *Ann. Rheum. Dis.* **67**, 1076–1083 (2008).
- 26 Hollox, E. J. Copy number variation of beta-defensins and relevance to disease. *Cytogenet. Genome. Res.* **123**, 148–155 (2008).
- 27 Fellermann, K., Stange, D. E., Schaeffeler, E., Schmalzl, H., Wehkamp, J., Bevins, C. L. *et al.* A chromosome 8 gene-cluster polymorphism with low human beta-defensin 2 gene copy number predisposes to Crohn disease of the colon. *Am. J. Hum. Genet.* **79**, 439–448 (2006).
- 28 Singleton, A. B., Farrer, M., Johnson, J., Singleton, A., Hague, S., Kachergus, J. *et al.* alpha-Synuclein locus triplication causes Parkinson's disease. *Science* **302**, 841 (2003).
- 29 Ibanez, P., Bonnet, A. M., Debarges, B., Lohmann, E., Tison, F., Pollak, P. *et al.* Causal relation between alpha-synuclein gene duplication and familial Parkinson's disease. *Lancet* **364**, 1169–1171 (2004).
- 30 Stefansson, H., Rujescu, D., Cichon, S., Pietilainen, O. P., Ingason, A., Steinberg, S. *et al.* Large recurrent microdeletions associated with schizophrenia. *Nature* **455**, 232–236 (2008).
- 31 Diskin, S. J., Hou, C., Glessner, J. T., Attiye, E. F., Laudenslager, M., Bosse, K. *et al.* Copy number variation at 1q21.1 associated with neuroblastoma. *Nature* **459**, 987–991 (2009).
- 32 Frank, B., Hemminki, K., Meindl, A., Wappenschmidt, B., Sutter, C., Kiechle, M. *et al.* BRIP1 (BACH1) variants and familial breast cancer risk: a case-control study. *BMC Cancer* **7**, 83 (2007).
- 33 Karypidis, A. H., Olsson, M., Andersson, S. O., Rane, A. & Ekstrom, L. Deletion polymorphism of the UGT2B17 gene is associated with increased risk for prostate cancer and correlated to gene expression in the prostate. *Pharmacogenomics J.* **8**, 147–151 (2008).
- 34 Iafate, A. J., Feuk, L., Rivera, M. N., Listewnik, M. L., Donahoe, P. K., Qi, Y. *et al.* Detection of large-scale variation in the human genome. *Nat. Genet.* **36**, 949–951 (2004).
- 35 Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Andrews, T. D. *et al.* Global variation in copy number in the human genome. *Nature* **444**, 444–454 (2006).
- 36 Sampieri, K., Meloni, I., Scala, E., Ariani, F., Caselli, R., Pescucci, C. *et al.* Italian Rett database and biobank. *Hum. Mutat.* **28**, 329–335 (2007).
- 37 Pegoraro, E., Schimke, R. N., Arahata, K., Hayashi, Y., Stern, H., Marks, H. *et al.* Detection of new paternal dystrophin gene mutations in isolated cases of dystrophinopathy in females. *Am. J. Hum. Genet.* **54**, 989–1003 (1994).
- 38 Hagberg, B., Hanefeld, F., Percy, A. & Skjeldal, O. An update on clinically applicable diagnostic criteria in Rett syndrome. Comments to Rett syndrome clinical criteria consensus panel satellite to European paediatric neurology society meeting, Baden Baden, Germany, 11 September 2001. *Eur. J. Paediatr. Neurol.* **6**, 293–297 (2002).
- 39 Sampieri, K., Amenduni, M., Papa, F. T., Katzaki, E., Mencarelli, M. A., Marozza, A. *et al.* Array comparative genomic hybridization in retinoma and retinoblastoma tissues. *Cancer Sci.* **100**, 465–471 (2009).
- 40 Yasui, D. H., Peddada, S., Bieda, M. C., Vallero, R. O., Hogart, A., Nagarajan, R. P. *et al.* Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction with active genes. *Proc. Natl. Acad. Sci. USA* **104**, 19416–19421 (2007).
- 41 Yang, J., Liu, X., Yue, G., Adamian, M., Bulgakov, O. & Li, T. Rootletin, a novel coiled-coil protein, is a structural component of the ciliary rootlet. *J. Cell. Biol.* **159**, 431–440 (2002).
- 42 Lee, J. H. & Gleeson, J. G. The role of primary cilia in neuronal function. *Neurobiol. Dis.* **38**, 167–172.
- 43 McClintock, T. S., Glasser, C. E., Bose, S. C. & Bergman, D. A. Tissue expression patterns identify mouse cilia genes. *Physiol. Genomics.* **32**, 198–206 (2008).
- 44 Badano, J. L., Mitsuma, N., Beales, P. L. & Katsanis, N. The ciliopathies: an emerging class of human genetic disorders. *Annu. Rev. Genomics. Hum. Genet.* **7**, 125–148 (2006).
- 45 Han, Y. G. & Alvarez-Buylla, A. Role of primary cilia in brain development and cancer. *Curr. Opin. Neurobiol.* **20**, 58–67.
- 46 Male, D. A., Ormsby, R. J., Ranganathan, S., Giannakis, E. & Gordon, D. L. Complement factor H: sequence analysis of 221 kb of human genomic DNA containing the entire fH, fHR-1 and fHR-3 genes. *Mol. Immunol.* **37**, 41–52 (2000).
- 47 Hughes, A. E., Orr, N., Esfandiari, H., Diaz-Torres, M., Goodship, T. & Chakravarthy, U. A common CFH haplotype, with deletion of CFHR1 and CFHR3, is associated with lower risk of age-related macular degeneration. *Nat. Genet.* **38**, 1173–1177 (2006).
- 48 Raychaudhuri, S., Ripke, S., Li, M., Neale, B. M., Fagerness, J., Reynolds, R. *et al.* Associations of CFHR1-CFHR3 deletion and a CFH SNP to age-related macular degeneration are not independent. *Nat. Genet.* **42**, 553–555. author reply 555–556.
- 49 Jozsi, M., Licht, C., Strobel, S., Zipfel, S. L., Richter, H., Heinen, S. *et al.* Factor H autoantibodies in atypical hemolytic uremic syndrome correlate with CFHR1/CFHR3 deficiency. *Blood* **111**, 1512–1514 (2008).
- 50 Derecki, N. C., Privman, E. & Kipnis, J. Rett syndrome and other autism spectrum disorders—Brain diseases of immune malfunction? *Mol. Psychiatry* **15**, 355–363.
- 51 Stevens, B., Allen, N. J., Vazquez, L. E., Howell, G. R., Christopherson, K. S., Nouri, N. *et al.* The classical complement cascade mediates CNS synapse elimination. *Cell* **131**, 1164–1178 (2007).
- 52 Armstrong, D. D. Neuropathology of Rett syndrome. *J. Child Neurol.* **20**, 747–753 (2005).
- 53 Iida, N. & Kozasa, T. Identification and biochemical analysis of GRIN1 and GRIN2. *Methods Enzymol.* **390**, 475–483 (2004).
- 54 Kamiji, M. M. & Inui, A. Neuropeptide Y receptor selective ligands in the treatment of obesity. *Endocr. Rev.* **28**, 664–684 (2007).
- 55 Sainsbury, A., Schwarzer, C., Couzens, M., Fetissov, S., Furlinger, S., Jenkins, A. *et al.* Important role of hypothalamic Y2 receptors in body weight regulation revealed in conditional knockout mice. *Proc. Natl. Acad. Sci. USA* **99**, 8938–8943 (2002).
- 56 Sha, B. Y., Yang, T. L., Zhao, L. J., Chen, X. D., Guo, Y., Chen, Y. *et al.* Genome-wide association study suggested copy number variation may be associated with body mass index in the Chinese population. *J. Hum. Genet.* **54**, 199–202 (2009).
- 57 Merla, G., Howald, C., Henrichsen, C. N., Lyle, R., Wyss, C., Zabol, M. T. *et al.* Submicroscopic deletion in patients with Williams-Beuren syndrome influences expression levels of the nonhemizygous flanking genes. *Am. J. Hum. Genet.* **79**, 332–341 (2006).
- 58 Gertler, F. B., Niebuhr, K., Reinhard, M., Wehland, J. & Soriano, P. Mena, a relative of VASP and Drosophila Enabled, is implicated in the control of microfilament dynamics. *Cell* **87**, 227–239 (1996).
- 59 Lanier, L. M., Gates, M. A., Witke, W., Menzies, A. S., Wehman, A. M., Macklis, J. D. *et al.* Mena is required for neurulation and commissure formation. *Neuron* **22**, 313–325 (1999).
- 60 Lin, Y. L., Lei, Y. T., Hong, C. J. & Hsueh, Y. P. Syndecan-2 induces filopodia and dendritic spine formation via the neurofibromin-PKA-Ena/VASP pathway. *J. Cell. Biol.* **177**, 829–841 (2007).
- 61 Li, W., Li, Y. & Gao, F. B. Abelson, enabled, and p120 catenin exert distinct effects on dendritic morphogenesis in Drosophila. *Dev. Dyn.* **234**, 512–522 (2005).
- 62 Kwiatkowski, A. V., Rubinson, D. A., Dent, E. W., Edward van Veen, J., Leslie, J. D., Zhang, J. *et al.* Ena/VASP is required for neurogenesis in the developing cortex. *Neuron* **56**, 441–455 (2007).
- 63 Dent, E. W., Kwiatkowski, A. V., Mebane, L. M., Philippar, U., Barzik, M., Rubinson, D. A. *et al.* Filopodia are required for cortical neurite initiation. *Nat. Cell Biol.* **9**, 1347–1359 (2007).
- 64 Nectoux, J., Bahi-Buisson, N., Guellec, I., Coste, J., De Roux, N., Rosas, H. *et al.* The p.Val66Met polymorphism in the BDNF gene protects against early seizures in Rett syndrome. *Neurology* **70**, 2145–2151 (2008).
- 65 Zeev, B. B., Bebbington, A., Ho, G., Leonard, H., de Klerk, N., Gak, E. *et al.* The common BDNF polymorphism may be a modifier of disease severity in Rett syndrome. *Neurology* **72**, 1242–1247 (2009).