

SCIENTIFIC REPORTS



OPEN

From Function to Phenotype: Impaired DNA Binding and Clustering Correlates with Clinical Severity in Males with Missense Mutations in *MECP2*

Received: 01 June 2016
Accepted: 11 November 2016
Published: 08 December 2016

Taimoor I. Sheikh^{1,2}, Juan Ausió³, Hannah Faghfoury⁴, Josh Silver⁴, Jane B. Lane⁵, James H. Eubanks^{6,7,8}, Patrick MacLeod⁹, Alan K. Percy⁵ & John B. Vincent^{1,2,10}

Mutations in the *MECP2* gene cause Rett syndrome (RTT). MeCP2 binds to chromocentric DNA through its methyl CpG-binding domain (MBD) to regulate gene expression. In heterozygous females the variable phenotypic severity is modulated by non-random X-inactivation, thus making genotype-phenotype comparisons unreliable. However, genotype-phenotype correlations in males with hemizygous *MECP2* mutations can provide more accurate insights into the true biological effect of specific mutations. Here, we compared chromatin organization and binding dynamics for twelve MeCP2 missense mutations (including two novel and the five most common MBD missense RTT mutations) and identified a correlation with phenotype in hemizygous males. We observed impaired interaction of MeCP2-DNA for mutations around the MBD-DNA binding interface, and defective chromatin clustering for distal MBD mutations. Furthermore, binding and mobility dynamics show a gradient of impairment depending on the amino acid properties and tertiary structure within the MBD. Interestingly, a wide range of phenotypic/clinical severity, ranging from neonatal encephalopathy to mild psychiatric abnormalities were observed and all are consistent with our functional/molecular results. Overall, clinical severity showed a direct correlation with the functional impairment of MeCP2. These mechanistic and phenotypic correlations of MeCP2 mutations will enable improved and individualized diagnostics, and may lead to personalized therapeutic interventions.

Rett syndrome (RTT; MIM#312750) is an X-linked neurological disorder present in ~1:10,000–15,000 girls¹. RTT children often appear to develop normally in early infancy; however, changes in patterns of mental and social development usually begin to appear between 6 and 18 months of age. Slower head growth and loss of muscle tone are common early symptoms, along with more general problems with gross motor development and coordination. Other clinical features and symptoms may include repetitive hand movements, scoliosis, constipation, excessive saliva, intellectual disability, stiff-legged gait, periodic breathing, seizures, and typically little or no verbal skills. Although RTT is known to be caused by dysfunction of the methyl-CpG-binding protein-2 gene, *MECP2*, there is a broad range in severity and co-morbidities. In affected heterozygous females, phenotypic

¹Molecular Neuropsychiatry & Development (MiND) Lab, Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, ON, Canada. ²Institute of Medical Science, University of Toronto, Toronto, ON, Canada. ³Department of Biochemistry and Microbiology, University of Victoria, BC, Canada. ⁴The Fred A. Litwin Family Centre in Genetic Medicine, University Health Network and Mount Sinai Hospital, Toronto, ON, Canada. ⁵Civitan International Research Center, University of Alabama at Birmingham, AL, USA. ⁶Division of Genetics and Development, Toronto Western Research Institute, University Health Network, Toronto, ON, Canada. ⁷Department of Physiology, University of Toronto, Toronto, ON, Canada. ⁸Department of Surgery (Neurosurgery), University of Toronto, Toronto, ON, Canada. ⁹Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada. ¹⁰Department of Psychiatry, University of Toronto, Toronto, ON, Canada. Correspondence and requests for materials should be addressed to J.B.V. (email: john.vincent@camh.ca)

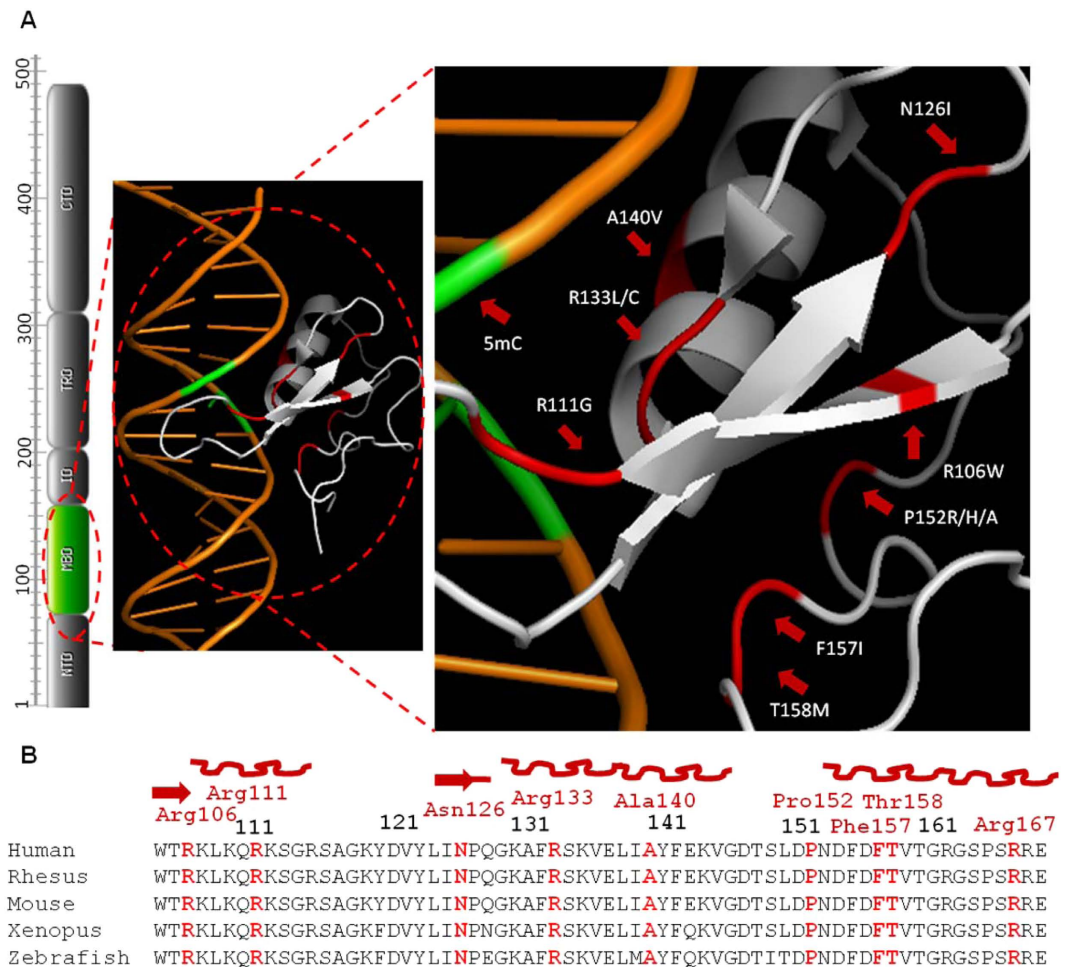


Figure 1. Structural analysis of human MeCP2 MBD bound at brain-derived neurotrophic factor (BDNF) promoter. (A) MeCP2 domain structure with approximate amino acid number in each domain³⁹ and the only highlighted domain (green) is MBD. (B) Human MECP2-MBD crystallographic structure¹¹ illustrated using PyMol Molecular Graphic System (PDB accession code: 3C2I) showing MBD bound at methylated double stranded DNA sequence of BDNF promoter. Location of amino acids substituted such as R106W, R111G, N126I, R133L/C, A140V, P152A/H/R, F157I, T158M are indicated with arrows (highlighted as Red). (C) Multiple sequence alignment (MSA)⁴⁰ of the full sequence of MBD of five vertebrates (NCBI GeneBank accession numbers: *Homo sapiens* gi|1708973|sp|P51608.1, *Macaca fascicularis* gi|50401118|sp|Q95LG8.1, *Mus musculus* gi|12585281|sp|Q9Z2D6.1, *Xenopus laevis* gi|4105999|gb|AAD02651.1, and *Danio rerio* gi|62204887|gb|AAH93116.1) to show conservation at the MeCP2 missense mutation residues (red).

severity may vary with patterns of non-random X chromosome inactivation across the brain and other relevant tissues². Detailed clinical criteria used for RTT diagnostics are described elsewhere³.

MECP2, on Xq28, was identified in 1999 as the gene responsible for RTT⁴. MeCP2 was initially identified due to the selective binding to DNA sequences with methylated CpG dinucleotides^{5,6}, however, MeCP2 has also shown moderate binding affinity to un-methylated DNA⁷. MeCP2 competes with histone H1 and binds linker DNA^{8,9}, which suggests that MeCP2 and histone H1 may share similar functions with respect to chromatin organization. MeCP2 is a 60–80% unstructured, intrinsically disordered protein that shows coil-like hydrodynamic properties in solution^{8,10}. The hydrophobic patch within the MBD forms hydrophobic interactions with methylated CpG. The X-ray crystallographic structure of the MBD bound with DNA has highlighted unique structural roles for various residues, which may be important in the DNA-binding dynamics of MeCP2¹¹ (Fig. 1A).

Mutations in the MeCP2 gene are commonly linked to RTT, but MeCP2 has also been associated with a number of other developmental disorders, including in males¹². This suggests that it may play a central role in the post-natal development of the human brain¹³. There are ~160 different single nucleotide variations (SNVs) that have been reported in ~1593 patients, out of which ~943 (~59.01%) have 60 different SNVs in the MBD¹⁴. A more comprehensive understanding of the functional effects of MeCP2/MBD missense mutations and genotype/phenotype will assist the search for therapeutic compounds targeting these deficits, as well as enable a clearer picture of the downstream molecular and/or cellular disruptions in common between RTT patients. T158, R133, R106 and P152 are common Rett related mutation hotspots within the MBD of MeCP2. According to Rettdbase, among MBD missense mutations, Thr158Met is the most common (8.79% of reported cases, also the most common

	Case I	Case II	Case III	Case IV	Case V	Case VI	Case VII	Case VIII	Case IX	Case X ^a	Case XI
Genotype	T158M	T158M	T158M	T158M	R133C	F157I	N126I	P152H	R167W	A140V	P152A
Condition at birth	Normal	Normal	Normal	Normal	Normal	Abnormal	Abnormal	Unknown	Normal	Normal	Normal
First Reported abnormality	<10 M	4–5 months	postnatal	postnatal	~8M	Neonatal period	Early infancy	4 YR	1.5 YR	2 YR & older	None ^b
Development	Delayed	Delayed	Delayed	Delayed	Delayed	Minimal	Minimal	Delayed	Delayed	Delayed	Normal
Head Circumference	Lower than normal	Lower than normal	Lower than normal	Not known	Normal at birth; <2% age 2 YR	Normal at birth; <2% age 9 M	Normal at birth; <2% age 6 M	Lower than normal	Normal	Normal	Lower than normal
Muscular abnormalities	Hypotonia	Hypotonia	Hypotonia	Hypertonia, limb spasticity	Hypotonia	Hypertonia	Hypotonia, then hypertonia	General muscle weakness	Normal	Hypotonia, then spastic paraparesis	None ^c
Motor skills	Never learned	Never learned	Poor head control	Regression of motor skills	Regression at age 2.5 YR	Never learned	Minimal skills	Regression	Mild delay	Mild delay	Normal
Epileptic Seizures	Yes	Yes	Not reported	Yes	Yes	Absent	Absent	No	Absent	Described in some	No
Age (Years)	Died at ~3	Died at ~1.5	Died at ~0.92	Died at ~1.2	13.33	11.0 ^d	Died 3.17	~40	21.17	27 – middle age	~28
Reason of death	Respiratory infection	Respiratory infection	Respiratory infection	Respiratory infection	N.A	N.A	Respiratory failure	N.A	N.A	N.A	N.A
Language skills	N.A	N.A	N.A	N.A	None	None	None	Several words	Short phrases	Short phrases or simple sentences	Able to write and read
CSS ^e	36	43	31	47	37	37	36	18	12	10 ^f ; 1–5 ^g	1
Ref.	15	15	16	17	Present study	Present study	Present study	Present study	Present study	Numerous ^{41–43}	18

Table 1. Clinical features of study participants. ^aHere we report an amalgamation of data for several Ala140Val cases from the literature^{41–43} plus one (Case X) from the RDCRC Natural History Study database. ^bNo physical abnormality. ^cNo muscular abnormality reported. ^dOn home ventilator. ^eClinical severity scores. ^fRDCRC Natural History Study Database. ^gbased on limited data from literature^{41–43}.

mutation overall), followed by R133C (4.24%), R106W (2.76%), P152R (1.41%) and A140V (0.66%)¹⁴. T158M has been reported in a few hemizygous male patients^{15–17}. P152R has been reported as a pathogenic variant only in females, showing classical, typical, and atypical Rett symptoms. In a single study, P152A was reported as a hypomorphic mutation in a male patient¹⁸. R106, R111, N126, R133, A140, P152, F157, T158 and R167 are conserved across all vertebrates (Fig. 1B), and substitution at these structurally important residues may affect the ability of MeCP2 to bind properly to its interacting targets, which may affect gene expression regulation and/or chromatin organization as well as its mobility in the cellular environment. Here, using cells transfected with several mutant MeCP2-GFP constructs in order to model the different mutations as a more practical approach than using patient-derived cells, we studied: 1) the effect of rare and common MBD mutations on chromatin organization and on DNA-binding kinetics; and 2) the correlation between clinical severity and the cellular effect for males with the MBD missense mutations N126I, R133C, A140V, P152A, P152H, F157I, T158M and R167W.

Results

Study participants' history and clinical characteristics. Detailed patient ascertainment and clinical descriptions are in the supplementary material. A comparison of clinical conditions of males with the hemizygous MeCP2 mutations N126I, R133C, A140V, P152A, P152H, F157I, T158M, and R167W are highlighted in Table 1.

MeCP2 localization at chromocenters. Figure 2A represents confocal stacks of all recombinant MeCP2 mutant and control proteins (GFP), DNA (counterstained with DAPI), and merged images representing the patterns of co-localization of both fluorophores. Although each mutant represents its unique binding and clustering pattern, we have divided the patterns into two scenarios, 1) mutants such as R111G, R133L and R106W showing complete mislocalization, likely due to impeded binding of mutant MeCP2 with DNA; 2) mutants such as P152R, F157I and N126I showing higher number and decreased size of chromocenters, likely due to impaired chromatin compaction (see below).

Merged images of chromatin DNA (DAPI) and MeCP2 fusion protein (GFP) for R111G, R106W, R133L and T158M show a small number of distinctively identifiable and clustered chromocenters (Fig. 2A). In the quantitative co-localization analysis of transfected MeCP2 and chromatin, median Pearson correlation coefficient (PCC) values (r_p) for the R111G, R106W, R133L and T158M constructs were calculated as $r_p = 0.26, 0.40, 0.46$ and 0.57 respectively ($p \leq 0.0005$). In contrast, constructs for P152R, F157I, P152H, P152A, N126I, R133C, P152A and R167W had PCC values of $r_p = 0.68, 0.69, 0.71, 0.72, 0.72, 0.82$ and 0.81 respectively (Fig. 2B,C). PCC values of $r_p \leq 0.60$ were not considered as mislocalized (Fig. 2C). Overall, we observed three different patterns among mutants, 1) impaired clustering; 2) abolished clustering; and 3) abolished binding (Fig. 2D).

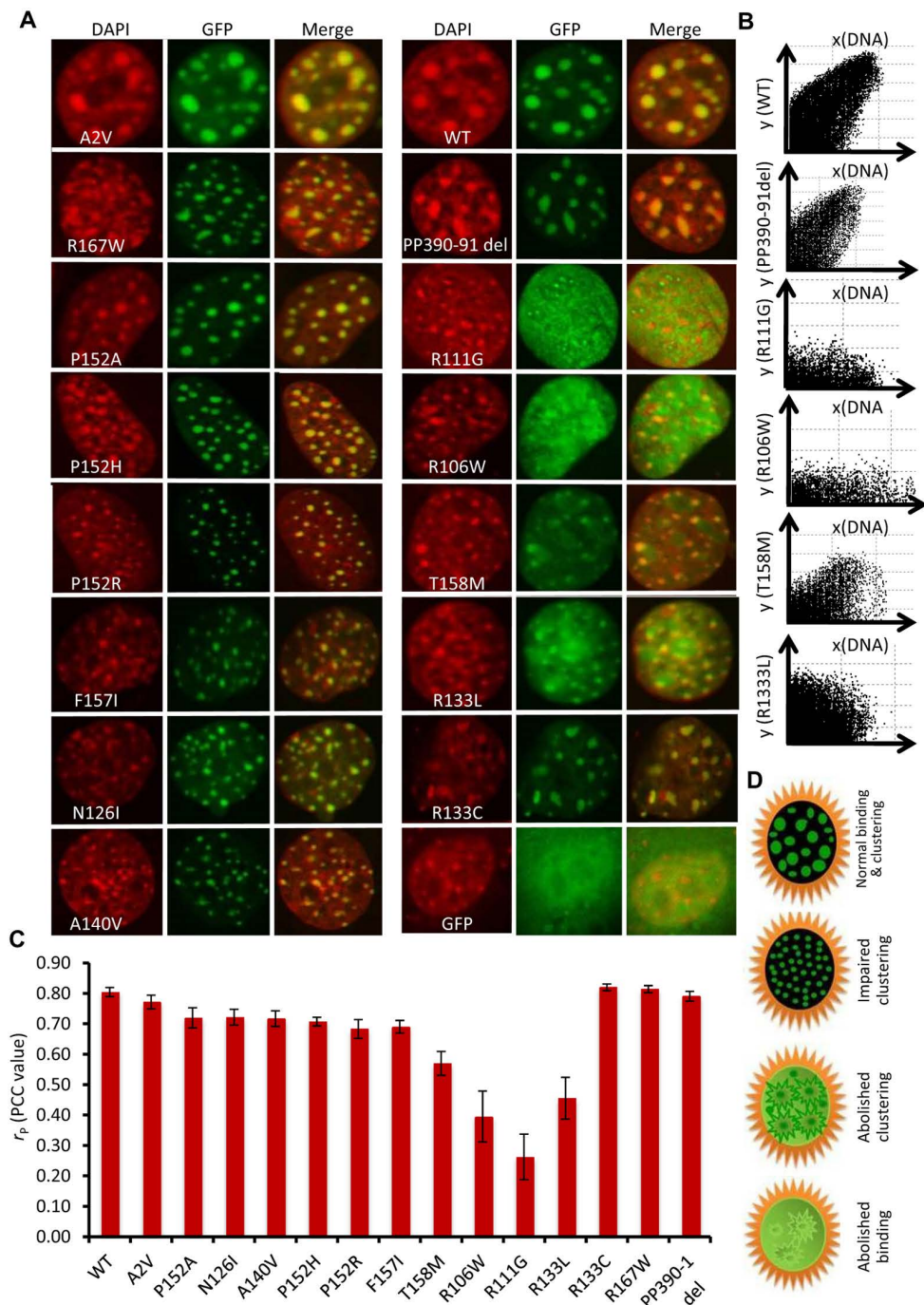


Figure 2. Comparison of the binding and clustering pattern of wild type and mutant MeCP2 expressing in C2C12 cells. (A) Confocal image stacks representing Left to right), column 1 DAPI (red), column 2 GFP (green), column 3 merge (red + green), column 4 DAPI (red), column 5 GFP (green) and column 6 merge (red + green). Left Row 1 A2V, Right Row 1 wild type, Left Row 2 R167W, Right Row 2 PP390-91del, Left Row 3 P152A, Right Row 3 R111G, Left Row 4 P152H, Right Row 4 R106W, and Left Row 5 P152R. Right Row 5 T158M, Left Row 6 F157I, Right Row 6 R133L, Left Row 7 N126I, Right Row 7 R133C, Left Row 8 A140V, and Right Row 8 GFP only. Note that all images shown here are representative of the authors' observations; averaged data from all images are presented in the quantitative analysis, illustrated in Fig. 3 and discussed in the main text. (B) Scatter plot of red and green pixel intensities showing co-localization of DNA (DAPI) and recombinant MeCP2 protein (GFP). (C) Pearson's correlation coefficient (PCC) r_p values of chromocenters (DAPI) and recombinant protein (GFP) were plotted to represent quantitative co-localization of DNA and recombinant MeCP2 ($n = 10$, 0.6 or higher = positive co-localization; < 0.6 weak positive co-localization; ≤ 0.5 No or negligible co-localization); \pm SEM is shown. (D) Cartoons shown (Right \rightarrow bottom) different binding and clustering patterns with normal binding and clustering (Top, wild type), impaired clustering (Middle 1, P152R/ N126I), completely abolished clustering (Middle 2, F157I/T158M) and abolished binding (Bottom Right, R111G/R133L).

MeCP2 binding at chromocenters. From the co-localization analysis of chromatin and mutant MeCP2, we found abolished co-localization for mutants such as R111G, R106W, T158M and R133L as mutant protein unable to localize at the chromocenter (Fig. 2A–D). Any further quantitative analysis of these mutants was not possible. In the quantitative analysis, comparisons of number and size of distinct identifiable chromocenters are shown in graphical form (Fig. 3). *P*-values (two tailed Mann Whitney non-parametric U test) are indicated where statistically significant (<0.05) for each MeCP2 mutation in comparison with WT MeCP2. In the initial quantitative analysis we compared transfected and untransfected cells using the DAPI channel to quantify chromocenters, and the GFP channel to quantify recombinant MeCP2_E1. Chromocenters of transfected cells over-expressing MeCP2_E1 showed significant clustering of chromocenters in comparison with untransfected cells (Fig. 3A). Quantitative number and size analysis reveal more but significantly smaller sized chromocenters in the untransfected cells (Fig. 3B).

We also compared number and size of chromocenters of wild type MeCP2_E1 and all mutant MeCP2 constructs included in this study. Due to extremely low quantity of distinctly identifiable chromocenters, R111G, R106W, R133L and T158M were not considered in the chromocenter number and size analysis. This lack of identifiable chromocenters suggests either loose or no binding, in addition to mislocalization, leading to a complete loss of chromatin clustering (Fig. 2). The chromocenter number and size for R133C was also slightly lower than WT, but non-significant (two tailed Mann Whitney non parametric U test *p* value ≥ 0.05 ; Fig. 3C,D).

Cells expressing N126I, A140V, P152R, P152H, F157I and R167W MeCP2 showed a significant increase in the number of chromocenters ($p \leq 0.005$). For P152A, on the other hand, the average chromocenter numbers were not significantly different from WT and other controls (Fig. 3C). Similarly, for the chromocenter size analysis, N126I, A140V, P152R, P152H, F157I and R167W showed a significant decrease in chromocenter size ($p \leq 0.005$), whereas for P152A chromocenter size was not significantly different from WT and other controls (Fig. 3D).

Real-time recovery and binding dynamics of Rett mutants. To further study the effect on protein mobility and binding dynamics, we selected several mutants that were typical for the distal category such as 1) P152A, 2) P152H and 3) P152R and two mutants typical for the proximal category such as T158M and R111G. Chromocenter fluorescence recovery after photobleaching (FRAP) was used to assess the binding dynamics of MeCP2 with the following missense mutations: P152A, P152H, P152R, T158M and R111G, in comparison to WT (Supplementary Videos). Figure 4A,B shows pre-bleach, bleached and post-bleached recovery images. R111G, T158M and P152R constructs showed a significantly more rapid rate of recovery (two tailed Mann Whitney non parametric U test; $p \leq 0.05$). Almost complete recovery was observed before the 60th frame, whereas for WT the recovery rate was slow and incomplete by the 160th frame. Values of recovery rates for P152A and P152H were not significantly different to WT. Slower recovery suggests tighter binding with heterochromatin, whereas faster recovery suggests looser binding, e.g. for T158M, R111G and P152R. However, concentration of available MeCP2 protein within the nucleus could also be influencing these rates. A comparative analysis of the averaged values of quantitative parameters t-half and mobile fraction is shown (Fig. 4D,E).

To evaluate statistically how well the model fits the data, R^2 values of each FRAP curve were calculated. Averaged ($n = 5$, independent FRAP experiments) value of the square of the correlation (R^2) between the response values and the predicted response values (goodness-of-curve-fit) were $R^2 = 0.992$ for WT, 0.99 for P152A, 0.976 for P152H, 0.8 for P152R, 0.948 for T158M, 0.674 for R111G and 0.982 for PP390-91del constructs. Averaged t-half recovery was $\sim 64.16 \pm 23.11$ sec, $\sim 52.61 \pm 2.27$ sec, $\sim 42.35 \pm 23.79$ sec, $\sim 12.44 \pm 3.86$ sec, $\sim 12.74 \pm 16.76$ sec, 15.85 ± 8.1 sec and 50.86 ± 22.69 for MeCP2 WT, P152A, P152H, P152R, T158M, R111G and PP390-91del constructs, respectively (Fig. 4D). The t-half for P152R, T158M and R111G was much lower than that for WT. Similarly, the numeric value of the mobile fraction (normalized to one) was $\sim 0.59 \pm 0.07$, $\sim 0.61 \pm 0.15$, $\sim 0.56 \pm 0.16$, $\sim 0.69 \pm 0.13$, $\sim 0.79 \pm 0.12$, 0.92 ± 0.09 and 0.69 ± 0.23 for MeCP2 WT, P152A, P152H, P152R, T158M, R111G and PP390-91del constructs, respectively (Fig. 4E).

Table 1 shows phenotypes in several very rare cases of hemizygous male patients with MeCP2 mutations. Clinical presentation such as condition at birth, age of reported first abnormality, history of development, head circumference, muscular abnormalities, motor skills, epileptic seizures, language and social skills, age and cause of death (if deceased) were compared and reported here. Interestingly, a wide range of phenotypic/clinical severity, ranging from neonatal encephalopathy to psychiatric disorders were observed. To determine the severity of clinical manifestations of the study participants, we used the clinical severity score (CSS), developed previously¹⁹. CSS were calculated and plotted in Fig. 5A. Where the same mutation was reported for multiple participants, the average CSS across these participants was plotted. Parameters such as chromocenter number (CN) and size (CS) were used where possible (i.e. where the mutation did not completely abolish chromocenter clustering, as seen for T158M, or binding as seen for R111G and R133L) to calculate Functional Severity Score (FSS) using the equation 1 below and were correlated with CSS (Fig. 5B,C).

$$FSS = |CN^{MUT} - CN^{WT}| |CS^{MUT} - CS^{WT}| \quad (1)$$

Discussion

We used C2C12 cells, which express relatively low levels of endogenous MeCP2 and have prominent heterochromatin/chromocenters suitable for qualitative and quantitative analysis of DNA binding proteins, with the aim of studying the effect of common and rare Rett MBD missense substitutions on the binding and clustering of the chromatin. The alternative approach involving studying the binding in patient-derived cells for each mutation is possible, but impractical. Using transfection of MeCP2_GFP fusion constructs presented us with the opportunity to expand the study to include a much larger number of mutations than in similar previous studies. It should be noted that the MeCP2-GFP fusion itself is unlikely to have different properties compared to native MeCP2, since

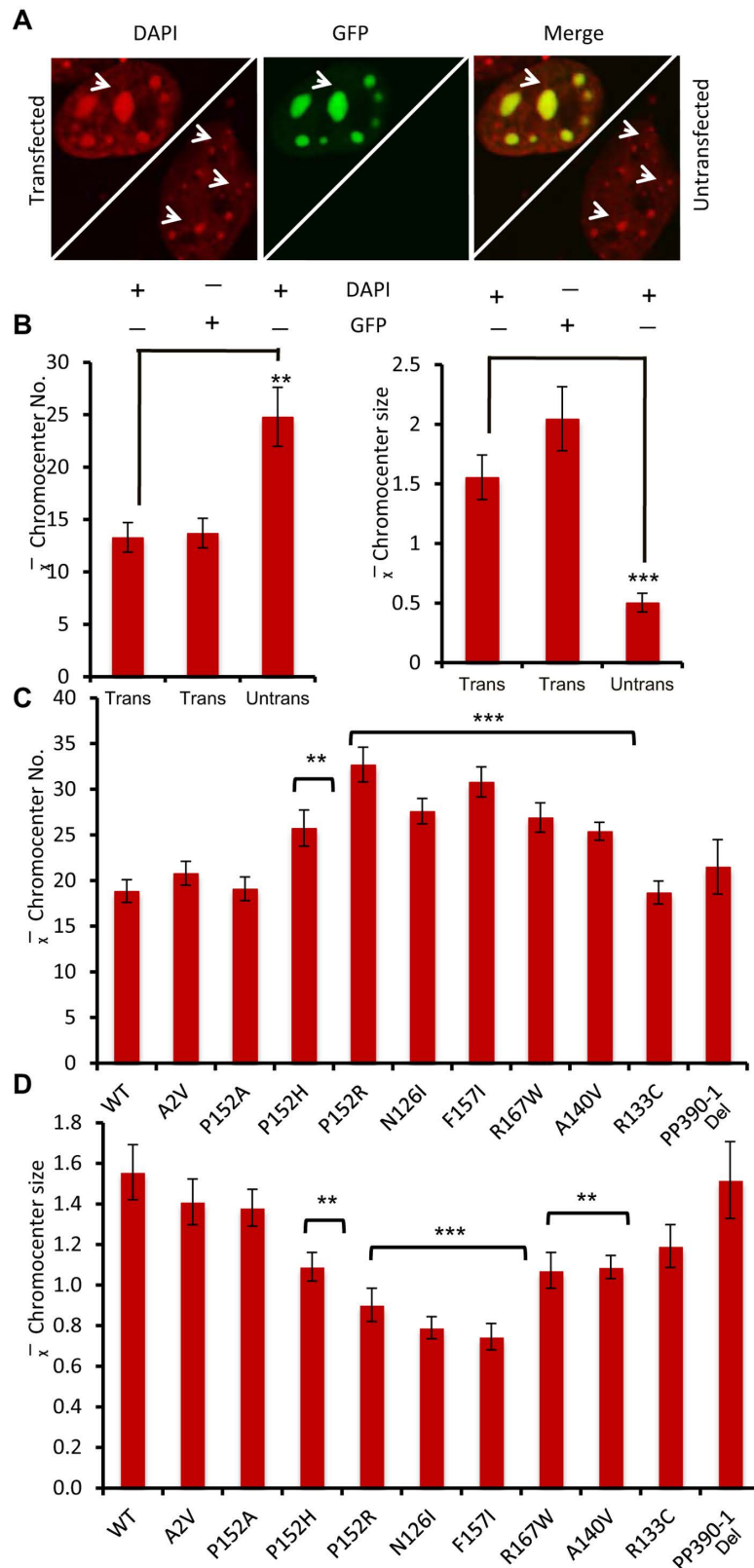


Figure 3. Quantitative analysis of chromocenter clustering in the transiently expressed mouse myoblast cells with mutant and wild type MeCP2 protein. (A) Comparison of chromocenter clustering in the transiently expressing wild type MeCP2 C2C12 cell (left corner) with neighbouring untransfected cell (right corner). (B) Quantitative analysis of chromocenter number and size using DAPI and GFP channels ($n = 10$); (C,D) Quantitative analysis of average chromocenter number from C2C12 cells transfected with wild type and each mutant recombinant MeCP2 protein ($n = 20$). Two tailed Mann Whitney non parametric U test. $+/-$ S.E. bars are shown. $***P \leq 0.005$ and $**P \leq 0.05$ and $+/-$ SEM is shown.

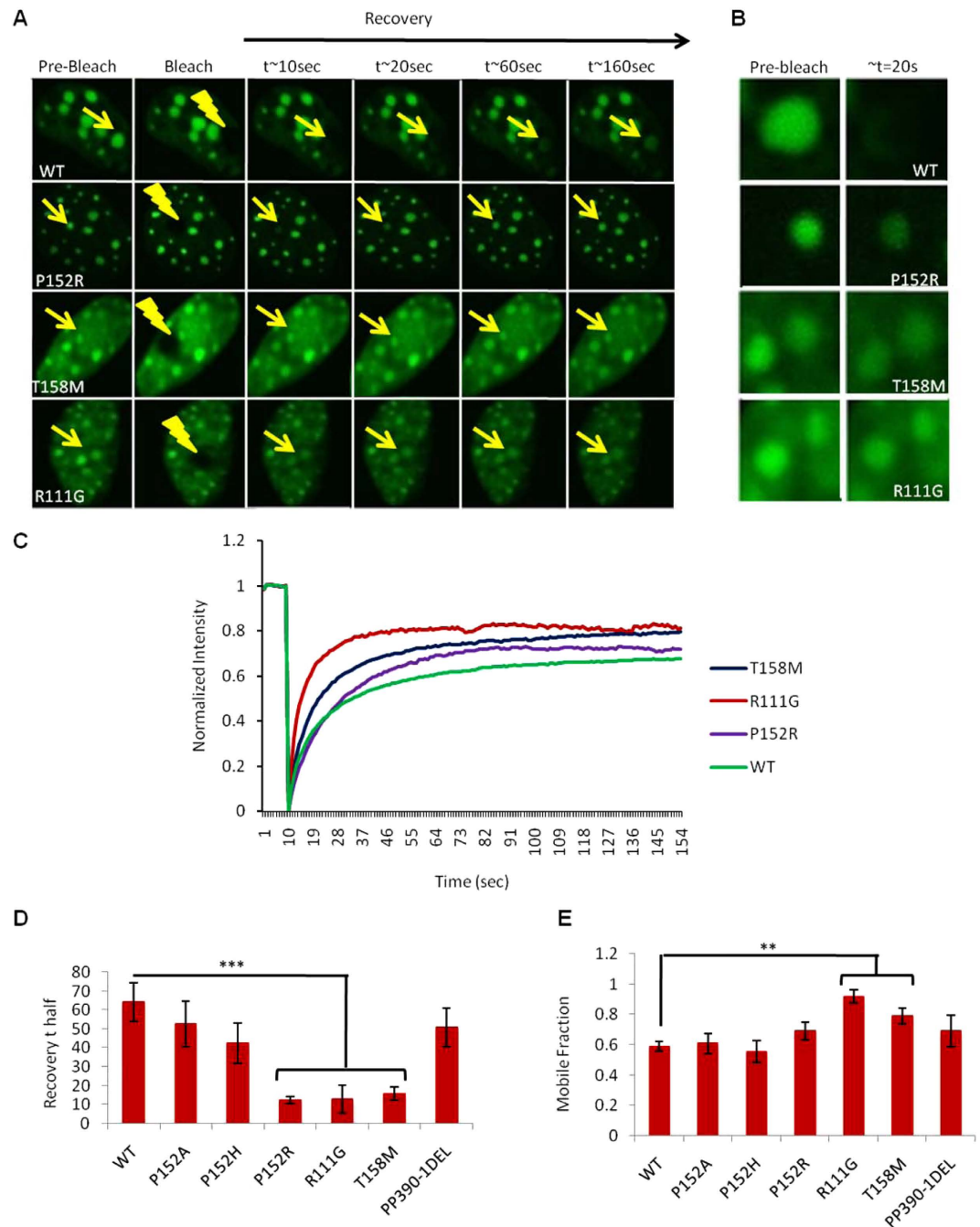


Figure 4. Chromocenter FRAP experiments to mobility and binding dynamics of wild type and mutant MeCP2. (A) Real-time post-bleach recovery of GFP tagged recombinant MeCP2 protein were captured after bleaching chromocenter at $\sim 200 \mu\text{sec}$. Yellow lightning indicate bleach spots. Rows (Top \rightarrow bottom): Representation of FRAP pre-bleach, bleach and post bleach real time recovery of MeCP2 wild type, P152R, T158M and R111G, respectively. Columns (Left \rightarrow Right): column 1 is representing preceding pre-bleach images; column 2 representing 1st images; column 3 representing 11th image; column 4 representing 21st images, column 5 representing 61th image and column 6 representing 161th image following bleach. where, 1 frame = 108.784 m.sec. (B) Chromocenter recovery at $\sim t = 20$ indicating different levels of recovery among wild type and mutant MeCP2; (C) FRAP recovery curves normalized to 1, showing chromocenter recovery in 151 frames. (D,E) Comparative illustration of half maximal recovery time and mobile fraction of wild type, P152A, P152H, P152R, T158M, R111G and PP390-91del, respectively. $n = 5$, $***p \leq 0.05$ two tailed Mann Whitney non parametric U test. \pm S.E. bars are shown.

transgenic mice expressing MeCP2-GFP are viable and almost undistinguishable from the wild type mice²⁰. In particular, we were interested in studying how the positional difference of each mutant in the tertiary structure of the MBD of MeCP2 affects its binding to the heterochromatin and mobility in the nuclear environment.

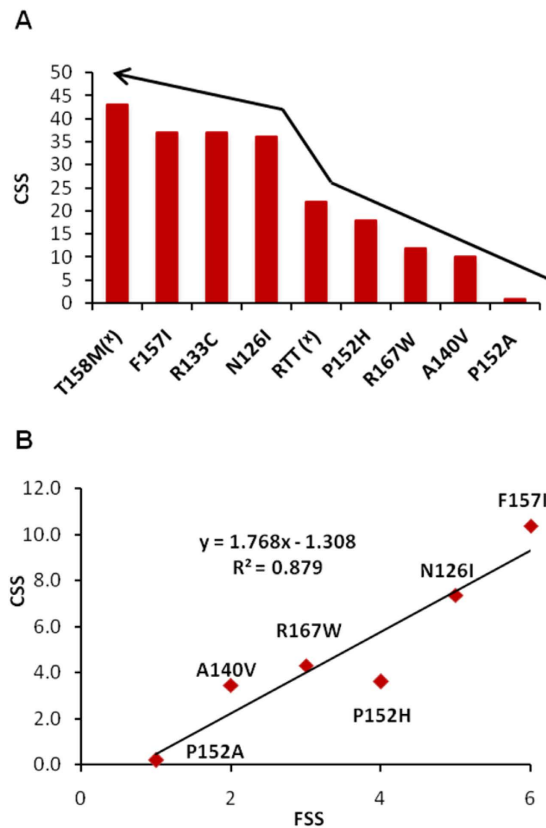


Figure 5. (A) Clinical severity scores (CSS) of F157I, N126I, P152H, P152A, † median CSS value of Rett girls and T158M cases (Case I, II, III, IV in Table 1). (B) Scatter plot representing correlation of functional Severity Score (FSS) with clinical severity score (CSS) of MeCP2 among male cases with missense mutations, as in Fig. 6A.

In the quantitative co-localization, the high degree of mislocalization of R111G, R106W and R133L MeCP2 suggest that MeCP2 is not binding at the chromocenters (Fig. 2B,C). Complete loss of the specific binding of MeCP2 to methylated DNA has already been reported for the R111G substitution²¹. R106W is present on the β -sheet site directing towards the DNA binding residue R111, which may disrupt the conformation of the MeCP2 DNA-binding pocket. For T158M, relatively higher PCC, indicates either a partial loss of binding or a complete loss of chromatin clustering (Fig. 2A,C). Overall, we can characterize these changes in the chromatin binding and clustering into three different patterns among mutants, 1) impaired clustering; 2) abolished clustering and 3) abolished binding (Fig. 2B).

Cells expressing N126I, A140V, P152R, P152H, F157I and R167W MeCP2 showed a significant increase in the number of chromocenters. Agarwal *et al.*²² previously reported severely affected heterochromatin clustering in proline substitutions P101R, P101H and P152R. Similarly, Adegbola *et al.*¹⁸ reported disrupted MeCP2-heterochromatin binding for P152A, P152R and T158M. Here, we confirm these observations and further report two novel substitutions, P152H and N126I, which also affect the clustering of heterochromatin. In the case of P152A, chromocenters were normal in number and size (Fig. 3C,D), and with relatively normal clustering of chromocenters, showing that some substitutions at P152 may have relatively minor consequences for DNA-clustering or compaction. Analysis of chromocenter numbers across all the mutants and WT showed a roughly two-fold range of variability (Fig. 3C). Interestingly, all the mutants that show a clustering defect (i.e. N126I, A140V, P152R, P152H and R167W) are located within the distal part of the MBD (Fig. 1A). This indicates the importance of the role of MBD, not only in binding to DNA, but also in the clustering of chromatin. It is not yet known whether MeCP2 performs this chromatin clustering and organizational function alone, or whether it is also dependent on interactions with its protein partners. However, it has been suggested previously that the MBD mutation A140V, which is present in the same distal pocket as P152R (Supplementary Fig. S1), loses the ability to interact with ATP-dependent helicase, ATRX, which is an important chromatin remodeling protein²³. These findings underpin the important architectural role of MeCP2 in the organization of chromatin.

There were no significant differences in the chromocenter size and number for the A2V (N-terminal) and PP390-91del (C-terminal) mutants compared with wild type MeCP2, suggesting that the patho-etiology in Rett patients with these mutants is independent of the MBD (Fig. 3C,D). For A2V, *in silico* analysis using TerminiNator (<http://www.isv.cnrs-gif.fr/terminator3/>) predicts decreased N-terminal methionine excision and N-acetylation and decreased stability/longevity of the protein²⁴, which most likely plays an important role in MeCP2 homeostasis. For PP390-91del there may be an effect on the binding of MeCP2 to other proteins due to deletion in the proline motif at the C-terminal domain (CTD)²⁵.

MeCP2 mutants such as R111G, T158M and P152R showed significantly higher recovery rates than wild type MeCP2, suggesting altered diffusion rate and mobility dynamics of the mutant protein. Previously, Kumar *et al.*²⁶ described the influence of different MeCP2 domain and RTT mutations on the DNA binding and clustering of chromatin using the same photobleaching approach. Here we further extended the study by using a site specific approach to describe how the binding and mobility dynamics of MeCP2 vary due to chemical and physical properties of the substituting amino acid, and how these differences relate to the clinical phenotypes in hemizygous males. R111G, T158M and P152R showed extremely fast recovery times ($p \leq 0.05$). Also, mobile fraction of R111G and T158M is significantly higher than wild type, whereas for P152R no significant amount of MeCP2 protein was mobile. This confirms that P152R does not affect the binding, instead it only affects MeCP2 mobility within the cellular environment. In addition, it is clear that different substitutions at P152 have different effects at the molecular level. Even though the recovery rates ($t_{1/2}$) for P152H and P152A are not significantly different, they were somewhat faster than wild type, which may go some way to explain the milder phenotypes associated with these mutations. Interestingly, for the first time a clear correlation can be seen with average $t_{1/2}$ recovery time in comparison to the clinical severity score (CSS) for the respective patients, with increasing severity: WT < P152A < P152H < P152R < T158M < R111G (Table 1; Figs 4D,E and 5A) and overall the clinical severity correlates highly with our functional/molecular tests, i.e. from mildest to severest effect: MeCP2 WT > P152A > A140V > R167W > P152H > median CSS for RTT girls > N126I > R133C > F157I > T158M (Fig. 5A,B).

To help explain the effects of the substitutions at various specific positions in the MBD tertiary structure (Fig. 1A and Supplementary Fig. S1), we compared the physical and chemical properties of the WT versus substituted amino acids (Supplementary Table S1) (www.ncbi.nlm.nih.gov/Class/Structure/aa/aa_explorer.cgi). For instance, the difference in the DNA binding at R111 and at R133 can be explained by the different chemical properties of the substituting amino acids. Arginine is one of the most chemically complex amino acids due to its enhanced ability to mediate molecular interactions. Guanidine is the functional group on the side chain of arginine, which makes arginine the most basic of the amino acids with a virtually permanent positive charge, which is crucial for binding with the negatively charged DNA backbone, and has a higher hydrogen bonding forming potential compared to other basic amino acid residues²⁷. For R111G, R133L and R133C, both glycine and leucine have non-polar hydrophobic side chains, but cysteine, on the other hand, is slightly polar and hydrophilic. Thus, the effect on MeCP2-DNA interaction is greater when the substituting amino acid is glycine and leucine, but milder when substituted with cysteine (Fig. 6A and Supplementary Table S1). This may also explain the milder phenotype of the R133C male patient in comparison with other mutants directly affecting the MeCP2-DNA binding site (Table 1), and an overall less severe phenotype reported for typical and atypical Rett girls with the R133C MeCP2 mutation^{28,29}. In addition to different chemical properties of the substitution, the effect may also be structural. For example, R106W is present on the β -sheet, directing towards the DNA binding residue R111. The substitution of arginine 106 with tryptophan may affect the MBD folding, as the non-polar hydrophobic properties of tryptophan alters both structure and charge, and thus adversely affect MeCP2-DNA binding (Figs 1A and 6B and Supplementary Table S1).

T158M is present on a region of the MBD potentially involved in both DNA binding and clustering of chromatin (Fig. 1B). The side chain of threonine is polar and hydrophilic and contains a hydroxyl group, whereas the methionine has a non-polar and hydrophobic side chain, and includes a S-methyl thioether side chain (Fig. 6E and Supplementary Table S1). Despite T158M MeCP2 previously reported to show close to normal electrophoretic band-shift with methylated DNA³⁰, the faster recovery time and high mobile fraction observed here and previously²⁶ suggests reduced binding, and in addition we show here reduced co-localization with chromocenters (Fig. 2B). The substitution of a chemically opposite amino acid most likely affects a MBD site that is important for chromatin clustering as well as DNA-binding (Fig. 6A,B and Supplementary Table S1). Interestingly, at the adjacent residue, F157I has a severe effect on chromocenter clustering (Fig. 2B). This may emphasize the effect of an aromatic side chain (phenylalanine) versus an aliphatic side chain (isoleucine) at this residue in the MBD³¹. Furthermore, this substitution may also affect a possible pi-pi interaction of phenylalanine-phenylalanine (Phe-Phe stack of aromatic rings) with the nearby residue F155 (Fig. 6C and Supplementary Table S1)³⁶. R167W which is a little further down from T158 and located in the intervening domain (ID) also appears to show a significant clustering defect, but the clinical phenotype is milder than T158M and F157I.

Among other mutants in the same pocket of the MBD such as P152R, P152H and P152A, it should be noted that the side chain of arginine and histidine both contain a full positive charge, whereas both proline and alanine are small, non-polar and hydrophobic (Fig. 6D and Supplementary Table S1). Furthermore, proline has an exceptional conformational rigidity compared to other amino acids which may help MeCP2 to maintain its structural integrity, whereas the substituted amino acids such as arginine fail to maintain the required structural arrangement (Figs 1B and 6D and Supplementary Table S1). The properties shared between proline and alanine may explain the relatively mild effect on the mobility dynamics for MeCP2 P152A (Fig. 4B) and on clinical severity (Fig. 5). Another example of a relatively mild phenotype at a structurally important residue is the A140V mutation, which is present within the middle of the alpha helix close to R133. Alanine, which has a non-polar, aliphatic methyl side chain, has very close properties in comparison with valine, except that valine has a slightly bulkier side chain due to a propyl group attached to C2. On another distal part of the MBD, N126I showed a similar clustering defect, likely due to the substitution of the polar and hydrophilic asparagine with a non-polar hydrophobic isoleucine (Figs 1B and 6F and Supplementary Table S1). Together, this may highlight the importance of exposed hydrophilic residues on distal parts of MBD in chromatin clustering, which are either required to maintain the structure of MeCP2 or involved in recruiting other proteins.

Evaluation of newly identified MeCP2 MBD missense variants through the interpretation of the chromatin binding and clustering effects, as reported here, may lead to improved and individualized diagnostics. Personalized therapeutic interventions may be developed based on our observations here, by selecting

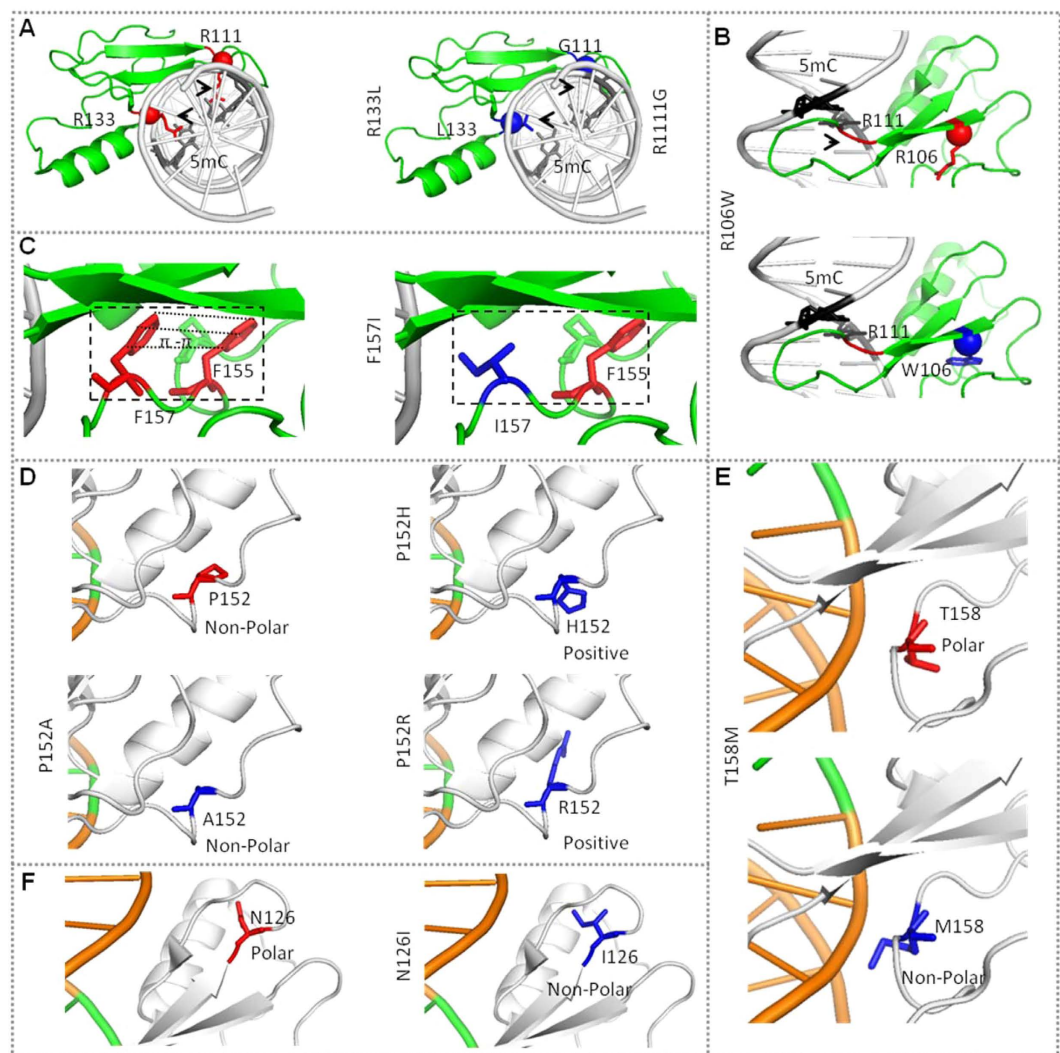


Figure 6. *In Silico* mutagenesis analysis of MeCP2 MBD model bound at BDNF promoter¹¹ using PyMol Molecular Graphic System (PDB accession code: 3C2I), showing structural and biochemical properties of the wild type and substituted residue (A) Disruption of arginine–DNA binding site at R111G and R133L (B) R106W and its potential effect on R111–DNA binding site; (C) F157I and its potential effect on the F155–F157 interaction due to absence of the aromatic ring of F157; (D–F) illustration of the effect on the structural and biochemical properties of the substitutions, including (D) P152A, P152H and P152R; (E) T158M and (F) N126I.

compounds on the basis of exclusive binding to the mutant versus wild-type MECP2 protein, then screening for improvements to the binding/chromocenter organization dynamics.

Methods

Experimental design.

The three-dimensional structure of MeCP2 MBD shows missense variants at: 1) the proximal part of MBD, which binds directly with the methylated DNA; and 2) the distal part of MBD, which appears not to be involved directly in DNA binding (Fig. 1A).

In this study we selected twelve MeCP2 missense mutations: R106W, R111G, N126I, R133C, R133L, A140V, P152A, P152H, P152R, F157I, T158M and R167W. Out of these twelve, there are two “hotspot” residues (R133, R111) that may interact with the DNA directly, and are present on or close to the MBD–methylated DNA interface, and residues (N126, P152) that are present on the distal part of the MBD of MeCP2 (Fig. 1A). R106W, F157I and R167W were selected due to their interesting position within the 3D structure. R106W, is present in the middle of a β -sheet, F157I is adjacent to T158M (Fig. 1A), whereas R167W is present a few residues away from T158M (Supplementary Figure S1); all three maybe crucial for DNA–MeCP2 binding and/or clustering of DNA (Fig. 1A and Supplementary Figure S1). To show that the binding and clustering effect is entirely due to the missense mutations in MBD of MeCP2, we generated two constructs, with a C-terminal deletion (PP390–391del) and an N-terminal mutation (MeCP2_e1: A2V), both with the normal MBD- as controls in addition to the wild type (WT) MeCP2. MECP2 is a major organizer of chromatin structure, and heterochromatin in particular, and helps

to re-organize DNA into chromocenters³². For all co-localization experiments, after fixation, cells were stained using DAPI which binds AT base pairs at the center of the DNA duplex³³. AT base pairs are present across the genome, but are enriched in heterochromatin compared to euchromatin. In order to assay the effects of MECP2 mutation on binding to and clustering of chromocenter DNA, we need to identify chromocenters. For this we have used co-localization of DAPI and MECP2-GFP.

Cell culture, vector system and mutagenesis. The mouse skeletal muscle cell line (myoblast) C2C12 (ATCC, Manassas, VA) was used as the cell system for this study due to the relatively prominent chromocenters and low endogenous MeCP2 expression²². In order to express MeCP2 in a mammalian system, the full length *MECP2_E1* gene was amplified using reverse transcriptase polymerase chain reaction followed by cloning into the expression construct pcDNA3.1 CT-GFP-TOPO (placing GFP at the C-terminal end of MeCP2), according to the manufacturer's instructions (Life Technologies, Carlsbad, CA). Mutations were introduced through PCR based site-directed mutagenesis according to the manufacturer's instructions (QuikChange Lightning Site-Directed Mutagenesis Kit, Agilent Technologies, Santa Clara, CA), in order to generate the MeCP2 missense substitutions R106W, R111G, N126L, R133L, R133C, A140V, P152A, P152R, P152H, F157I, T158M, also an N-terminal missense substitution A2V, and a C-terminal deletion P390_P391del, for use as positive controls, in addition to the wild type (WT) MeCP2.

Fluorescence confocal microscopy. Wild type and mutant MeCP2_E1-GFP fusion protein was expressed in C2C12 cells (mouse myoblasts; ATCC, Manassas, VA) after cloning in the pcDNA3.1 CT-GFP-TOPO vector (Life Technologies). Cells were transfected with a polyethylenimine transfection reagent (jetPEI, PolyPlus, Illkirch, France) following the manufacturer's instructions. Cells were fixed using 4% paraformaldehyde followed by 2X washes of ice-cold phosphate buffer saline (PBS), and the transfected cells were then stained with the AT-selective DNA dye DAPI (4'-6'-diamidino-2-phenylindol). Fixed-cell confocal microscopy was used to study MeCP2 accumulation at chromocenters. Z-stacks were acquired with a frame size of 1024 × 1024 pixels of MeCP2-GFP expressing fixed C2C12 cells.

Identification of single cell nuclei and chromocenters. Identification of single cell nuclei and chromocenters was performed by intensity-based thresholding and by implementation of the Water algorithm³⁴. GFP fusion protein and heterochromatin were captured by confocal microscopy (Olympus FV1200) using 63×/1.3 NA oil objective at 405 nm diode pumped solid state for DAPI, and 488 nm Argon for GFP.

Fluorescence recovery after photobleaching. To study protein binding dynamics in live cells, C2C12 cells were cultured in chambered cover glass culture plates (Nalge Nunc International, Rochester, NY). Normal and mutant MeCP2_E1-GFP proteins were transfected and expressed as described above. Fluorescence recovery after photobleaching (FRAP) was used to capture binding and diffusion dynamics. Live cell images were captured on an Olympus FV1200 confocal microscope using 63x/1.3 NA oil objective at 488 nm argon for GFP. Ten independent FRAP experiments in the temperature (37 °C)- and CO₂ (5%) controlled incubation chambers were performed on the recombinant protein for each of the MeCP2 missense substitutions P152A, P152A, P152H, T158M, R111G and wild type MeCP2 in the C2C12 cells. FRAP imaging series of all the GFP fusion protein constructs were recorded as follow: 25 pre-bleach frames followed by 200 μs bleach time with 405 nm laser line at 100% transmission, and at least 200 post-bleach frames were recorded at equal time intervals using time lapse imaging. For confocal time-lapse imaging, a frame size of 512 × 512 pixels and 488 nm laser excitation with 0.05 transmission were used to capture GFP-tagged protein post bleach recovery frames.

Data analysis. The cell images were captured using OlympusFluoView software. Pearson's correlation coefficient (PCC) for the co-localization of chromocenters (DAPI) and recombinant protein (GFP) were calculated using FluoView software. The identified chromocenters' size and their number were calculated using Fiji Particle Analysis plugin³⁵. FRAP data was analyzed on an automated MATLAB based windows application, easyFRAP³⁷. The descriptive statistical data analysis was calculated using Microsoft Excel and two tailed Mann Whitney non-parametric U test for statistical significance were calculated using GraphPad online tool. The questionnaire used to calculate clinical severity score (CSS) can be found in Supplementary Material. Written, informed consent was provided for all human subjects, and approval for use of human subjects was given by the respective institutions' research ethics board (Centre for addiction and Mental Health (CAMH) and Civitan International Research Center). Where study participants were minors or not able to provide consent, consent was obtained from a parent or legal guardian.

All experiments were performed in accordance with under relevant ethics and laboratory safety guidelines, under approved experimental protocols (CAMH Research Ethics Board protocol #194/2005, and CAMH Biosafety license #2014-02).

References

- Laurvick, C. L. *et al.* Rett syndrome in Australia: a review of the epidemiology. *J. Pediatr.* **148**, 347–352 (2006).
- Dragich, J., Houwink-Manville, I. & Schanen, C. Rett syndrome: a surprising result of mutation in MECP2. *Hum. Mol. Genet.* **9**, 2365–2375 (2000).
- Neul, J. L. *et al.* Rett Search Consortium: Rett syndrome: revised diagnostic criteria and nomenclature. *Ann. Neurol.* **68**, 944–950 (2010).
- Amir, R. E. *et al.* Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.* **23**, 185–188 (1999).
- Lewis, J. D. *et al.* Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell.* **69**, 905–914 (1992).

6. Nan, X., Meehan, R. R. & Bird, A. Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Res.* **21**, 4886–4892 (1993).
7. Meehan, R. R., Lewis, J. D. & Bird, A. P. Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. *Nucleic Acids Res.* **20**, 5085–5092 (1992).
8. Ghosh, R. P., Horowitz-Scherer, R. A., Nikitina, T., Shlyakhtenko, L. S. & Woodcock, C. L. MeCP2 Binds Cooperatively to its Substrate and Competes with Histone H1 for Chromatin Binding Sites. *Mol. Cell Biol.* **30**, 4656–4670 (2010).
9. Thambirajah, A. A. *et al.* MeCP2 binds to nucleosome free (linker DNA) regions and to H3K9/H3K27 methylated nucleosomes in the brain. *Nucleic Acids Res.* **40**, 2884–2897 (2012).
10. Hite, K. C., Adams, V. H. & Hansen, J. C. Recent advances in MeCP2 structure and function. *Biochem. Cell Biol.* **87**, 219–227 (2009).
11. Ho, K. L. *et al.* MeCP2 binding to DNA depends upon hydration at methyl-CpG. *Mol. Cell.* **29**, 525–531 (2008).
12. Ham, A. L., Kumar, A., Deeter, R. & Schanen, N. C. Does genotype predict phenotype in Rett syndrome? *J. Child Neurol.* **20**, 768–778 (2005).
13. Cheval, H. *et al.* Postnatal inactivation reveals enhanced requirement for MeCP2 at distinct age windows. *Hum. Mol. Genet.* **21**, 3806–3814 (2012).
14. Christodoulou, J., Grimm, A., Maher, T. & Bennetts, B. RettBASE: The IRSA MECP2 variation database - A new mutation database in evolution. *Hum. Mutat.* **21**, 466–472 (2003).
15. Lundvall, M., Samuelsson, L. & Kyllerman, M. Male RTT phenotype in T158M and R294X MeCP2 mutations. *Neuropediatrics*, **37**, 296–301 (2006).
16. Villard, L. *et al.* Two affected boys in a Rett syndrome family: Clinical and molecular findings. *Neurology*, **55**, 1188–1193 (2000).
17. Lynch, S. A., Whatley, S. D., Ramesh, V. & Sinha, S. & Ravine, D. Sporadic case of fatal encephalopathy with neonatal onset associated with a T158M missense mutation in MECP2. *Arch. Dis. Child Fetal Neonatal.* **88**, 250–252 (2003).
18. Adegbola, A. A., Gonzales, M. L., Chess, A., LaSalle, J. M. & Cox, G. F. A novel hypomorphic MECP2 point mutation is associated with a neuropsychiatric phenotype. *Hum. Genet.* **124**, 615–623 (2009).
19. Neul, J. L. *et al.* Developmental delay in Rett syndrome: data from the natural history study. *J. Neurodev. Disord.* **6**, 20 (2014).
20. Schmid, R. S. *et al.* A methyl-CpG-binding protein 2-enhanced green fluorescent protein reporter mouse model provides a new tool for studying the neuronal basis of Rett syndrome. *Neuroreport*, **5**, 393–398 (2008).
21. Kudo, S. *et al.* Heterogeneity in residual function of MeCP2 carrying missense mutations in the methyl CpG binding domain. *J. Med. Genet.* **40**, 487–493 (2003).
22. Agarwal, N., Becker, A., Jost, L. K., Haase, S. & Thakur, B. K. MeCP2 Rett mutations affect large scale chromatin organization. *Hum. Mol. Genet.* **20**, 4187–4195 (2011).
23. Nan, X. *et al.* Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation. *Proc. Natl. Acad. Sci. USA* **104**, 2709–2714 (2007).
24. Martinez, A. *et al.* Extent of N-terminal modifications in cytosolic proteins from eukaryotes. *Proteomics*, **8**, 2809–2831 (2008).
25. Buschdorf, J. P. & Strätling, W. H. A WW domain binding region in methyl-CpG-binding protein MeCP2: impact on Rett syndrome. *J. Mol. Med.* **82**, 135–143 (2004).
26. Kumar, A. *et al.* Analysis of protein domains and Rett syndrome mutations indicate that multiple regions influence chromatin-binding dynamics of the chromatin-associated protein MECP2 *in vivo*. *J. Cell Sci.* **121**, 1128–1137 (2008).
27. Ausio, J., Greulich, K. O., Haas, E. & Wachtel, E. Characterization of the fluorescence of the protamine thynnine and studies of binding to double-stranded DNA. *Biopolymers*, **23**, 2559–2571 (1984).
28. Leonard, H. *et al.* Patients with the R133C mutation: is their phenotype different from patients with Rett syndrome with other mutations? *J. Med. Genet.* **40**, 1–7 (2003).
29. Cuddapah, V. A. *et al.* Methyl-CpG-binding protein 2 (MECP2) mutation type is associated with disease severity in Rett syndrome. *J. Med. Genet.* **51**, 152–160 (2014).
30. Ghosh, R. P., Horowitz-Scherer, R. A., Nikitina, T., Gierasch, L. M. & Woodcock, C. L. Rett syndrome- causing mutations in human MECP2 result in diverse structural changes that impact folding and DNA interactions. *J. Biol. Chem.* **283**, 20523–20534 (2008).
31. Mason, P. E., Dempsey, C. E., Neilson, G. W., Kline, S. R. & Brady, J. W. Preferential interactions of guanidinium ions with aromatic groups over aliphatic groups. *J. Am. Chem. Soc.* **131**, 16689–16696 (2009).
32. Singleton, M. K. *et al.* MeCP2 is required for global heterochromatic and nucleolar changes during activity-dependent neuronal maturation. *Neurobiol. Dis.* **43**, 190–200 (2011).
33. Larsen, T. A., Goodsell, D. S., Cascio, D., Grzeskowiak, K. & Dickerson, R. E. The structure of DAPI bound to DNA. *J. Biomol. Struct. Dyn.* **7**, 477–91 (1989).
34. Harmon, B. & Sedat, J. Cell-by-Cell Dissection of Gene Expression and Chromosomal Interactions Reveals Consequences of Nuclear Reorganization. *PLoS Biol.* **3**, 450–462 (2005).
35. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods*, **9**, 676–682 (2012).
36. Hunter, C. A., Singh, J., Thornton, J. M. & Hunter, C. A. π - π interactions: the geometry and energetics of phenylalanine-phenylalanine interactions in proteins. *J. Mol. Biol.* **218**, 837–846 (1991).
37. Rapsomaniki, M. A. *et al.* easyFRAP: an interactive, easy-to-use tool for qualitative and quantitative analysis of Fluorescence Recovery After Photobleaching data. *Bioinformatics* **28**, 1800–1801 (2012).
38. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protocols*, **10**, 845–858 (2015).
39. Sheikh, T. I., Mittal, K., Willis, M. J. & Vincent, J. B. A synonymous change, p.Gly16Gly in MECP2 Exon 1, causes a cryptic splice event in a Rett syndrome patient. *Orphanet J. Rare Dis.* **8**, 1–7 (2013).
40. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using ClustalOmega. *Mol. Syst. Biol.* **7**, 1–6 (2011).
41. Orrico, A. *et al.* MECP2 mutation in male patients with non-specific X-linked mental retardation. *F. E. B. S. Letters.* **481**, 285–288 (2000).
42. Cohen, D., Lazar, G., Couvert, P., Desportes, V. & Lippe, D. MECP2 mutation in a boy with language disorder and schizophrenia. *Am. J. Psychiatry.* **159**, 148–149 (2002).
43. Klauck, S. M. *et al.* A mutation Hot Spot for Nonspecific X-linked Mental Retardation in the MECP2 \Gene Causes the PPM-X Syndrome. *Am. J. Hum. Genet.* **70**, 1034–1037 (2002).

Acknowledgements

We would like to thank all study participants and their families. Funding: This work was supported by a grant from the Ontario Rett Syndrome Association (to JBV and JA). TIS is supported by a Margaret and Howard Gamble Research Grant, and a Dalton Whitebread Scholarship. AKP and JBL are supported by NIH U54 grants RR019478 and HD061222, Office of Rare Disease Research, funds from the International Rett Syndrome Foundation and the Civitan International Research Center. The Rett Syndrome Natural History Study (U54 HD061222) is a part of the National Institutes of Health (NIH) Rare Disease Clinical Research Network (RDCRN), supported through collaboration between the NIH Office of Rare Diseases Research (ORDR) at the National Center for Advancing Translational Science (NCATS), and the Eunice Kennedy Shriver Child Health and Human Development Institute (NICHD).

Author Contributions

Study design by T.I.S. and J.B.V., with support from J.H.E. and J.A. Molecular studies and data analysis by T.I.S. Clinical studies and analyses by P.M., H.F., J.S., J.B.L. and A.K.P. Manuscript preparation by J.B.V. and T.I.S., with assistance from J.A. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The Centre for Addiction and Mental Health (JBV) holds a patent, and receives royalties, related to diagnostic screening of MECP2.

How to cite this article: Sheikh, T. I. *et al.* From Function to Phenotype: Impaired DNA Binding and Clustering Correlates with Clinical Severity in Males with Missense Mutations in *MECP2*. *Sci. Rep.* **6**, 38590; doi: 10.1038/srep38590 (2016).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016