
Hayley Archer, Julie Evans, Helen Leonard, Lyn Colvin, David Ravine, John Christodoulou, Sarah Williamson, Tony Charman, Mark E S Bailey, Julian Sampson, Nicholas de Klerk, Angus Clarke

Introduction: Rett syndrome (RTT) is an X-linked dominant neurodevelopmental disorder, usually caused by mutations in the MECP2 gene. The most common mutations in the gene are p.R168X and p.T158M. The influence of X-chromosome inactivation (XCI) on clinical severity in patients with RTT with these mutations was investigated, taking into account the extent and direction of skewing.

Methods: Female patients and their parents were recruited from the UK and Australia. Clinical severity was measured by the Pineda Severity and Kerr profile scores. The degree of XCI and its direction relative to the X chromosome parent of origin were measured in DNA prepared from peripheral blood leucocytes, and allele-specific polymerase chain reaction was used to determine the parental origin of mutation. Combining these, the percentage of cells expected to express the mutant allele was calculated.

Results: Linear regression analysis was undertaken for fully informative cases with p.R168X (n = 23) and p.T158M (n = 20) mutations. A statistically significant increase in clinical severity with increase in the proportion of active mutated allele was shown for both the p.R168X and p.T158M mutations.

Conclusions: XCI may vary in neurological and haematological tissues. However, these data are the first to show a relationship between the degree and direction of XCI in leucocytes and clinical severity in RTT, although the clinical utility of this in giving a prognosis for individual patients is unclear.

Rett syndrome (RTT) is an X-linked dominant neurodevelopmental disorder, usually caused by mutations in the MECP2 gene. Mutations often arise at CpG hotspots, and the most common mutations in MECP2 found in RTT cases are p.T158M and p.R168X (RettBASE, http://mecp2.chw.edu.au/). RTT has a wide clinical variability in terms of its severity. Studies investigating the association between genotype and phenotype were originally quite inconsistent in their findings. However, with larger studies and increasing numbers of publications, evidence for definite relationships between genotype and phenotype is becoming clearer. Apparent differences in study results often occurred because of the use of different means of classifying and recording clinical severity. Additionally, the effects of X-chromosome inactivation (XCI) status and other epigenetic influences on MECP2 function are likely to have a real influence on the variation in phenotype associated with specific mutations.

XCI occurs early in embryogenesis at the blastula stage, and is usually a random process. The inactive X chromosome is determined as cells become pluripotent; once this has happened, lineages derived from each of these cells will all have the same X chromosome inactivated through a process of methylation. Some studies have shown that there is an increased tendency for skewing of XCI in lymphocytes in RTT when compared with age-matched controls, and that this usually confers a protective effect. The most striking clinical examples of the effects of XCI in RTT are seen in twins with disparate severity and in healthy carrier mothers with skewed XCI (presumed favourable) with affected daughters.

In this study, we adopted a new approach to investigate genotype–phenotype relationships in RTT by exploring the association between clinical severity and the proportion of active mutated allele for the two common MECP2 mutations, p.R168X and p.T158M.

METHODS

Data sources

Patients were recruited from three sources in two countries (table 1): UK cases from two RTT research studies (Cardiff and Glasgow) and Australian cases from the Australian Rett Syndrome Database (www.ichr.uwa.edu.au/rett/aussierett/).

UK patients were recruited from those referred to Cardiff for molecular genetic analysis, or were selected from those participating in a large genotype-phenotype study. Cardiff provides one of the two main diagnostic molecular testing services within the UK. By July 2005, 245 patients had a MECP2 mutation identified and all were invited to participate in the Cardiff Rett Study. DNA samples were already available from the proband, and a blood sample was obtained from at least one parent where this was not already available. In total, samples were available from 18 of 25 Cardiff cases with p.T158M and 28 of 37 cases with p.R168X. An additional seven p.T158M cases and three p.R168X cases were obtained from Scotland (the majority identified in Glasgow).

The Australian Rett Syndrome Database is an ongoing population-based registry, established in 1993, of Australian RTT cases born since 1976. Data are collected from families and clinicians on enrolment and, in combination with molecular

Abbreviations: PCR, polymerase chain reaction; RTT, Rett syndrome; SNP, single-nucleotide polymorphism; 3′ UTR, 3′ untranslated region; XCI, X-chromosome inactivation
X-chromosome inactivation and severity in Rett syndrome

Allele-specific PCR was performed to amplify a fragment of the MECP2 gene. The HUMARA locus was then amplified by polymerase chain reaction (PCR) using fluorescent primers, and analysed using an ABI 3100 automated sequencer and Genotyping software (Applied Biosystems, Foster City, California, USA). Allele peak areas were compared for HpaII digested, McrBC digested and undigested DNA to determine the degree of XCI. HpaII predigestion will prevent PCR amplification of the HUMARA allele on the unmethylated (active) X chromosome; conversely, McrBC predigestion will prevent PCR amplification of the allele on the methylated (inactive) X chromosome. This allows identification of the allele that is present on the more active X chromosome. DNA samples from the parents were genotyped to determine the parental origin of the alleles, and thereby from which parent the more active X chromosome was inherited.

Table 1 Distribution of sample sources and informative data for mutations p.T158M and p.R168X

<table>
<thead>
<tr>
<th>Source</th>
<th>Cases (n)</th>
<th>Informative and clinical data available</th>
</tr>
</thead>
<tbody>
<tr>
<td>T158M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiff</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td>Scotland</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Australia</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td>20</td>
</tr>
<tr>
<td>R168X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiff</td>
<td>28</td>
<td>15</td>
</tr>
<tr>
<td>Scotland</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Australia</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>43</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>83</td>
<td>43</td>
</tr>
</tbody>
</table>

RESULTS

Molecular assay

Of the 83 cases recruited, XCI analysis failed in 17 cases due to poor DNA quality or insufficient DNA. Of the remaining 66 cases, 9 were homozygotic at the androgen receptor locus, and in 14 cases it was not possible to ascertain which parental allele was most active because of missing paternal samples or an approximate 50:50 ratio of XCI. Among these non-contributory samples, two mutations were of maternal origin and three of paternal origin. In the 43 cases in which the most active allele was ascertained, 21 mutations were of paternal origin and none were of maternal origin. The normal allele was favoured in 18 of 21 (86%) cases. In the other 22 cases there were no informative SNP data to allow determination of parental origin.

Clinical data

For the 28 UK cases there were 29 missing individual items, equivalent to 5.2% of all items, for the Kerr profile and 26 missing items, equivalent to 9.3% of all items, for the Pineda scale. In the Kerr profile items relating to height, joint contractures, mood and sleep disturbance, and in the Pineda scale, items relating to age at sitting and acquisition and loss of ambulation, had the most missing items.

Phenotype scoring, data management and statistical analysis

Genotype–phenotype investigations by our groups have until now relied on scoring systems previously established by others. These scales use simple and objective measures that are not dependent on clinical skill or intuition. We refer to these scoring systems as the Kerr profile and the Pineda and Percy scores. As our previous work has shown that the Pineda score, which takes into account historical data as well as current clinical status, is the most sensitive to change, we used both the Kerr profile and Pineda Severity scores. These scores had already been generated for the Australian cases based on data collected in 2000. For the UK cases, when necessary, telephone interviews were arranged with families (conducted by HA) to obtain the necessary information and to supplement the details available from clinical sources.

Statistical analysis was undertaken using Stata V.9.0. For the 15 Australian cases the previously generated scores, where missing data were replaced with the average score for that item for the Rett classification type (ie, classic, mild atypical or early-onset atypical), were used. For the 28 UK cases, an imputation routine, in which missing values for each scale were imputed on the basis of other values in that scale, was used to estimate missing values and to maximise the use of the scarce data. For this, imputation data that had been collected on a further 102 UK cases (either with p.T158M or p.R168X mutations but non-informative XCI data, or with another mutation) were included.

Linear regression was undertaken with the Kerr profile and Pineda scales as continuous variables and the percentage of the mutant allele as the independent variable. A case with both p.R168X and trisomy 21 was excluded from the analysis to avoid the potential confounding effects arising from coexisting Down’s syndrome.

Statistical analysis, data management and genetic analysis

Allele-specific PCR was performed to amplify a fragment including the position of the mutation. The allele-specific PCR products were sequenced to determine which haplotype contained the mutation. It was then possible to determine the parent of origin by genotyping the SNP in the parents. By combining the parent of origin with the XCI results, the percentage of cells expected to be expressing the mutant MECP2 was calculated. Where parent of origin could not be determined, but it was known from which parent the more active X chromosome was inherited, it was assumed that the mutation was of paternal origin.

Determination of the parental origin of the mutation

Intron 3 and the 3’ untranslated region (3’UTR) of MECP2 were sequenced in the patients to find informative single-nucleotide polymorphisms (SNPs). When an SNP was found, allele-specific PCR primers were designed so that the base on the 3’ end of each primer was complementary to only one allele of the SNP (table 2). SNPs are numbered relative to the position of the stop codon (+878C→G) or the end of exon 3 (+266C→T, +516C→T and +648A→G). All SNPs are available in dbSNP (http://www.ncbi.nlm.nih.gov/SNP/): +878C→G = rs3027924, +266C→T = rs2075596, +516C→T = rs3027931 and +648A→G = rs3850326.

Allele-specific PCR was performed to amplify a fragment including the position of the mutation. The allele-specific PCR products were sequenced to determine which haplotype contained the mutation. It was then possible to determine the parent of origin by genotyping the SNP in the parents. By combining the parent of origin with the XCI results, the percentage of cells expected to be expressing the mutant MECP2 was calculated. Where parent of origin could not be determined, but it was known from which parent the more active X chromosome was inherited, it was assumed that the mutation was of paternal origin.

Phenotype scoring, data management and statistical analysis

Genotype–phenotype investigations by our groups have until now relied on scoring systems previously established by others. These scales use simple and objective measures that are not dependent on clinical skill or intuition. We refer to these scoring systems as the Kerr profile and the Pineda and Percy scores. As our previous work has shown that the Pineda score, which takes into account historical data as well as current clinical status, is the most sensitive to change, we used both the Kerr profile and Pineda Severity scores. These scores had already been generated for the Australian cases based on data collected in 2000. For the UK cases, when necessary, telephone interviews were arranged with families (conducted by HA) to obtain the necessary information and to supplement the details available from clinical sources.

Statistical analysis was undertaken using Stata V.9.0. For the 15 Australian cases the previously generated scores, where missing data were replaced with the average score for that item for the Rett classification type (ie, classic, mild atypical or early-onset atypical), were used. For the 28 UK cases, an imputation routine, in which missing values for each scale were imputed on the basis of other values in that scale, was used to estimate missing values and to maximise the use of the scarce data. For this, imputation data that had been collected on a further 102 UK cases (either with p.T158M or p.R168X mutations but non-informative XCI data, or with another mutation) were included.

Linear regression was undertaken with the Kerr profile and Pineda scales as continuous variables and the percentage of the mutant allele as the independent variable. A case with both p.R168X and trisomy 21 was excluded from the analysis to avoid the potential confounding effects arising from coexisting Down’s syndrome.
Correlation between proportion of mutated allele activity and clinical severity

Among cases with the p.R168X mutation, clinical severity increased significantly as the proportion of cells with an active mutant allele increased (fig 1, table 3). This effect was observed with both the Pineda ($p = 0.05$) and Kerr ($p = 0.014$) scoring systems. When the data were adjusted for age, the effect with the Kerr profile became more significant ($p = 0.005$) and the correlation was not significant ($p = 0.15$).

As the proportion of cells with an active mutant allele increased in cases with the p.T158M mutation, the Pineda score increased significantly ($p = 0.04$), again indicating greater severity (fig 2, table 3), but the correlation was not significant in the case of the Kerr profile score. The magnitude of the association was the same as with the p.R168X mutation, both with and without age adjustment.

For both mutations, repeating the analysis using the logistic transformation of the percentage of the mutant allele active had minimal effect on the results. In our dataset, these two mutations have very similar Kerr and Pineda scores, indicating that on average they are of similar severity. Therefore, we also assessed the effect of X inactivation on the two groups combined. This has shown that the increasing severity indicated by both the Pineda and the Kerr scales was associated with the effects of X inactivation both when not adjusted ($p = 0.005$ for either scale) and when adjusted for age ($p = 0.005$ for Pineda and $p = 0.001$ for Kerr), as well as with the direction of XCI skewing. The $R^2$ (square of the correlation coefficient) was 19% for both; the direction of XCI has therefore been shown to account for 19% of the variance in phenotypic outcome, a substantial effect.

**DISCUSSION**

In this collaborative international study, our data show a relationship between the extent and direction of skewing of XCI and clinical severity for cases with the p.T158M or p.R168X mutation in the MECP2 gene. Both p.R168X and p.T158M are typically associated with a severe presentation, but both have also been reported in mildly affected cases with the preserved speech variant of RTT. This suggests that the clinical variability associated with these mutations is broad, and that this may be due to the role of epigenetic influences, including XCI. This contrasts with those with the p.R133C mutation (usually mildly affected) or those with the p.R255X mutation (usually more severely affected), where there is less variability in terms of severity. The sensitivity of the available severity scales may not be great enough to detect subtle differences caused by the effects of XCI status.

An RNA-based approach to determine the parent of origin would have resulted in a higher ascertainment, but was not possible in this study as fresh patient samples were not available for RNA extraction. The method used to determine XCI status and parental origin of the mutation in this study was straightforward and did not require further fresh blood samples from patients. However, there were losses at each stage of the testing process. The androgen receptor assay is usually uninformative in approximately 20–30% of cases, although
in our experience this was true for 9 of 66 (15%) cases with DNA of suitable quantity and quality for analysis. Further cases (14/66, 21%) in this study were lost because of difficulty in ascertaining the most active parental allele, usually in cases where only one parental sample was available. In samples where we could determine the parental origin of the MECP2 mutation, 92% (24/26) were of paternal origin, as found in previous studies. Paternal origin of mutation was assumed in 51% (22/43) of cases included in the investigation, so it is likely that a small number of these cases had the incorrect allele assigned as the active one. It is interesting that so few cases were identified in which >60% of the mutant MECP2 allele was active. There might be a selective advantage in growth and cell division for those haematopoietic cells using the X chromosome with the intact MECP2 locus, but this is not confirmed. There may also be some selection against the long-term survival of females, with >60% of their cells using the altered copy of the MECP2 locus, or their phenotype may be less typical of classic RTT. In either case, they might be under-represented in our clinically selected group. An additional part of the explanation may lie in the fact that, where there is skewing of XCI, it is usually the paternal X chromosome that is preferentially inactivated, and, combined with the fact that most mutations are paternal, this leads to the conclusion that the cases that are more heavily skewed are less likely to express the mutant copy of MECP2.

In this study, we found that XCI almost always favours the normal allele, as has been reported in mouse models of RTT, and this is the pattern commonly seen in other X-linked diseases. Although making assumptions based on XCI status in lymphocytes might be considered a compromise in this study, it was neither realistic nor ethical to obtain neurological DNA of suitable quantity and quality for analysis. Further cases were identified in which MECP2 mutations associated in females with classic RTT. We might expect any male patient to represent the most severe phenotype associated with a particular mutation. There is a published report of a male patient with an early truncating mutation, which would be expected to be associated with a severe phenotype, yet he was relatively mildly affected at 11 years. There was no evidence for somatic mosaicism, although this could not be completely ruled out. This example suggests that there may be other background genetic and epigenetic influences that contribute significantly to the clinical manifestations of RTT, but the exact nature of these additional factors is yet to be determined.

Despite the limitations, we were able to show in DNA extracted from peripheral leucocytes that there was a correlation between the proportion of the mutated allele used by an individual and the clinical severity of a predominantly neurodevelopmental phenotype. Combining the efforts of three countries was essential for this study, but wider collaborative efforts will be required to study the phenotype associated with other common mutations in a similar way. We had hoped to look at the effects of the parent of origin based on the suggestion that this may also modulate the phenotype. As mutations of maternal origin are infrequent, substantially greater numbers of patients would be required to deal with this question.

ACKNOWLEDGEMENTS

We thank all the families and clinicians who have contributed to the study, the Rett Syndrome Association UK, Dr Alison Kerr and all other clinicians, who recruited patients from the UK, and the Australian Paediatric Surveillance Unit and the Rett Syndrome Association of Australia, who facilitated case ascertainment in Australia. We also acknowledge the molecular work undertaken by Dr Mark Davis.

Authors’ affiliations

Hayley Archer*, Julie Evans*, Julian Sampson, Angus Clarke, Institute of Medical Genetics, Cardiff University, University Hospital of Wales, Cardiff, UK
Helen Leonard, Lyn Colvin, Nicholas de Klerk, Telethon Institute of Child Health Research, Centre for Child Health Research, The University of Western Australia, Perth, Western Australia, Australia
David Ravine, Western Australian Institute of Medical Research, School of Medicine and Pharmacology, University of Western Australia, Perth, Western Australia, Australia
John Christodoulou, Sarah Williamson, Western Sydney Genetics Program, The Children’s Hospital at Westmead and Discipline of

---

### Table 3

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Scoring system</th>
<th>Cases (n)</th>
<th>Mean (SD)</th>
<th>Minimum value</th>
<th>Maximum value</th>
<th>r</th>
<th>*Reg. coefft.</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>p.R168X</td>
<td>Pineda scale unadjusted</td>
<td>23</td>
<td>16.87 (3.29)</td>
<td>10.93</td>
<td>21.69</td>
<td>0.42</td>
<td>0.11</td>
<td>0.05</td>
</tr>
<tr>
<td>p.R168X</td>
<td>Pineda scale age adjusted</td>
<td>23</td>
<td>22.65 (4.40)</td>
<td>14.25</td>
<td>30</td>
<td>0.49</td>
<td>0.16</td>
<td>0.02</td>
</tr>
<tr>
<td>p.T158M</td>
<td>Kerr profile unadjusted</td>
<td>20</td>
<td>17.64 (4.78)</td>
<td>9.68</td>
<td>25</td>
<td>0.45</td>
<td>0.16</td>
<td>0.05</td>
</tr>
<tr>
<td>p.T158M</td>
<td>Pineda scale age adjusted</td>
<td>20</td>
<td>24.90 (6.93)</td>
<td>10</td>
<td>33</td>
<td>0.35</td>
<td>0.17</td>
<td>0.13</td>
</tr>
<tr>
<td>p.T158M</td>
<td>Kerr profile age adjusted</td>
<td>43</td>
<td>17.48 (4.17)</td>
<td>9.68</td>
<td>25</td>
<td>0.44</td>
<td>0.13</td>
<td>0.003</td>
</tr>
<tr>
<td>p.R168X</td>
<td>Pineda scale age adjusted</td>
<td>43</td>
<td>23.70 (5.76)</td>
<td>10</td>
<td>33</td>
<td>0.44</td>
<td>0.17</td>
<td>0.003</td>
</tr>
<tr>
<td>p.R168X</td>
<td>Kerr profile age adjusted</td>
<td>43</td>
<td>23.70 (5.76)</td>
<td>10</td>
<td>33</td>
<td>0.44</td>
<td>0.17</td>
<td>0.003</td>
</tr>
</tbody>
</table>

*Coefficient of the regression of severity score on percentage of mutant allele active—represents the increase in severity score for each percentage point increase.

---

* Indicates significance. Values were compared using Spearman’s rank correlation coefficient to determine the strength and direction of the relationship.
REFERENCES


