ARTICLE

Dimensional phenotypic analysis and functional categorisation of mutations reveal novel genotype–phenotype associations in Rett syndrome

Tony Charman*[,1], Tracey CS Neilson[2], Veronica Mash[1], Hayley Archer[3], Mary T Gardiner[2], Gun PS Knudsen[4], Aoibhinn McDonnell[2], Jacqueline Perry[2], Sharon D Whatley[5], David J Bunyan[6], Kirstine Ravn[7], Rebecca H Mount[1], Richard P Hastings[8], Maj Hulten[9], Karen Helene Ørstavik[4,10], Sheena Reilly[11], Hilary Cass[1], Angus Clarke[3], Alison M Kerr[12] and Mark ES Bailey[2]

[1]Institute of Child Health, University College London, London, UK; [2]Division of Molecular Genetics, Institute of Biomedical and Life Sciences, University of Glasgow, Glasgow, UK; [3]Department of Medical Genetics, University of Wales College of Medicine, Cardiff, UK; [4]Faculty Division, Rikshospitalet, University of Oslo, Oslo, Norway; [5]Department of Medical Biochemistry and Immunology, University Hospital of Wales, Cardiff, UK; [6]National Genetics Reference Laboratory (Wessex), Salisbury District Hospital, Salisbury, UK; [7]Department of Clinical Genetics, University Hospital, Rigshospitalet, Copenhagen, Denmark; [8]School of Psychology, University of Wales Bangor, Bangor, UK; [9]Department of Biological Science, University of Warwick, Warwick, UK; [10]Department of Medical Genetics, Rikshospitalet, University of Oslo, Oslo, Norway; [11]School of Human Communication Sciences, La Trobe University, Melbourne, Australia; [12]Department of Psychological Medicine, University of Glasgow, Glasgow, UK

We aimed to improve the understanding of genotype–phenotype correlations in Rett syndrome (RS) by adopting a novel approach to categorising phenotypic dimensions – separating typicality of presentation, outcome severity and age of onset – and by classifying *MECP2* mutations strictly by predicted functional attributes. *MECP2* mutation screening results were available on 190 patients with a clinical diagnosis of RS (140 cases with classic RS, 50 with atypical RS). 135 cases had identified mutations. Of the 140 patients, 116 with classic RS (82.9%) had an identified mutation compared with 19 of 50 patients (38%) with an atypical presentation. Cases with early onset of regression and seizures, and those with clinical features that might indicate alternative aetiologies, were less likely to have mutations. Individuals with late truncating mutations had a less typical presentation than cases with missense and early truncating mutations, presumably reflecting greater residual function of MECP2 protein. Individuals with early truncating mutations had a more severe outcome than cases with missense and late truncating mutations. These findings held when restricting the analysis to cases over 15 years of age and classic cases only. Previous findings of variation in severity among the common mutations were confirmed. The approach to phenotypic and genotypic classification adopted here allowed us to identify genotype–phenotype associations in RS that may aid our understanding of pathogenesis and also contribute to clinical knowledge on the impact of different types of mutations.

*Correspondence: Dr T Charman, Behavioural & Brain Sciences Unit, Institute of Child Health, 30 Guilford Street, London WC1N 1EH, UK. Tel: +44 207 905 2163; Fax: +44 207 831 7050;

E-mail: t.charman@ich.ucl.ac.uk

Introduction

Rett syndrome (RS) (genetic locus RTT1; MIM No. 312750) is an X-linked dominant neurodevelopmental disorder predominantly affecting females.[1,2] The prevalence is estimated to be approximately one in 10 000/15 000 females.[3,4] The majority (80–90%)[5,6] of classic RS cases are the result of sporadic mutations in the methyl-CpG-binding protein 2 gene (MECP2), while a smaller proportion of atypical RS cases have identified MECP2 mutations. To date, more than 200 different mutations have been identified in MECP2.[7,8]

The core diagnostic features of RS include subtle developmental deviation from birth but with some progress such that infants may be considered normal, normal head circumference at birth followed by deceleration of head growth, a period of regression characterised by loss of hand skills, severe intellectual disability and language delay, onset of stereotypic midline hand movements, and gait dyspraxia.[9,10] Associated features include breath holding and hyperventilation, EEG abnormalities and epilepsy, spasticity and scoliosis, growth retardation and hypotrophic feet.[10] Atypical presentations of RS can occur.[11,12] These include both milder (eg 'preserved speech' variant, late regression) and more severe (eg frank abnormality from birth, early-onset seizures, early regression) presentations.

One current focus of investigation is the extent to which the clinical phenotype in RS is dependent on the nature of the molecular lesion (ie mutation) that causes the disorder. Findings from initial genotype–phenotype studies were somewhat inconsistent.[7] In part, this was due to small sample sizes and the use of different measures of both the phenotype and genotype. However, a more consistent picture has emerged in larger, more recent studies.[13–17] For example, missense mutations have been found to be associated with a less severe phenotype compared to truncating mutations in several studies.[13,15–17] Recent studies have had large enough samples to examine severity of the phenotype in patients with common individual mutations. p.R133C is associated with a milder phenotype[13,18] and p.R255X and p.R270X mutations (both located within the transcription repression domain-nuclear localisation signal (TRD-NLS) region) are associated with a more severe phenotype.[13,16] However, there are still inconsistencies across studies; for example, p.R306C has been associated with a less severe phenotype in some studies[16] but not in others.[13] Several genetic factors (eg skewing of X chromosome inactivation (XCI), genetic background) as well as clinical factors (typicality of the Rett presentation and selection criteria for cases, age at sampling, the items included in any phenotypic 'severity score') will affect any genotype–phenotype association found, and it is clear that there is considerable variability in phenotype even for groups of individuals with the same mutation.[16]

One approach to enhancing identification of genotype–phenotype associations is to identify phenotypic dimensions that are, at least in part, separable. There are at least three potentially separable features of the Rett phenotype. The first is the presence or absence of the cardinal diagnostic features, a measure of the 'typicality' of the profile of the disorder. The second is the severity of the associated clinical outcomes that cause impairment and present management difficulties: hand stereotypies, feeding problems, muscle tone, locomotor ability, seizures and scoliosis. Lastly, there is variability in the course of the disorder in terms of age of onset (of regression, for example). These dimensions are not orthogonal (see Schanen et al[16] for examples of correlated features), but may improve our ability to partition the phenotype to enable recognition of the true nature and extent of genotype–phenotype relationships in RS.

Methods

Participant selection and recruitment

In all, 240 patients clinically judged to have RS or a variant thereof were recruited to the study after selection (retrospective and biased towards over-representation of atypical phenotypes) from previously compiled research databases.[19–21] All recruitment, gathering of clinical information and sampling procedures were approved by the MREC for Scotland and all participants' families gave informed consent. MECP2 mutation screening results were available on 190 patients (all female) with a clinical diagnosis of RS (140 cases with classic RS, 50 cases with atypical RS). Patients screened for this study in Glasgow (set A, $N = 62$), and 96 patients in set B (see below) have not previously been reported. Mutation results for 32 patients in set B have been published previously.[22–25] The mean age of the sample was 13.0 years (SD = 8.9 years, range 2–42 years).

Phenotypic information

Comprehensive clinical information from the British Isles Survey of RS (BIRS)[19,20] was available on 180 patients. Information about characteristic Rett behaviours from the parent-completed Rett Syndrome Behaviour Questionnaire (RSBQ)[21] was available on 169 patients. For the purpose of the present analysis, information from the BIRS and RSBQ was extracted to attempt to characterise the Rett phenotype in the manner described below.

Typicality of presentation First, each patient was given a score (from 0 to 7) for the number of necessary diagnostic features present (Diagnostic Criteria Work Group,[10] with allowance made for the recent updates to this schema) and also a score for the number of supportive diagnostic features present (from 0 to 6), extracting this information from the BIRS. For cases where data were missing a minimum score was calculated, on the basis that diagnostic

and supportive features are explicitly requested in the BIRS questionnaire, and are consistently absent if unrecorded, in the experience of the authors (AMK). Almost all patients had been seen by AMK.

Severity of outcome The BIRS includes a clinical severity rating that combines severity ratings (0 = none, 1 = mild, 2 = severe) for feeding problems, muscular tone, locomotor ability, seizures and scoliosis, yielding a total severity score from 0 (least severe) to 10 (most severe).[20] The RSBQ 'hand behaviours' factor, which discriminated most highly (effect size > 2SD) between girls with RS and girls with other types of severe and profound intellectual disability,[21] was chosen as a second measure of severity. Parents/carers were asked to rate each item on a three point scale: 0 (not true), 1 (somewhat or sometimes true) or 2 (very true) according to how well it described their child's behaviour over the last 6 months. The hand behaviours factor included items such as: 'Does not use hands for purposeful grasping', 'Restricted repertoire of hand movement' and 'Hand movements uniform and monotonous'. It yields a factor score between 0 (least severe) and 12 (most severe).

Age of onset Two BIRS items record onset parameters: age at the start of regression (excluding the $N = 21$ cases who did not show regression) and age of first seizures (excluding the $N = 77$ cases who had not had seizures). The principal interest was in identifying cases with particularly early or particularly late onset. For age of onset of regression, regression before age 6 months was deemed 'early' and regression after 30 months was deemed 'late'. For age of first seizures, seizures in the first year of life were deemed 'early' and seizures after 5 years of age were deemed 'late' (see Table 2a).

A number of other clinical variables from the BIRS were also used for investigation of clinical characteristics that might indicate possible genetic or nongenetic aetiology (other event or illness that might have caused neurological damage and facial dysmorphism).

Mutation screening

MECP2 mutation data were either collated with permission from UK NHS molecular diagnostics laboratories and from previous research studies ($N = 128$), or were generated for this study in Glasgow ($N = 62$). All previously published mutation results are indicated in Supplementary Table ST2. For the Glasgow mutation screening, *MECP2* exonic regions were amplified using the PCR (primer sequences are given in Supplementary Table ST1). Exons 2–4 were screened by a combination of restriction digestion for several of the common mutations and direct sequencing of PCR products. All mutations and variants/polymorphisms were confirmed by two methods, or by sequencing product

from at least two independent amplifications in forward and reverse orientations. Further details are given in the Supplementary Information.

MECP2 mutations identified were categorised into 3 classes – missense, early truncating and late truncating (see Supplementary Information for a full explanation of the rationale). Briefly, the late truncating class included all nonsense mutations and frameshifting deletions/indels located downstream of the TRD-NLS (amino acids 255–271), and consisted largely of p.R294X and the group of C-terminal truncating deletions. All these mutations are expected to cause premature truncation of the protein, but may leave it with at least partial function. The early truncating class consisted of nonsense and frameshifting mutations up to and including the TRD-NLS, all larger deletions removing this region, all DNA 'null' mutations involving deletions spanning large parts of the gene, and the single splice site mutation in the dataset (affects splicing of exons 3_4). The TRD-NLS was chosen as the cutoff, as proteins that do not incorporate it are likely to be excluded from the nucleus and are therefore unlikely to be able to perform the major functions of MECP2 as they are currently understood. The missense mutations were treated as a single group in the analyses presented here, as their phenotypic effects are not expected to vary consistently with position in the protein's primary structure.

Statistical analysis

For those cases with identified *MECP2* mutations, group mean scores on the dependent variables were compared using analysis of covariance (ANCOVA), covarying for age. The group mean scores on the dependent variables for cases with common individual mutations were compared using Kruskal–Wallis and Mann–Whitney nonparametric tests, appropriate for small group sizes. Categorical comparisons were conducted using the χ^2-test. Alpha was set at $P < 0.05$ (two-tailed) throughout, with appropriate Bonferroni corrections for multiple comparisons being employed in *post hoc* tests.

Results
MECP2 mutation screening
The mutation data employed in this study were generated from two sources: Set A – generated in Glasgow specifically for this study ($N = 62$); and Set B – patients already screened for *MECP2* mutations ($N = 128$) by other UK clinical genetics and research labs.

Mutation data for both sets of patients are detailed in Supplementary Table ST2. In set A, screened specifically for this study, we identified mutations in 51/57 classic RS cases (89.5%), including two cases found to have large deletions of the region of *MECP2* containing exons 3 and 4 (identified using the MLPA assay). We also identified a mutation in one of the five atypical cases. In Set A, 10

patients had no identifiable mutations, despite MLPA screening in all but three cases (appropriate samples being unavailable). We did not screen exon 1/promoter regions, as recent reports have indicated that only a tiny proportion of RS patients, classic or atypical, harbour clearly pathogenic mutations there.[26]

Six of the patients screened in Glasgow for this study were found to carry previously undescribed mutations (according to the RettBASE website; http://mecp2.chw.edu.au/), c.495_1163del669 (an in-frame deletion of the entire transcriptional repression domain and interdomain-encoding regions), c.792_804del13, c.1054_1259del206, c.1101_1201del101, c.1126_1159del34ins39 and c.1157_1196del40.

Overall, combining the new results generated for this study (Set A) with those gathered from clinical genetics and other research labs (Set B), there were 135 patients with identified mutations. 116 of 140 patients (82.9%) diagnosed clinically with classic and 19 of 50 patients (38%) with an atypical presentation had an identified mutation. Of the remaining classic Rett patients (24 cases with no identified mutation), 13 had been screened for large deletions. The relatively low proportion (compared with other published studies) of cases with p.R255X and p.R270X mutations and the relatively high proportion of cases with the p.R133C mutation can largely be ascribed to the fact that the patient series was nonrandom with respect to the distribution of classic and atypical RS phenotypes. Table 1 summarises mutation detection rates in the different cohorts by clinical diagnosis and the age of each subgroup by the presence/absence of identified mutation, mutation type and common identified mutations.

Characteristics associated with the absence of identified mutations

Cases with early (<6 months) regression or early (<12 months) seizures were less likely to have an identified mutation (see Table 2a; $\chi^2 = 17.1$, $P < 0.001$ and $\chi^2 = 19.6$, $P < 0.001$, respectively). Only 33.3% of early regression cases were *MECP2* mutation-positive and only 18.8% of the early seizures group had identified mutations, significantly

Table 1 Age of sample in relation to diagnosis, mutation presence/absence, mutation category and common individual mutations

	N	Have mutation (%)	No mutation
(a) Proportion of patients with identified MECP2 mutations			
Clinical diagnosis			
Set A (Glasgow)			
Classic Rett syndrome	57	51 (89.5%)	6
Atypical Rett syndrome	5	1 (20%)	4
Set B (other)			
Classic Rett syndrome	83	65 (78%)	18
Atypical Rett syndrome	45	18 (40%)	27
Overall			
Classic Rett syndrome	140	116 (82.9%)	24
Atypical Rett syndrome	50	19 (38%)	31

	N	Mean age in years (SD)
(b) Age distribution by mutation category		
Mutation present/absent		
Mutation identified	135	13.4 (9.1)
No mutation identified	55	11.9 (8.4)
Mutation category		
Missense	50	15.3 (9.0)
Early truncating	56	10.8 (8.4)
Late truncating	29	14.7 (9.9)
Common individual mutations		
R133C	9	17.1 (8.0)
R168X	17	9.0 (10.0)
R255X	8	16.8 (9.0)
R270X	9	12.5 (6.5)
R306C/H	12	16.2 (9.0)
T158M	13	14.6 (9.4)

Table 2 Analysis of phenotypic characteristics of cases with and without identified mutations

	Mutation identified N (row %)	No mutation identified N (row %)	Total N
(a) Age of onset for cases with and without identified mutations			
Age of onset of regression[a]			
Before 6 months	5 (33.3%)	10 (66.7%)	15
6–30 months	115 (81.0%)	27 (19.0%)	142
After 30 months	9 (75.0%)	3 (25.0%)	12
Age of first seizures[b]			
Before 12 months	3 (18.8%)	13 (86.7%)	16
12–60 months	50 (74.6%)	17 (25.4%)	67
After 60 months	22 (75.9%)	7 (24.1%)	29
(b) Unusual clinical features for cases with and without identified mutations			
Event or illness that may have caused neurological deficit[c]			
Yes	11 (44.0%)	14 (56.0%)	25
No	120 (74.1%)	42 (25.9%)	162
Facial dysmorphism			
Yes	15 (55.5%)	11 (44.4%)	26
No	100 (73.5%)	36 (26.5%)	136

[a]$P<0.001$.
[b]$P<0.001$.
[c]$P<0.01$.

fewer than cases with more typical onset. Four cases had both early regression and early seizures, three of whom had no identified mutations. As shown in Table 2, events/ illnesses that might have caused neurological damage ($\chi^2 = 9.7$; $P<0.01$) and facial dysmorphism (marginal significance, $P=0.10$) were both more frequent in cases with no identified mutations. Possible events or illness included a case with cleft palate and possible Pierre Robin syndrome, a case with a thin corpus callosum and a case with prematurity in which a twin died.

Genotype–phenotype associations
The three clinical dimensions – profile, severity and onset – were investigated in relation to type/location of mutation (Table 3).

(i) *Typicality*: cases in the three mutation categories differed in their necessary diagnostic features scores (Table 3a; ANCOVA, $F(3,130) = 4.57$, $P<0.05$; 6.6% variance explained). *Post hoc* tests indicated that cases with early truncating mutations had a higher necessary diagnostic features score than cases with late truncating mutations ($P<0.01$). The groups did not differ in their supportive diagnostic features score.

(ii) *Severity*: The BIRS severity score differed according to mutation type and location (Table 3b; $F(3,123) = 8.42$, $P<0.001$; 12.3% variance explained). *Post hoc* tests indicated that the cases with Early truncating muta-tions had higher Severity scores than both the cases

with Late truncating mutations ($P<0.001$) and cases with missense mutations ($P<0.01$). Mutation type/ location was also significantly associated with RSBQ hand factor score ($F(3,117) = 6.03$, $P<.01$; 9.7% variance explained). *Post hoc* tests indicated that the cases with early truncating mutations had higher Hand factor scores than the cases with late truncating mutations ($P<.001$).

(iii) *Age of onset*: Mutation type/location was not associated with the age of onset variables (Table 3c).

Role of age
The full clinical picture in RS emerges with age. To take account of this, we repeated the ANCOVA analysis using only patients ($N = 48$) who had reached the age of 15, at which time all necessary and supportive features of RS are expected to have appeared if they are going to in each patient.[27,28] Both the necessary and supportive diagnostic features scores differed according to mutation type/loca-tion ($F(3,44) = 5.32$, $P<0.01$; 20.6% variance explained and $F(3,44) = 4.61$, $P<0.05$; 17.8% variance explained, respec-tively). *Post hoc* tests (again, Bonferroni-corrected, $P<0.017$) indicated that on the Necessary diagnostic features score the cases with Early truncating mutations ($N = 14$, mean(SD) = 6.93 (0.27)) scored more highly than the cases with Late truncating mutations ($N = 11$, 6.00 (1.10)). On the Supportive diagnostic features score both the cases with Early truncating mutations ($N = 14$, mean (SD) = 5.14 (0.77)) and those with missense mutations

Table 3 Analysis of genotype–phenotype relationships in mutation positive cases by mutation category

	Early truncating		Missense		Late truncating	
	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N
(a) Typicality of presentation by mutation type and location						
Number of necessary diagnostic features present[a]	6.73 (0.62)	56	6.64 (0.63)	50	6.25 (0.93)	28
Number of supportive diagnostic features present	4.27 (1.29)	56	4.54 (1.33)	50	4.07 (0.77)	28
(b) Severity of outcome by mutation type and location						
BIRS severity score[b]	6.78 (2.66)	51	5.44 (2.76)	48	4.43 (2.46)	28
RSBQ hand factor score[c]	8.89 (1.80)	52	7.93 (2.72)	42	6.76 (3.11)	28

(c) Age of onset for cases by mutation type and location

	Early Truncating N (row %)	Missense N (row %)	Late Truncating N (row %)	Total N
Age of onset of regression				
Before 6 months	4 (80%)	1 (20%)	0 (0%)	5
6–30 months	46 (40.4%)	43 (37.7%)	24 (21.1%)	114
After 30 months	3 (30.0%)	4 (40.0%)	3 (30.0%)	10
Age of first seizures				
Before 12 months	2 (67.3%)	0 (0%)	1 (33.3%)	3
12 to 60 months	21 (42.0%)	20 (40.0%)	9 (18.0%)	50
After 60 months	7 (31.8%)	11 (50.0%)	4 (18.2%)	22

[a] $P < 0.05$.
[b] $P < 0.001$.
[c] $P < 0.01$.

($N = 23$, 5.17 (0.89)) scored more highly than the cases with Late truncating mutations ($N = 11$, 4.27 (0.79)). There were also differences in severity score according to mutation type and location for the BIRS Severity variable ($F(3,43) = 3.97$, $P < 0.05$; 15.6% variance explained) but not the RSBQ Hand factor score. Cases with Early truncating mutations had marginally higher BIRS Severity scores ($N = 14$, mean(SD) = 7.07 (2.34)) than cases with Late truncating mutations ($N = 11$, 4.36 (2.77), $P = 0.02$).

Role of diagnosis
To test whether the associations between mutation type/location and the 'typicality' and 'severity' scores were affected by the inclusion of a range of atypical cases, we repeated the analysis again including cases with classic RS only ($N = 116$). The necessary diagnostic features score differed according to mutation type/location ($F(3,111) = 3.11$; $P < 0.05$; 5.5% variance explained) but none of the *post hoc* comparisons reached significance. There were also differences in severity score according to mutation type/location for the BIRS severity variable ($F(3,105) = 4.27$, $P < 0.05$; 9.4% variance explained) and the RSBQ hand factor score ($F(3,100) = 4.85$, $P < 0.05$; 8.9% variance explained). Cases with Early truncating mutations had significantly higher BIRS severity scores and RSBQ hand factor scores than cases with Late truncating mutations (both $P < 0.01$).

Common individual mutations
For mutations found in eight or more cases, we examined diagnostic, severity and onset scores. Figures 1, 2 and 3 show boxplots of the Necessary and Supportive diagnostic features, the BIRS Severity and RSBQ hand factor scores, and age of regression onset and age of seizure onset, respectively, for cases with p.R133C, p.R168X, p.R255X, p.R270X, p.R306C/H and p.T158M mutations. The number of cases with each type of mutation who had each phenotypic variable differed slightly for different analyses. Appropriate Bonferroni corrections to significance criteria ($P < 0.05/15 = P < 0.003$) for multiple comparisons were made.

Figure 1 shows that all cases with p.R255X mutations score the maximum of seven on the necessary diagnostic features variable and (with the exception of one outlier in each group) the same is true for the mutations p.R270X, p.R306C/H and p.T158M. The cases with p.R133C mutations show the most variation in necessary diagnostic features score. Every group has more variability in Supportive diagnostic features score, although the cases with p.R255X scored 5 or 6 out of 6 with the exception of one case scoring 4. Kruskal–Wallis tests indicated that across the 6 groups the necessary diagnostic features score differed significantly ($P < 0.01$) but none of the *post hoc* comparisons reached significance. The 6 groups did not differ in their supportive diagnostic features score. Figure 2 shows that the cases with p.R133C mutations tended to
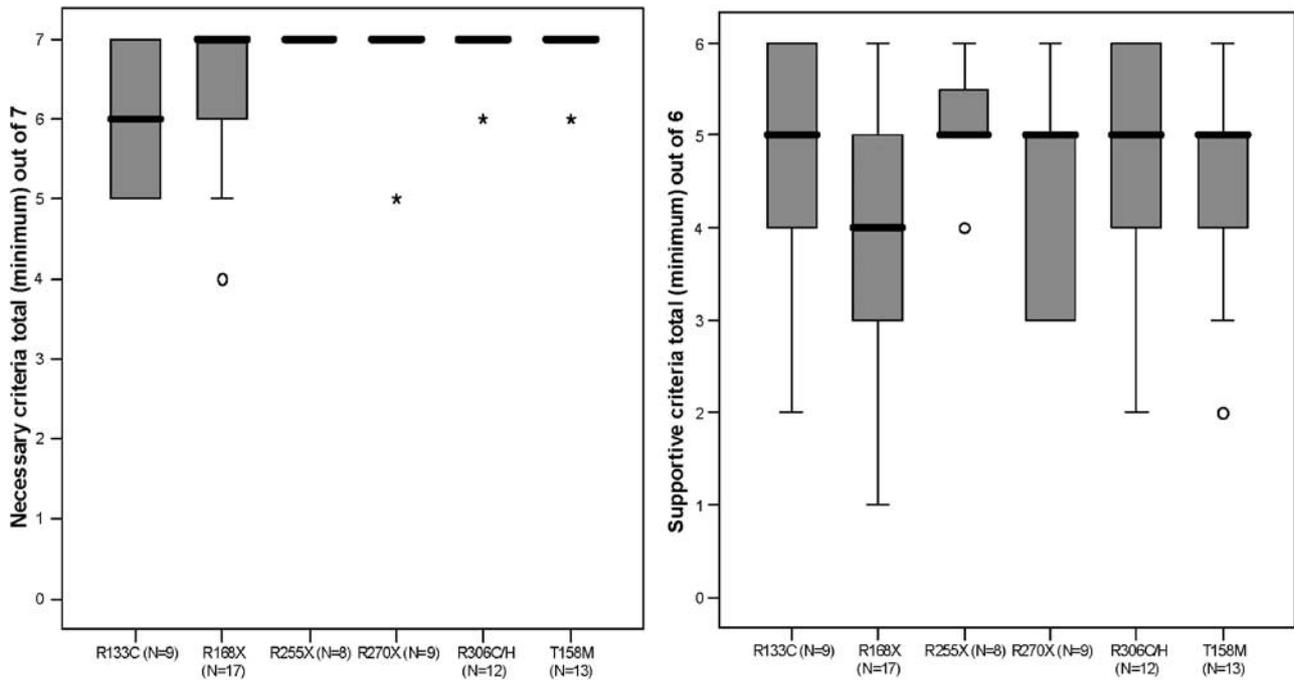
Figure 1 Necessary and supportive diagnostic features scores for common individual mutations. Boxplot displaying number of necessary diagnostic criteria (left) and number of supportive criteria (right) for patients with each of six common mutations. Shaded box indicates the interquartile range and thick black line the median of each distribution. Whiskers indicate the highest and lowest values observed, except for outliers (O = 1.5–3 boxlengths from the upper or lower edge of the box; * = >3 boxlengths from the upper or lower edge of the box). N = number of patients with each mutation.
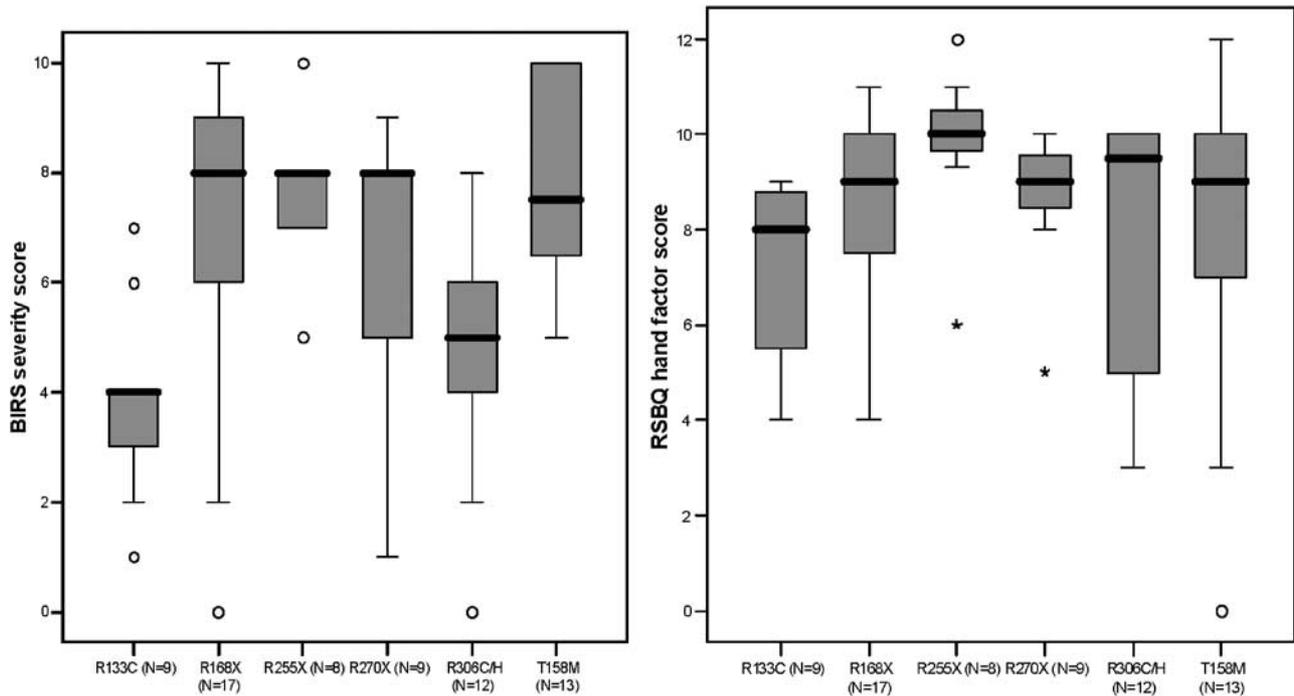

Figure 2 BIRS severity and RSBQ hand factor scores for common individual mutations. Boxplot displaying number of BIRS severity score (left) and RSBQ hand factor score (right) for patients with each of six common mutations. Shaded box indicates the interquartile range and thick black line the median of each distribution. Whiskers indicate the highest and lowest values observed, except for outliers (O = 1.5–3 boxlengths from the upper or lower edge of the box; * = >3 boxlengths from the upper or lower edge of the box). N = number of patients with each mutation.
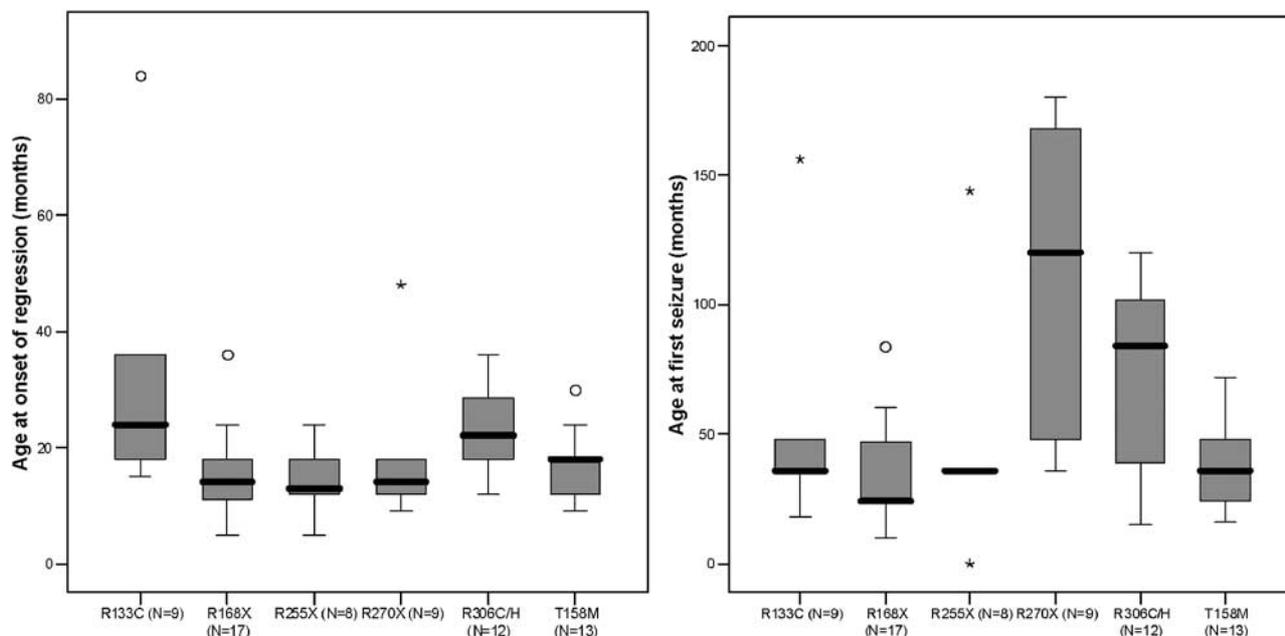
Figure 3 Age of onset of regression and seizures for common individual mutations. Boxplot displaying age at onset of regression in months (left) and age at onset of first seizure (right) for patients with each of six common mutations. Shaded box indicates the interquartile range and thick black line the median of each distribution. Whiskers indicate the highest and lowest values observed, except for outliers (O = 1.5–3 boxlengths from the upper or lower edge of the box; * = >3 boxlengths from the upper or lower edge of the box). N = number of patients with each mutation.

have the lowest BIRS severity scores (with the exception of a few outliers) but that some other mutations, notably p.R168X and p.R270X, show a wide distribution of severity. Across the six groups the BIRS severity score differed significantly ($P<0.001$) and Mann–Whitney *post hoc* comparisons showed that the cases with p.R133C had a significantly lower BIRS severity score than cases with p.R255X and p.T158M mutations. Cases with p.R306C/H had lower BIRS severity scores than cases with p.R255X mutations and those with p.T158M mutations. The pattern was somewhat different for the RSBQ hand factor score: cases with p.R255X and p.R270X mutations tended to have high scores with only one lower scorer (outlier) in each group, while the remaining groups had a wider distribution of scores. Across the six groups the RSBQ hand factor score did not differ significantly. Figure 3 shows that the cases with p.R133C mutations tended to have later age of onset of regression but although the overall six group comparison was significant ($P<0.05$) none of the group-by-group *post hoc* tests reached significance.

Discussion

Characterising the Rett phenotype in dimensional terms allowed us to identify separable associations between the genotype and typicality of presentation, severity of outcome and onset. Firstly, early onset of regression and

seizures were both associated with the lack of an identified mutation. The aetiology of this group of cases with Rett-like presentations requires further investigation. Some may prove to have as yet undetected *MECP2* mutations of substantial effect on gene function, as would be predicted for a high 'severity' phenotype, but it has been determined that a separate X-linked gene (*CDKL5*) accounts for a small proportion,[29] particularly among cases with early-onset seizures. There are 15 cases with early seizures in our dataset. Of these, three have *MECP2* mutations (two truncating, one late truncating). The single case in our study screened for *CDKL5* mutations thus far was negative.[29] Among the cases used in this study, five of those with no identified mutation by sequencing had later been shown to have a large deletion by MLPA or qPCR.[30] Of the 55, 29 'No mutation' cases have been screened for large deletions (8 by MLPA and 21 by qPCR), but a proportion of the remaining 26 cases may turn out to have large deletions/rearrangements on further investigation.[30,31] Whether this subset can be identified phenotypically with any accuracy will await larger studies. Lack of identified mutations was also associated with presence of an illness or event that might be associated with neurological disorder. This may be suggestive of an alternative genetic or nongenetic aetiology despite the fact that some of these cases had a classic Rett-like presentation.

Cases with missense and early truncating mutations had a more typical Rett presentation than cases with late

truncating mutations, although this held for the 'Necessary' but not the 'Supportive' diagnostic criteria. A different pattern was found for the outcome severity scores. Early truncating mutations were associated with more severe outcomes than both missense mutations and late truncating mutations. The size of this association was considerable, accounting for over 10% of the variance in outcome on some measures. This finding was independent of age and held for cases age > 15 years and for cases with a classic Rett diagnosis only. Although our age-covariance approach contrasts with that of Huppke et al[15] who used information about the presentation of girls with RS at the age of 5 years, they found a similar pattern of results, with missense and late truncating mutations being identified with milder phenotypes than other truncating mutations (their mutation categories differed slightly from those used here).

Examining common individual mutations, almost all cases with p.R255X, p.R270X, p.R306C/H and p.T158M mutations had seven out of seven of the necessary diagnostic features. However, in terms of outcome severity several of these mutations did show considerable variability, in line with Schanen et al.[16] In general, the most severe outcomes were found for cases with p.R255X and p.R270X mutations, although p.R168X and p.T158M showed comparable severity on some measures. The least severe outcomes were found for cases with p.R133C mutations. This is largely in line with other studies,[13,16,18] although some findings differ (eg Schanen et al[16] did not find mutation p.R133C associated with a mild phenotype). Thus, whilst individuals with identical mutations can have very different outcomes, mutation type and location do account for a significant proportion of variance in outcome.

Part of the explanation for this individual variation in outcome may be ascribed to the effects of skewing of the XCI ratio in the brain. It has been suggested that XCI ratios that show extreme skewing will be associated either with a visibly milder or visibly more severe phenotype, depending on the direction of skewing, towards or away from expression of the mutated allele in the majority of cells.[14] Such effects of extreme skewing have been reported in several studies,[32–34] but we would also predict that this effect should operate in a graded or threshold-influenced fashion through all or part of the skewing spectrum. Although informative XCI ratio data were available for > 80 patients in this study, we have not presented an analysis here, for two main reasons: (i) a preliminary analysis of the non-directional ratio data did not reveal any striking effects of skewing on the phenotype; and (ii) data on direction of skewing relative to mutant and normal alleles are not yet available. A full understanding of the relationship between blood cell XCI ratio and phenotypic variables will have to await the generation of directional XCI data, a task we are currently engaged in and which will be reported in a future publication.

Conclusion

In this study, we have found that focusing on partly separable aspects of the Rett phenotype has enabled us to detect genotype–phenotype associations that differ to some extent across the phenotypic dimensions considered, providing clues to the effects of genotype on the course, presentation and clinical severity of RS. Early onset was strongly associated with lack of identified mutations, indicating either an alternative genetic or nongenetic cause or the presence of unidentified large deletions. Early truncating mutations (all such, not just those in the TRD-NLS region, p.R255X and p.R270X) that are predicted to disrupt MECP2 protein function to the greatest extent lead both to a more characteristic Rett phenotype and to a more severe clinical outcome. Improvements in our ability to predict outcome in RS patients (both untreated and after therapeutic intervention) will depend on an understanding of the roles and interactions of all relevant factors governing outcome, including definition of phenotypic groupings, appropriately delineated mutation categories, and the effects of XCI ratio and direction of skewing. To achieve this, further studies using very large samples are essential, both to maximise power to detect subtle effects and to reveal associations at the extremes of the phenotypic distribution. Such studies will assist in the identification of the pathogenic processes by which mutations in the MECP2 gene lead to RS.

References

1 Rett A: Uber ein eigenartiges hirnatrophisches syndrome bei hyperammonamie im kindsalter. Wiener Medizinische Wochenschrift 1966; 116: 723–726.
2 Hagberg B, Aicardi J, Dias K, Ramos O: A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand movements in girls: Rett syndrome: report of 35 cases. Ann Neurol 1983; 14: 471–479.
3 Hagberg B, Hagberg G: The Swedish Rett Syndrome series updated March 1996. Geographical distribution at ages 3–19 years. Eur Child Adolesc Psychiatr 1997; 6: 12–13.

4 Kerr AM, Stephenson JB: Rett's syndrome in the west of Scotland. *BMJ* 1985; **291**: 579–582.

5 Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY: Rett syndrome is caused by mutations in X-linked *MECP2*, encoding methyl-CpG-binding protein 2. *Nat Genet* 1999; **23**: 185–188.

6 Kammoun F, de Roux N, Boespflug-Tanguy O *et al*: Screening of *MECP2* coding sequence in patients with phenotypes of decreasing likelihood for Rett syndrome: a cohort of 171 cases. *J Med Genet* 2004; **41**: e85.

7 Christodoulou J, Weaving LS: MECP2 and beyond: Phenotype–genotype correlations in Rett syndrome. *J Child Neurol* 2003; **18**: 669–674.

8 Miltenberger-Miltenyi G, Laccone F: Mutations and polymorphisms in the human methyl CpG-binding protein MECP2. *Hum Mutat* 2003; **22**: 107–115.

9 Kerr AM, Nomura Y, Armstrong D *et al*: Guidelines for reporting clinical features in cases with *MECP2* mutations. *Brain Dev* 2001; **23**: 208–211.

10 Rett Syndrome Diagnostic Criteria Working Group: Diagnostic criteria for Rett syndrome. *Ann Neurol* 1988; **23**: 425–428.

11 Hagberg BA, Skjeldal OH: Rett variants: a suggested model for inclusion criteria. *Paediatr Neurol* 1994; **11**: 5–11.

12 Hagberg B: Clinical delineation of Rett syndrome variants. *Neuropediatr* 1995; **26**: 62.

13 Colvin L, Leonard H, de Klerk N *et al*: Refining the phenotype of common mutations in Rett syndrome. *J Med Genet* 2004; **41**: 25–30.

14 Hoffbuhr K, Devaney JM, LaFleur B *et al*: MeCP2 mutations in children with and without the phenotype of Rett syndrome. *Neurology* 2001; **56**: 1486–1495.

15 Huppke P, Held M, Hanefeld F, Engel W, Laccone F: Influence of mutation type and location on phenotype in 123 patients with Rett syndrome. *Neuropediatric* 2002; **33**: 63–68.

16 Schanen C, Houwink EJF, Dorrani N *et al*: Phenotypic manifestations of *MECP2* mutations in classical and atypical Rett syndrome. *Am J Med Genet* 2004; **126A**: 129–140.

17 Weaving LS, Williamson SL, Bennetts B *et al*: Effects of *MECP2* mutation type, location and X-inactivation in modulating Rett syndrome phenotype. *Am J Med Genet* 2003; **118A**: 103–114.

18 Leonard H, Colvin L, Christodoulou J *et al*: Patients with the R133C mutation: is their phenotype different from patients with Rett syndrome with other mutations? *J Med Genet* 2003; **40**: e52.

19 Kerr AM, Armstrong DD, Prescott RJ, Doyle D, Kearney DL: Analysis of deaths in the British Rett Survey. *Eur Child Adolesc Psychiatr* 1997; **6**: 71–74.

20 Kerr AM, Burford B: Towards a full life with Rett syndrome. *Paediatr Rehab* 2001; **4**: 157–168.

21 Mount RH, Charman T, Hastings RP, Reilly S, Cass H: The Rett Syndrome Behaviour Questionnaire (RSBQ): refining the behavioural phenotype of Rett syndrome. *J Child Psychol Psychiatr* 2002; **43**: 1099–1110.

22 Cheadle JP, Gill H, Fleming N *et al*: Long-read sequence analysis of the MECP2 gene in Rett syndrome patients: correlation of disease severity with mutation type and location. *Hum Mol Genet* 2000; **9**: 1119–1129.

23 Ravn K, Nielsen JB, Skjeldal OH, Kerr A, Hulten M, Schwartz M: Large genomic rearrangements in *MECP2*. *Hum Mutation* 2005, mutations in brief # 791: 1–5.

24 Hampson K, Woods CG, Latif F, Webb T: Mutations in the MECP2 gene in a cohort of girls with Rett syndrome. *J Med Genet* 2000; **37**: 610–612.

25 Vacca M, Filippini F, Budillon A *et al*: Mutation analysis of the MECP2 gene in British and Italian Rett syndrome females. *J Mol Med* 2001; **78**: 648–655.

26 Evans JC, Archer HL, Whatley SD, Kerr A, Clarke A, Butler R: Variation in exon 1 coding region and promoter of *MECP2* in Rett syndrome and controls. *Eur J Hum Genet* 2005; **13**: 124–126.

27 Cass H, Reilly S, Owen L *et al*: Findings from a multidisciplinary clinical case series of females with Rett syndrome. *Dev Med Child Neurol* 2003; **45**: 325–337.

28 Kerr AM: Outcome in Rett syndrome; in Goodyear I, Howlin P (eds): *Outcomes in Neurodevelopmental and Genetic Disorder*. Cambridge: Cambridge University Press, 2002, pp 241–247.

29 Weaving LS, Christodoulou J, Williamson SL *et al*: Mutations of *CDKL5* cause a severe neurodevelopmental disorder with infantile spasms and mental retardation. *Am J Hum Genet* 2004; **75**: 1079–1093.

30 Laccone F, Junemann I, Whatley S, Morgan R, Butler R, Huppke P, Ravine D: Large deletions of the *MECP2* gene detected by gene dosage analysis in patients with Rett syndrome. *Hum Mutat* 2004; **23**: 395.

31 Erlandson A, Samuelsson L, Hagberg B, Kyllerman M, Vujic M, Wahlstrom J: Multiplex ligation-dependent probe amplification (MLPA) detects large deletions in the *MECP2* gene of Swedish Rett syndrome patients. *Genet Test* 2003; **7**: 329–332.

32 Wan M, Lee SS, Zhang X *et al*: Rett syndrome and beyond: recurrent spontaneous and familial *MECP2* mutations at CpG hotspots. *Am J Hum Genet* 1999; **65**: 1520–1529.

33 Young JI, Zoghbi HY: X-chromosome inactivation patterns are unbalanced and affect the phenotypic outcome in a mouse model of Rett syndrome. *Am J Hum Genet* 2004; **74**: 511–520.

34 Braunschweig D, Simcox T, Samaco RC, LaSalle JM: X-chromosome inactivation ratios affect wild-type MeCP2 expression within mosaic Rett syndrome and $Mecp2^{-/+}$ mouse brain. *Hum Molec Genet* 2004; **13**: 1275–1286.