TITLE

A multi-model approach defines function altering MECP2 missense variants identified in individuals with autism spectrum disorder.

AUTHOR LIST

Eric Chen[1]*, Jessica Schmitt[1]*, Graeme McIntosh[1]*, Barry P. Young[1], Tianshun Lian[1], Jie Liu[1], Kexin K. Chen[2], J. Beatrice Liston[1], Lily MacDonald[1], Bill Wang[1], Sonia Medina Giro[1], Benjamin Boehme[1], Mriga Das[1], Seevasant Indran[1], Jesse T. Chao[3], Sanja Rogic[2], Paul Pavlidis[2], Douglas W. Allan[1]#, and Christopher J.R. Loewen[1]#

[1] Department of Cellular and Physiological Sciences, The Life Sciences Institute, The University of British Columbia, 2350 Health Sciences Mall, Vancouver, BC, Canada, V6T 1Z3

[2] Department of Medical Biophysics, Sunnybrook Research Institute, The University of Toronto, 2075 Bayview Avenue, E240, Toronto, Ontario, Canada, M4N 3M5

[3] Department of Psychiatry and Michael Smith Labs, The Michael Smith Institute, The University of British Columbia, 301-2185 East Mall, Vancouver, BC, Canada, V6T 1Z4

(co-first *, co-corresponding #)

Running title: Altered MECP2 function in autism

ABSTRACT

Pathogenic variants in MECP2 commonly lead to Rett syndrome, where MECP2's function as a DNA cytosine methylation reader is believed critical. MECP2 variants are also catalogued in individuals with autism spectrum disorder (ASD), including nine missense variants which had no known clinical significance at the start of this study. To assess these nine variants as risk alleles for ASD, we developed MECP2 variant functional assays using budding yeast and *Drosophila*. We calibrated these assays with known pathogenic and benign variants. Our data predict that four ASD variants are loss of function and five are functional. Protein destabilization offers insight into the altered function of some of these variants. Notably, yeast and *Drosophila* lack DNA methylation, yet all Rett pathogenic and ASD variants located in the methyl DNA binding domain that we analyzed proved to be loss of function, suggesting a clinically-relevant role for non-methyl DNA-binding by MECP2.

INTRODUCTION

The X-linked protein MECP2 was first identified in screens for proteins recruited to methylated DNA [1], and loss of function (LoF) alleles of MECP2 were subsequently found as causative in the neurodevelopmental disorder, Rett syndrome [2]. Since then variants in MECP2 have been found in most cases of Rett, but also in other neurodevelopmental disorders, including autism spectrum disorder (ASD) [3–6]. The origin and phenotypic consequence of MECP2-damaging variants differ notably between sexes. In females, the vast majority of these variants are *de novo*, while in males, they can also be inherited from unaffected mothers [7]. Variants that typically result in Rett syndrome in females tend in males to cause severe neonatal encephalopathy, often leading to

44 death in the first year of life. However, in males with two populations of X chromosomes, such
45 as those with Klinefelter syndrome (47XXY) or somatic mosaicism, these variants result in a
46 milder Rett-like condition known as male RTT encephalopathy. Males can also be affected by
47 variants that are seemingly benign in females, which can lead to moderate to severe intellectual
48 disability [8].
49
50 MECP2 has largely been examined as a DNA cytosine methylation reader via its methyl cytosine
51 binding domain (MBD) that recruits co-repressor complexes to DNA to repress transcription via
52 its transcriptional repressor domain (TRD) [9–12]. The MBD, and to a lesser extent the TRD,
53 harbors most Rett LoF variants [13], resulting in variants which are believed to reduce binding to
54 methylated DNA or reduce recruitment of co-regulator complexes, respectively [14,15]. While
55 numerous MECP2 variants have been found in patients that exhibit Rett with ASD, recent
56 reports have annotated five MECP2 missense variants in ASD without an associated Rett
57 diagnosis (R91W, G195S, A202G, E282G, S411R) [16–20]. However, it is unclear whether any of
58 these variants are function altering and clinically significant. Two of them, R91W and A202G,
59 are classified by ClinVar [21] as variants of uncertain significance (VUS) in cases of severe
60 neonatal onset encephalopathy with microcephaly. E282G is classified by ClinVar as likely
61 pathogenic. The remaining two variants (G195S and S411R) have no current ClinVar
62 classification.
63
64 While MECP2 is understood to primarily be a methyl-cytosine DNA (DNAm) binding protein, it
65 has become clear it also binds unmethylated DNA with a preference for GC-rich sequences [22–25].
66 This has been proposed as an alternate hypothesis for why MECP2 appears to preferentially bind

67 methylated GC-rich regions of the genome, as opposed to binding DNAm methylation *per se*. It
68 can be challenging to segregate DNAm-dependent from DNAm-independent functions of
69 MeCP2 *in vivo*, as DNAm is both abundant and essential in mammals. Consequently, it is still
70 unclear what fraction of MeCP2's functions are dependent on its interaction with methylated
71 and/or unmethylated DNA. Such ambiguity leaves unresolved the question of whether disease
72 relevant variants in MECP2 selectively disrupt binding to methylated versus unmethylated DNA.
73

74 There are extremely low levels of methylated cytosine in the *Drosophila* genome, and none in
75 *Saccharomyces cerevisiae* [26,27]. Thus, fly and yeast offer avenues to explore the interactions of
76 MECP2 with unmethylated DNA, and to test if Rett and ASD variants alter MECP2 binding to
77 unmethylated DNA. Moreover, while neither organism has an ortholog of *MECP2*, both
78 organisms are robust genetic models that have proven well suited to comparing the function of
79 large numbers of gene variants [28–31]. MECP2 has been expressed in *Drosophila*, where it
80 associates with chromatin and mediates transcriptional repression via interactions with orthologs
81 of known MECP2 physical interactors in humans, including Sin3A, N-CoR, REST as well as
82 other chromatin modifiers [32]. In fly neurons, MECP2 impacts dendritic branching and
83 behavioural output [33,34] and C-terminal truncation mutants (R294X specifically) promote
84 neuronal apoptosis [35]. In yeast, MECP2 has recently been shown to bind chromatin and mediate
85 widespread transcriptional repression [36].

86 Our goal in this study was to use the *Drosophila* and budding yeast models to test the functional
87 impact of these ASD MECP2 variants to determine whether they may confer autism risk, by
88 virtue of their altered function. We calibrated our assays with MECP2 variants classified by
89 ClinVar as pathogenic or benign, in accordance with American College of Medical Genetics and

90 Genomics (ACMG) guidelines for experimental assays [37]; first to validate our assays for variant
91 assessment, and second to provide a rubric for determining if test ASD variants are likely
92 pathogenic or likely benign. All calibration variants functioned as expected in the *Drosophila*
93 assay, indicating that we could functionally assess any test variants. The yeast assay could be
94 calibrated appropriately for the MBD domain, but the MBD proved to be the only domain
95 showing the expected separation between benign and pathogenic calibration variants. Our
96 analyses indicate four of the nine test ASD variants were LoF, making these variants candidate
97 risk alleles for ASD. Interestingly, these variants mostly proved to be less severe LoF than Rett-
98 associated pathogenics, providing a possible explanation for how MECP2 mutations can lead to
99 ASD, but not Rett.

100

101 **METHODS**

102 **MECP2 isoform selection**

103 *MECP2* has two major isoforms, E1 and E2. We selected to work with *MECP2* isoform E2
104 (ENST00000303391, NM_004992) based on the recommendation by the ClinGen
105 Rett/Angelman-like Expert Panel [38]. This transcript is annotated as the APPRIS (Annotation of
106 Principal and Alternative Splice Isoforms [39]) principal isoform in Ensembl genome browser,
107 based on the annotations for protein structure, function and cross-species conservation, and is the
108 highest expressed transcript in the brain (GTEx portal; **gtexportal.org**).

109

110 **Variant aggregation, annotation and prioritization**

111 We utilized our in-house computational pipeline to aggregate *MECP2-E2* missense variants from
112 ClinVar and VariCarta and perform comprehensive annotation. We collected ClinVar variants

113 (downloaded in January 2022) that have at least one gold review star, i.e., the review status
114 (CLNREVSTAT) has to have one of the following designations: "criteria provided, single
115 submitter", "criteria provided, conflicting interpretations", "criteria provided, multiple
116 submitters, no conflicts" or "reviewed by expert panel". For each of these variants, we also
117 obtain its interpretation (clinical significance; CLNSIG) and reported condition (disease name;
118 CLNDN). The ASD missense variants were obtained from VariCarta, a database of variants
119 found in individuals diagnosed with ASD and reported in the scientific literature, that we created
120 [40]. We annotated the aggregated variants using the Ensembl VEP tool [41]. We also added protein
121 domain annotation relevant in the context of Rett Syndrome [42].
122

123 ***MECP2-E2* cloning**

124 A *MECP2-E2* cDNA clone was obtained (Horizon, Clone Id: 3956518, # MHS6278-
125 202830559), and *MECP2* coding sequences were amplified with iProof DNA polymerase (Bio-
126 Rad, California USA, #1725301) using the following primers:
127 GA-attL1-MECP2-522-F (tgtacaaaaaagcaggctccggcaccatggtagctgggatgttag)
128 GA-attL1-MECP2-522-R (ttgtacaagaaagctgggtctatcagctaactctctcggtc)
129 Amplicons were cloned in PCR8-GW-TOPO (Invitrogen, ThermoFisher Scientific, Waltham,
130 USA, catalog #K250020) using NEBuilder HiFi DNA Assembly Master Mix (New England
131 Biolabs, USA, # M5520). The resulting construct was named MECP2-pENTRY. A sequence of
132 CCGGCACC was introduced in primer GA-attL1-MECP2-522-F (underlined) preceding the
133 start codon for *MECP2-E2*. After LR cloning of MECP2-pEntry into pDEST vectors, this
134 CCGGCACC sequence can function as a Kozak sequence to initiate translation, or as an in-
135 frame linker for N-terminal peptide sequences. Missense variants were generated in MECP2-

136 pENTRY either using the Q5 mutagenesis kit (New England Biolabs, USA, # E0554) in house,
137 or were made by GenScript (New Jersey, USA). For MECP2 expression clone construction,
138 MECP2-pENTRY plasmids were subcloned in pAG426GAL-ccdB (a kind gift from a gift from
139 Susan Lindquist; Addgene plasmid # 14155) or pGW-HA.attB [43] using the Gateway LR Clonase
140 reaction (Thermo Fisher Scientific, Waltham, USA, catalog # 11791020). All transgenic *UAS-*
141 *MECP2* variants (on pGW-HA.attB) *Drosophila* strains used were verified by Sanger sequencing
142 (UBC NAPS, Vancouver, Canada) or whole amplicon sequencing (Plasmidsaurus, CA, USA)
143 (Supplemental Data 1).
144

145 **Sentinel interaction mapping (SIM) in yeast**

146 SIM screening for *MECP2* was performed as described previously [30]. A singer RoToR HDA
147 colony arraying robot was used to copy arrays of yeast cells. All steps were performed at a
148 density of 1,536 colonies per plate and plates were incubated for 24 hours at 30 °C prior to the
149 next pinning step, except where indicated otherwise. The SGA query yeast strain Y7092 [44] was
150 transformed with plasmid pSI7338-MECP2 and arrayed on synthetic defined medium lacking
151 uracil. The haploid yeast deletion mutant array (DMA) was arrayed across a total of four plates
152 on YPD medium containing 200 mg/L G418 sulfate (YPD+G418). The query array was mated to
153 the DMA by subsequently pinning both arrays on to fresh YPD plates. Diploids were selected by
154 pinning colonies to synthetic defined medium lacking uracil containing 200 mg/L of G418
155 sulfate and subsequently pinned to YPD+G418 plates. Each plate was then pinned in triplicate to
156 enriched sporulation media [44] and incubated at 22 °C for seven days. Following this, *MAT*a
157 haploid cells were germinated by pinning to synthetic complete medium lacking histidine,
158 arginine and lysine with the addition of 50 mg/L canavanine sulfate, 50 mg/L thialysine and

159 incubating for two days. To select for deletion strains carrying the MECP2-Ref plasmid, colonies
160 were pinned to two pairs of plates ("control" and "experimental") containing synthetic complete
161 medium lacking histidine, arginine, lysine and uracil, containing 50 mg/L canavnine sulfate, 50
162 mg/L thialysine, 200 mg/L G418 sulfate ("HURK+G418" medium). The experimental set was
163 supplemented with 10 nM -estradiol to induce expression of *MECP2*; this was omitted from the
164 control set. The final plates used for imaging were generated by copying each of these plates
165 once more to the same medium. Plates were scanned using a CanoScan 8600F flatbed scanner
166 and subsequently analyzed using *Balony* software [45] to compare the normalized colony size for
167 each strain in the presence or absence of estradiol. Genetic interactions were defined as those
168 strains where the ratio of experimental:control colony size was below a threshold determined by
169 *Balony* in all three replicates and with a p-value < 0.05. The strains *pho85*Δ*, rad50*Δ*, slx5*Δ*,*
170 *swi4*Δ*, swi6*Δ*,* and *xrs2*Δ were confirmed to be correct by PCR amplification of down-tag
171 barcodes with the standard primers B-U1 and B-D1 and sequencing.
172
173 **Yeast quantitative liquid growth assays**
174 Yeast strain JHY716 was transformed with various plasmids containing *MECP2* variants placed
175 downstream of the *GAL1* promoter. Four independent clones were picked from transformation
176 plates and transferred to SD-URA media containing 2% raffinose, 2% galactose to induce
177 expression of *MECP2*. For each clone, a 2 mL culture in SD-URA with 2% raffinose, 2%
178 galactose was grown overnight at 30 °C in a shaking incubator. The following morning, cultures
179 were diluted to an OD600 of ~0.25 in the same media. After 6 hours of incubation at 30 °C, log
180 phase cultures were diluted to an OD600 of 0.125. Three aliquots of 200 µl were taken from each
181 culture and transferred to a well of a flat-bottomed 96-well dish. Growth of cultures was

182 measured overnight at 30 °C with shaking at 800 rpm in an Agilent Logphase 600 plate reader.
183 All plates contained strains expressing wild type *MECP2* and the corresponding empty vector as
184 controls. The exponential growth constant k was determined for each well by fitting to an
185 exponential growth equation during the exponential phase (typically between 5 and 10 hours post
186 dilution).
187
188 **Protein abundance assays**
189 To determine the abundance of MECP2 variants in yeast, strains were grown in SD-URA media
190 containing 2% raffinose to early log phase. MECP2 expression was induced by the addition of
191 2% galactose to the growth medium. Cells were harvested three hours later by centrifugation and
192 stored at -80 °C for subsequent processing. Cell pellets were resuspended in Laemmli sample
193 buffer, and heated to 95 C for 5 minutes. Cell lysis was achieved by the addition of 50 µl acid-
194 washed glass beads followed by two rounds of disruption in a MP Biomedical FastPrep-24
195 homogenizer (maximum speed for 60 s with 5 minutes cooling between cycles). Samples (0.2
196 OD600 equivalents) were separated by SDS-PAGE on Bio-Rad 4-15% pre-cast gels and
197 transferred to nitrocellulose membranes. MECP2 was immunoblotted using Sigma antibody
198 (M9317), with PGK1 (Abcam AB113687) used as a loading control. For visualization,
199 secondary antibodies conjugated to Cy3 (MECP2) and AlexaFluor-488 (PGK1) were used and
200 imaged using a Bio-Rad ChemiDoc MP Imaging System. MECP2 band volume was normalized
201 to PGK1 band volume for each variant.
202
203 ***Drosophila* strains**

204 All flies were maintained on standard cornmeal food at 25 °C and 70% humidity unless
205 otherwise specified. *P{GAL4}A9* (BL8761) (referred to as *A9-GAL4*) was obtained from
206 Bloomington *Drosophila* Stock Centre (Indiana, USA). *UAS-MECP2* reference (*UAS-MECP2-*
207 *ref*) and variant sequences (collectively termed *UAS-MECP2.#*) were integrated into embryos of
208 genotype *y[1] w[] P{y[+7.7]=nos-phiC31\int.NLS}X; P{y[+t7.7]=CaryP}attP40* (referred to as
209 attP40) by Genome ProLab (Quebec, Canada), and maintained as *w[1118]; P{y[+t7.7]*
210 *w[+mC]=UAS-MECP2.#.GW.*}attP40***,** where # represents the variant and * represents
211 integrant line number (Supplemental Data 1).
212
213 **Crossing scheme**
214 Experimenters were double blinded to genotype throughout. Homozygous *UAS-MECP2* males
215 and a control *attP40* line were crossed to *A9-GAL4* virgins (ratio 4 males : 12 females), and
216 maintained at room temperature (~21 °C) for 72 hours. Crosses were then transferred into fresh
217 'assay' vials, and transferred each 24 hrs each for three days, generating three 'assay' vials per
218 cross. After removal of adults from each vial, vials were placed into a 29 °C incubator (70%
219 humidity). Three days after eclosion, adult flies were sorted by sex and placed in pure
220 isopropanol at -20 °C. All progeny flies for comparison were tested in a single batch.
221
222 **Wing imaging**
223 Experimenters were double blinded to genotype throughout. Flies were removed from
224 isopropanol, dried on paper towels, and 10-25 left wings were removed per genotype per assay
225 vial (one wing per fly). Wings were mounted on microscope slides in Canadian Balsam Fir
226 (1.01691.0025, CAS-No: 8007-47-4, Sigma Aldrich, MA, USA), coverslipped and sealed with

227 nail polish. Females have larger wings and the X-linked *A9-GAL4* driver is expressed at higher
228 levels in males, due to X-chromosome male-hyperactivation. Thus, it is important to analyze
229 males and females separately. While the impact of MECP2 variants in the wing was similar
230 between males and females, the phenotype was substantially stronger in males due to the higher
231 GAL4 levels; hence, higher MECP2 levels. To simplify analysis, we therefore only fully
232 analyzed male wings. Slides were compressed while drying for 24 hrs. Mounted wings were
233 scanned using a Panoramic Midi II digital slide scanner (3DHISTECH, Budapest, Hungary).
234 Scanned fly wing images were cropped at 15x zoom using CaseViewer (3DHISTECH,
235 Budapest,Hungary). Images were imported into Imaris v10.0 software (Bitplane, MA, USA) for
236 measurement of wing vein area per wing area, using Labkit machine learning and Fiji software
237 (University of Wisconsin-Madison, USA). In Labkit, we developed
238 "BIGSMOOTHCLASSYBOI", a machine learning algorithm to automate measurement of wing
239 vein area for calculation of wing vein area/wing area (details provided in Supplemental Methods
240 1).
241
242 **Statistical analysis of variant effects**
243 Yeast growth constants (k) and fly relative wing vein area data were analyzed using methods
244 similar to Post et al. [29]. Briefly, linear mixed-effects models were fit to each data set using the
245 *lmerTest* R package [46] to estimate the phenotypic effects of each variant compared to the
246 *MECP2-ref* constructs, treating experimental batches as random effects (for R scripts see
247 Supplemental Methods 2 and 3). For visualization, values were standardized to the range [0-1]
248 where 0 represents the mean phenotype measured for an empty vector, and 1 represents the mean
249 phenotype with the *MECP2-ref* construct. Four biological replicates, each with three technical

250 replicates were performed for yeast variant growth assays, while for *Drosophila* wing assays 10-
251 25 wings, each from a different animal were imaged per genotype.
252
253 **DATA AVAILABILITY STATEMENT**
254 Strains and plasmids are available upon request. The authors affirm that all data necessary for
255 confirming the conclusions of the article are present within the article, figures, tables and
256 supplemental data files.
257
258
259 **RESULTS**
260 ***Selection of variants***
261 We selected benign and pathogenic variants with clear clinical classifications to calibrate our
262 functional assays, and nine test (ASD) variants (Figure 1, Supplemental Data 1). We focused on
263 ClinVar variants with "pathogenic", "likely_pathogenic", "benign" and "likely_benign" clinical
264 interpretation as calibration variants. Priority was given to those with either "reviewed by expert
265 panel" or "criteria provided, multiple submitters" review status. A few exceptions were made.
266 D121E, a single-submitter benign variant, and T148A, originally classified as benign, but now
267 listed as conflicting were selected because there were no other benign variants in the MBD that
268 matched the aforementioned criteria. When selecting pathogenic variants, in addition to using
269 review status as a criterion, we also prioritized the ones with "Rett syndrome" condition
270 annotation as well as the ones that have been frequently observed in Rett patients [42]. In total, we
271 selected 37 calibration variants, 10 benign and 27 pathogenic. We selected ASD MECP2 variants
272 from VariCarta as test variants as long as they did not have a clear clinical interpretation in

273 ClinVar, for the total of 9 test (ASD) variants. Since the time of our initial selection, six test
274 (ASD) variants have been given updated classifications in ClinVar: P152L, P217L and E282G
275 becoming pathogenic/ likely pathogenic, and G195S, P376S and P388S becoming benign
276 (Figure 1 displays the initial classifications).
277
278 ***MECP2 produces a slow growth phenotype in yeast***
279 *MECP2-ref* was cloned into a high-copy 2-micron plasmid designed to give controlled
280 expression in response to addition of estradiol. Using an anti-MECP2 antibody we monitored
281 MECP2 protein levels by Western blot in response to increasing estradiol (Figure 2A). In the no-
282 estradiol condition we detected low levels of MECP2 protein, which increased with increasing
283 estradiol and then saturated at ~10 nM. Next, we measured the growth rate of WT yeast
284 containing either the *MECP2-ref* plasmid or an empty vector control in response to estradiol
285 addition (Figure 2B). Compared to empty vector which had little effect in response to increasing
286 estradiol, yeast containing MECP2 showed a dose-dependent decrease in exponential growth
287 rate. We compared this to another vector designed for strong overexpression utilizing the
288 *Gal1/10* promoter in a 2-micron vector and observed an even greater decrease in growth rate in
289 the presence of MECP2 (Figure 2C). Thus, expression of MECP2 reduced the fitness of WT
290 yeast in a dose-dependent manner.
291
292 ***MECP2 expression affects DNA-related processes in yeast***
293 To understand how MECP2 might be functioning in yeast and if this is related to the normal
294 function of MECP2 we employed a genetic interaction mapping approach we previously
295 developed to study the function of human genes in yeast, called Sentinel Interaction Mapping

296 (SIM) [30]. For SIM we express the human gene of interest in a collection of nearly 4,800 yeast
297 deletion mutants and systematically identify genetic interactions, which indicate functional
298 associations between the human gene and yeast cellular pathways. Hence, these genetic
299 interactions will inform us on the function of the human gene in yeast and if this is relevant to its
300 function in humans. We performed a SIM screen with *MECP2-ref* in which MECP2, under
301 control of the estradiol responsive promoter in the 2-micron vector, was expressed in ~ 4,800
302 different yeast deletion mutant strains in high-density arrays (see Methods for details). Colony
303 sizes were scored using Balony software and genetic interactions were identified by comparing
304 control arrays containing no estradiol (uninduced) to experimental arrays containing 10 nM
305 estradiol (induced). We chose 10 nM estradiol because MECP2 protein expression saturated at
306 this concentration and the effect on growth was not too severe, which would otherwise hinder
307 identification of genetic interactions.
308
309 The SIM screen identified 123 aggravating genetic interactions in 3/3 biological replicates
310 (Supplemental Data 2). We examined functional enrichment for these interactions using Gene
311 Ontology (GO) analysis (Figure 3A; also see Supplemental Data 3 for gene lists). This revealed a
312 network of interconnected functions related to DNA and chromosome stability, with the most
313 significant enrichment being for the terms "Cellular response to stress" (30 genes) and "DNA
314 damage response" (24 genes). Functions related to DNA binding (27 genes), cell cycle (21
315 genes), Golgi (8 genes) and nitrogen compound utilization (17 genes) were also enriched for.
316 The large enrichment for DNA-related functions was consistent with MECP2 being a DNA
317 binding protein with roles in transcriptional repression and chromatin organization. We focused
318 on a subset of interactions from GO terms most relevant to MECP2 function. Shown in Figure

319 3B is the effect on growth of the individual mutants from the screen upon MECP2 induction for
320 the GO terms “Chromatin organization”, “DNA binding” and “DNA damage response”. Because
321 these are aggravating interactions, MECP2 induction (10 nM estradiol) resulted in decreased
322 growth rate of the mutants relative to no induction (0 nM estradiol). Growth was normalized to
323 account for the effect of MECP2 induction on WT yeast growth.
324
325 Looking more closely at these genes, we identified five out of the ten members of the *RAD52*
326 epistasis group, *RAD50/51/52/54* and *XRS2* (*RAD55/57/59*, *RDH54* and *MRE11* were not
327 identified) (Figure 3B). Mutants in this group are defective in the repair of DNA damage,
328 maintenance of telomere length and meiotic and mitotic recombination [47]. We also identified
329 three out of four of the subunits of the highly conserved MRX endo/exonuclease complex, which
330 functions in repair of DNA double strand breaks, detection of damaged DNA, DNA damage
331 checkpoint activation, telomerase recruitment and suppression of chromosomal defects [48]. We
332 identified Rad50, Xrs2 and Sae2, but not *MRE11* in our screen (Figure 3B). *RAD50* and *XRS2*
333 were among the strongest interactions such that MECP2 expression reduced growth of the
334 *rad50*Δ and *xrs2*Δ mutants to almost undetectable levels.
335
336 We also identified both subunits of the SBF transcription factor complex, Swi4 and Swi6. Swi6
337 is also a component of the MBF complex (MBF comprises Swi6 and Mbp1), which along with
338 SBF functions to regulate transcription of the G1/S gene cluster as well as transcription of genes
339 involved in DNA synthesis and repair [49]. *swi6*Δ cells grow slowly, have a delayed G1/S
340 transition and are sensitive to DNA damaging agents. Expression of MECP2 in the *swi6*Δ mutant
341 severely reduced growth indicating a strong interaction. Additionally, we identified a strong

342 interaction with the cyclin-dependent kinase *PHO85*, which regulates many cellular processes in
343 response to stress and also plays an important role in G1/S transition via interaction with various
344 G1 cyclins [50]. Δ*pho85* cells grow slowly, have increased mitotic chromosome loss and shortened
345 telomeres. Taken together, these strong aggravating interactions with genes with central
346 functions in DNA synthesis and repair, chromatin organization and G1/S cell cycle progression
347 suggested MECP2 expression severely antagonized these processes which resulted in reduced
348 fitness.
349
350 ***Assessment of MECP2 variants in yeast***
351 We decided to simply exploit the robust slow growth phenotype observed upon MECP2
352 expression in WT yeast to determine the effects of MECP2 variants, anticipating variants that
353 interfered with MECP2 function would no longer reduce growth. We measured yeast growth in
354 liquid culture for all 46 variants using the galactose-inducible 2-micron vector, because this
355 vector produced the largest growth defect. We calculated exponential growth rates for WT yeast
356 expressing MECP2 variants and using a linear mixed-effects model estimated the effects of the
357 variants on MECP2 function (raw data is provided in Supplemental Data 4). Shown in Figure 4A
358 is the model's estimates of MECP2 function, arranged by variant amino acid position. Values
359 were standardized to the range [0-1] where 0 represents the mean phenotype measured for an
360 empty vector, and 1 represents the mean phenotype with the *MECP2-ref* construct. Hence,
361 variants with values near to 1 functioned similarly to MECP2-ref, whereas values greater or less
362 than 1 indicated increased or decreased function, respectively (error is plotted as 95% confidence
363 intervals).
364

365 Benign calibration variants had little effect on function as expected across all MECP2 domains.
366 The one exception was the D121E variant which was nearly complete LoF, for which we have
367 additional supporting evidence in fly leading to reclassification of this variant as pathogenic (see
368 below). For the MBD domain, all seven pathogenic calibration variants were complete or near-
369 complete LoF (we define a variant is complete LoF when its 95% confidence interval overlaps
370 with 0). This supported that our assay was disease-relevant for variants in the MBD. L100V,
371 R106W, R133C, and T158M have been previously reported to have reduced DNA binding *in*
372 *vitro* and *in vivo* [51], suggesting our assay reported on this function of the MBD. Residue R133 is
373 part of the DNA interface and interacts directly with guanine of the mCpG island, and T158 is
374 part of the Asx-ST motif that interacts with the DNA backbone [51]. The two test (ASD) variants in
375 this domain, R91W and P152L, were found to have decreased function, although not as severe as
376 the calibrating pathogenics. In contrast, calibrating pathogenics outside the MBD generally did
377 not show decreased function. This indicated the yeast assay could not reliably detect altered
378 function for variants outside the MBD.
379
380 ***Protein level of variants in yeast***
381 We measured protein levels for all variants in the MBD to determine if decreased stability/
382 increased turnover might be a factor in our assessment of function. We also included additional
383 non-MBD calibrating benigns from the TRD (T228S) and the CTD (S356G) for comparison; and
384 P388S because this variant showed the most reduced function of all non-MBD variants. MECP-2
385 protein levels from yeast extracts were measured on Western blots and normalized to MECP2-ref
386 (Figure 4B and Supplemental Figure 1). For the MBD, 8 out of 13 MBD variants had reduced
387 levels relative to MECP2-ref, while 5 were not significantly different from MECP2-ref. Variants

388 with decreased protein levels ranged from ~65% (T148A) to ~37% (R106W) of MECP2-ref. The
389 two calibrating benigns outside the MBD, T228S and S356G, did not show decreased protein
390 levels, whereas the test (ASD) variant P388S had substantially reduced protein levels (~ 25% of
391 MECP2-ref).
392
393 Comparing protein levels to function of variants there appeared to be little correlation. Some
394 variants had normal protein levels yet were non-functional (L100V, P101S, D121E, R133C).
395 Other variants showed both decreased protein levels and decreased function (R106W, T158M,
396 G163W, P388S). This suggested in some cases the variant affected MECP2 DNA binding
397 directly leading to decreased function (e.g., R133C), whereas in other cases decreased function
398 might also be a result of destabilization/ increased turnover (e.g., R106W). To investigate this in
399 more detail we looked to variants outside the MBD, because these should not directly affect
400 DNA binding. P388S was the only variant outside the MBD that showed moderate LoF (~ 50%;
401 Figure 3A) and we found it to have greatly reduced protein levels, whereas a variant in the same
402 domain, S356G, was functional and had normal protein levels. The modest decrease in function
403 of P388S relative to its low protein levels, suggested our functional assay was relatively
404 insensitive to destabilization. The T148A variant, which had reduced protein levels (~65% of
405 MECP2-ref), but was completely functional in the assay, further supported this conclusion.
406 Overall, these data suggested the yeast functional assay primarily reported on the intrinsic DNA
407 binding ability of the MECP2 MBD.
408
409 ***Using a fly wing vein assay for assessing MECP2 variant function***

410 Expression of *MECP2* in *Drosophila* tissues has a wide range of phenotypes [32–35]. From these
411 reports, we chose to assess MECP2 variants in developing wing tissues. This was because (1) the
412 reported ectopic wing vein phenotype could be exploited here to provide a quantitative readout
413 of MECP2 variant protein function, and (2) *MECP2* was shown to genetically interact with
414 known MECP2 co-repressor complexes in the context of this specific phenotype [32], suggesting
415 the wing vein phenotype was relevant to MECP2 function in humans. These provided strong
416 precedence for using this wing vein assay for assessing the function of ASD variants in MECP2.
417
418 We integrated *UAS-MECP2-ref* into the *attP40* locus on the *pGW.HA.attB* transgenic vector
419 backbone to control for genome position effect and copy number [43], to reduce variability in
420 expression levels between integrants. In pilot experiments we tested a range of conditions
421 including wing imaginal disc *GAL4* drivers, incubation temperatures and sex of progeny (as
422 females have larger wings). We found that the ideal set of variables to produce a robust ectopic
423 wing vein phenotype were the *A9-GAL4* driver in males (X-chromosome hyperactivation in male
424 flies results in higher GAL4/UAS activity) raised at 29 ºC (Figure 5, arrows). We also noted
425 *MECP2-ref* expression resulted in smaller wings (Figure 5). Thus, we could proceed with
426 assessing the effects of variants using this assay.
427
428 ***Assessment of 25 MECP2 variants using the wing vein assay***
429 In total 25 transgenic lines were generated, including eight calibrating benign, eight calibrating
430 pathogenic and nine test (ASD) variants. For all transgenic lines we passaged mating adults
431 through three consecutive 'assay' vials. Adult progeny were then collected 3 days post-eclosion
432 from each assay vial, and a single wing from a minimum of 10 adult male progeny per assay vial

433 was removed and slide-mounted, for a total of 30 wings per variant. Mounted wings were
434 scanned and total wing vein area / wing area was analyzed using a Labkit machine learning
435 algorithm that semi-automated the quantitation of wing vein area (see Methods).
436
437 Next, we examined wing vein phenotypes induced by expression of all 25 MECP2 variants, with
438 *attP40* and *MECP2-ref* controls in parallel. Representative images for controls, selected
439 calibrating benign and pathogenic variants and all test (ASD) variants are shown in Figure 5
440 (also see Supplemental Figure 2 for representative images of all MECP2 variants and the masked
441 area that was considered to contain vein tissue). Benign calibration variants all showed increased
442 ectopic wing veins (T203M and T228S are shown) whereas calibrating pathogenics showed very
443 little (L100V and T158M are shown), and test (ASD) variants had a range of outcomes (Figure
444 5). For example, P376S showed increased ectopic wing veins similar to *MECP2-ref* whereas
445 P152L had few extra wing veins.
446
447 To quantify this phenotype, we trained two machine learning algorithms to quantify (1) the total
448 area of the wing vein region and (2) the total area of the wing (see Supplemental Data 5 for
449 details). Then using a linear mixed effects model these two values were used to calculate the
450 estimated effect of variants relative to *MECP2-ref* and were normalized to *MECP2-ref* (at 1) and
451 *attp40* (at 0) (Figure 6, error bars are 95 % confidence intervals). All benign calibration variants
452 functioned similar to *MECP2-ref*, except for D121E. This variant was classified in ClinVar as
453 benign by a single contributor, but we find both yeast and fly assays characterized it as a strong
454 LoF, hence we predict D121E is likely pathogenic. All pathogenic calibration variants showed
455 varying degrees of LoF, indicating the wing vein assay reported on disease relevant functions for

456 MECP2. Most were strong LoF variants, except for P225R, which had reduced function but not a
457 complete LoF. This successful calibration of the assay indicated clinically pathogenic and benign
458 variants can be correctly classified in the fly model. We were particularly confident for variants
459 in the MBD and TRD domains, because we tested multiple calibrating pathogenics in these
460 domains, and both DNA and co-repressor interactions through these domains are validated in this
461 assay. In the same experimental batch as the calibrating benign and pathogenic variants
462 (controlling for food, temperature and humidity), we determined the wing vein phenotype for all
463 nine test (ASD) variants. We found that both MBD variants R91W and P152L were LoF. The
464 TRD variants P217L and E282G exhibited a partial LoF with data distributed similarly to the
465 TRD pathogenic variant P225R. Finally, the Interdomain (G195S, A202G) and CTD variants
466 (P376S, P388S and S411R) all exhibited normal function, similar to the calibrating benigns.
467 Thus, we conclude that only test (ASD) variants in the MBD and TRD domains were LoF.
468
469 Finally, we compared yeast and fly assay results for MBD variants only, because the yeast assay
470 could only reliably measure MBD variants. We observed remarkable similarity in results
471 between the two assays. Of particular interest was that both test (ASD) variants R91W and
472 P152L appeared more functional in both yeast and fly assays than the pathogenic calibration
473 variants. This suggested ASD variants may retain more DNA binding capacity than Rett variants,
474 given these assays likely reported on the intrinsic ability of MECP2 to bind DNA.
475
476 **DISCUSSION**
477 ***Clinical classification of variants***

478 Using our assays calibrated with variants classified in ClinVar as pathogenic and benign, we are
479 able to provide functional assessments for nine test (ASD) variants, five of which have been
480 found in cases of ASD without Rett (underlined): R91W, P152L, G195S, A202G, P217L,
481 E282G, P376S, P388S, S411R (Figure 1). We found that R91W and P152L were LoF variants
482 (yeast and fly) and we classify these as likely pathogenic. Our classification for P152L agrees
483 with the updated pathogenic ClinVar classification and we re-classify R91W from a VUS in
484 ClinVar. G195S had a modest gain of function (fly), hence we classify it as likely benign. Our
485 likely pathogenic classification for G195S agrees with the updated benign ClinVar classification.
486 A202G was found to be functional (fly). The ClinVar classification for this variant is conflicting
487 (one benign and one VUS report) and we propose this variant be classified as likely benign.
488 P217L and E282G showed moderately reduced function (fly) consistent with their updated likely
489 pathogenic ClinVar classifications. P376S, P388S and S411R were found to be functionally
490 normal (fly). For P376S and P388S this is consistent with their updated benign classifications in
491 ClinVar. However, since P388S appeared to be a destabilized protein in yeast we suggest P388S
492 is worthy of further examination. We classify S411R as likely benign providing the first
493 classification for this variant. Importantly, we provide functional evidence to classify two
494 variants found in ASD but not Rett as likely pathogenic (R91W, E282G), suggesting certain
495 mutations in MECP2 may cause ASD while others will cause Rett. Further, our data suggest
496 variants found in ASD but not Rett may be less severe than Rett causing mutations, since the
497 ASD-only variants we tested were more functional than Rett causing variants.
498
499 Finally, we provide evidence for reclassification of two variants, D121E and T148A, currently
500 classified in ClinVar as benign and conflicting, respectively. D121E was LoF in both yeast and

501 *Drosophila* assays, hence we propose D121E be reclassified as likely pathogenic. Interestingly,
502 our data are consistent with *in vitro* studies showing a ~7-fold reduction in the DNA binding
503 capacity of the D121E variant [52]. This variant is neither in VariCarta nor RettBASE and it has
504 one report in ClinVar, but this report lacks a description for how a benign classification was
505 made. Interestingly, the computational predictors SIFT, PolyPhen and CADD all predict this
506 variant is likely to be LoF (Supplemental Data 1), supporting our likely pathogenic classification.
507 For T148A, the current ClinVar classification is conflicting. We find this variant to be functional
508 (yeast), but to have modestly decreased protein levels (~65% of MECP2-ref), suggesting it may
509 have reduced stability. SIFT, PolyPhen and CADD all predict this variant to be damaging. We
510 propose T148A be reclassified as likely benign, but suggest further characterization is warranted.
511
512 ***Calibrating functional assays by protein domain***
513 Often proteins contain multiple domains which contribute distinct functions to the protein. Our
514 results indicate that when developing functional assays for human proteins in evolutionarily
515 distant model organisms it is important to calibrate assays for each domain. This can be achieved
516 through inclusion of calibration variants with known clinical/biochemical designations covering
517 each domain. MEPC2 is a multi-domain protein and is largely unstructured outside the MBD
518 domain [53]. We initially focused on identifying calibrating variants in the MBD and TRD
519 domains, as these domains harbour many of the clinically relevant variants.
520
521 The yeast assay only calibrated for the MBD and failed to calibrate for all other domains,
522 indicating we could only assess functional effects of variants found within the MBD. A likely
523 explanation is that the MECP2 MBD binds DNA in yeast without the need of accessory factors.

524 Indeed, Brown *et al.*, show this is likely to be the case [36]. The TRD mediates interactions with
525 co-repressor complexes, including Sin3A, N-Cor and REST, and Rett-causing mutations disrupt
526 these interactions [54]. It is likely the assay failed to calibrate for TRD variants because these
527 complexes are either not present in yeast (N-Cor and REST), or the TRD did not interact with the
528 yeast ortholog (Sin3A/HDAC). The results of the MECP2 SIM screen also provided valuable
529 insight into MECP2's function in yeast. The strong aggravating genetic interactions we observed
530 with genes involved in DNA damage response and DNA repair pathways (e.g., *RAD50*) suggest
531 MECP2 expression either directly or indirectly results in DNA damage. These effects of MECP2
532 likely disrupted the cell cycle (supported by the interactions with cell cycle genes, e.g., *PHO85*)
533 leading to overall decreased cell growth. Since variants in the MBD known to disrupt DNA
534 binding failed to inhibit growth, this further suggested DNA binding by MECP2 contributed to
535 defects in DNA and/or chromatin.
536
537 *Drosophila* was successful in stratifying pathogenic and benign calibration variants in both the
538 MBD and TRD domains, indicating we can make high confidence predictions about test ASD
539 variants located in these domains. For variants in the MBD, it is likely these disrupt DNA
540 binding, although DNA binding by MECP2 has yet to be demonstrated in *Drosophila*. For the
541 TRD, our data suggest MECP2 is able to interact with *Drosophila* orthologs of co-repressor
542 complexes. First, R306C prevents interaction with co-repressors and was complete LoF in our
543 assay [54]. Second, co-repressor complex mutants suppress phenotypes resulting from MECP2
544 over-expression in the *Drosophila* wing [32]. For example, mutation of the Polycomb group co-
545 repressor ortholog of Asxl1, *Asx*, prevents ectopic wing vein growth resulting from MECP2
546 expression [32]. For the Interdomain and CTD domain, testing additional pathogenic variants in

547 these domains would increase confidence in our predictions for test (ASD) variants within these
548 domains. Thus, when developing functional assays in surrogate model systems with considerable
549 evolutionary distance from humans, it is important to consider calibrating assays for each domain
550 before making conclusions about the effects of test variants.
551
552 ***MECP2 protein stability***
553 The stability of a handful of MECP2 variants has been studied both using *in vitro* assays and *in*
554 *vivo* mouse models [51,55–57]. We were able to compare seven variants located in the MBD between
555 our yeast results and these studies (Supplemental Table 1). R106W, P152R and T158M are all
556 destabilized in mouse models and *in vitro* systems, and based on protein levels in yeast were
557 predicted to be also destabilized in yeast also. R133C is found to be stable in *in vitro* systems and
558 is predicted to be stable in yeast, but is destabilized in mouse. L100V and P101S are destabilized
559 *in vitro*, but are predicted to be stable in yeast. R133H is stable *in vitro*, but is predicted to be
560 destabilized in yeast. Hence, the discrepancies between these data emphasize the value of
561 comparing variant function across multiple species and assay systems. MECP2 turnover in cells
562 is regulated by ubiquitylation by the E3 ligase RNF4 [58], indicating it may be an *in vivo*-specific
563 modifier of variant stability. Interestingly, we identified the yeast orthologs of RNF4, *SLX5* and
564 *SLX8*, as aggravating genetic interactions in our SIM screen (Figure 3B), suggesting MECP2
565 levels may also be regulated by this E3 ligase in yeast. Thus, regulated turnover *in vivo* is likely
566 an important factor contributing to MECP2 variant function and may vary across assay systems.
567 Interaction of MECP2 with DNA and co-repressors also influences MECP2 intrinsic stability [53],
568 hence differences in these interactions could be another factor contributing to differences in
569 MECP2 variant function/stability between assay systems.

570

571 ***MECP2 and DNA methylation***

572 It is noteworthy that all pathogenic calibration variants in the MBD were LoF in our functional
573 assays despite there being no, or extremely low levels of DNA cytosine methylation in yeast and
574 *Drosophila*, respectively. Our data correlate well with Brown *et al.* [36], showing that MECP2
575 binds the unmethylated yeast genome and mediates widespread transcriptional suppression. This
576 raises the possibility that ASD and Rett MECP2 variants may not be specifically deficient in
577 their ability to bind methylated DNA, but may be generally deficient in binding to unmethylated
578 DNA and/or chromatin. It is challenging to specify differential impacts of specific Rett variants
579 on methyl versus non-methyl DNA binding in models where DNA methylation is present. We
580 propose this may have obscured the impact that pathogenic variants have on non-methylated
581 DNA binding and its importance in healthy and disease states, and we propose that the yeast and
582 fly models offer valuable platforms to further examine these mechanisms.

583

584 **REFERENCES**

585 1. Lewis, J. D. *et al.* Purification, sequence, and cellular localization of a novel chromosomal
586 protein that binds to Methylated DNA. *Cell* **69**, 905–914 (1992).
587 2. Amir, R. E. *et al.* Rett syndrome is caused by mutations in X-linked MECP2, encoding
588 methyl-CpG-binding protein 2. *Nat. Genet.* **23**, 185–188 (1999).
589 3. Wan, M. *et al.* Rett Syndrome and Beyond: Recurrent Spontaneous and Familial MECP2
590 Mutations at CpG Hotspots. *Am. J. Hum. Genet.* **65**, 1520–1529 (1999).
591 4. Chahrour, M. & Zoghbi, H. Y. The Story of Rett Syndrome: From Clinic to Neurobiology.
592 *Neuron* **56**, 422–437 (2007).

593 5. Nomura, Y. Early behavior characteristics and sleep disturbance in Rett syndrome. *Brain*
594 *Dev.* **27**, S35–S42 (2005).
595 6. Amir, R. E. *et al.* Influence of mutation type and X chromosome inactivation on Rett
596 syndrome phenotypes. *Ann. Neurol.* **47**, 670–679 (2000).
597 7. Ananth, A. *et al.* MECP2 Variants in Males: More Common than Previously Appreciated.
598 *Pediatr. Neurol.* **161**, 263–267 (2024).
599 8. Pascual-Alonso, A., Martínez-Monseny, A. F., Xiol, C. & Armstrong, J. MECP2-Related
600 Disorders in Males. *Int. J. Mol. Sci.* **22**, 9610 (2021).
601 9. Ng, H. H. & Bird, A. DNA methylation and chromatin modification. *Curr. Opin. Genet.*
602 *Dev.* **9**, 158–163 (1999).
603 10. Jones, P. L. *et al.* Methylated DNA and MeCP2 recruit histone deacetylase to repress
604 transcription. *Nat. Genet.* **19**, 187–191 (1998).
605 11. Harikrishnan, K. N. *et al.* Brahma links the SWI/SNF chromatin-remodeling complex with
606 MeCP2-dependent transcriptional silencing. *Nat. Genet.* **37**, 254–264 (2005).
607 12. Lunyak, V. V. *et al.* Corepressor-Dependent Silencing of Chromosomal Regions Encoding
608 Neuronal Genes. *Science* **298**, 1747–1752 (2002).
609 13. Ehrhart, F. *et al.* A catalogue of 863 Rett-syndrome-causing MECP2 mutations and lessons
610 learned from data integration. *Sci. Data* **8**, 10 (2021).
611 14. Yusufzai, T. M. & Wolffe, A. P. Functional consequences of Rett syndrome mutations on
612 human MeCP2. *Nucleic Acids Res.* **28**, 4172–4179 (2000).
613 15. Ballestar, E., Yusufzai, T. M. & Wolffe, A. P. Effects of Rett Syndrome Mutations of the
614 Methyl-CpG Binding Domain of the Transcriptional Repressor MeCP2 on Selectivity for
615 Association with Methylated DNA. *Biochemistry* **39**, 7100–7106 (2000).

616 16. Leblond, C. S. *et al.* Operative list of genes associated with autism and neurodevelopmental
617 disorders based on database review. *Mol. Cell. Neurosci.* **113**, 103623 (2021).
618 17. Wen, Z. *et al.* Identification of autism-related MECP2 mutations by whole-exome
619 sequencing and functional validation. *Mol. Autism* **8**, 43 (2017).
620 18. Wang, T. *et al.* De novo genic mutations among a Chinese autism spectrum disorder cohort.
621 *Nat. Commun.* **7**, 13316 (2016).
622 19. Guo, H. *et al.* Inherited and multiple de novo mutations in autism/developmental delay risk
623 genes suggest a multifactorial model. *Mol. Autism* **9**, 64 (2018).
624 20. Schaaf, C. P. *et al.* Oligogenic heterozygosity in individuals with high-functioning autism
625 spectrum disorders. *Hum. Mol. Genet.* **20**, 3366–3375 (2011).
626 21. Landrum, M. J. *et al.* ClinVar: improving access to variant interpretations and supporting
627 evidence. *Nucleic Acids Res.* **46**, D1062–D1067 (2018).
628 22. Cheng, T.-L. *et al.* MeCP2 suppresses nuclear microRNA processing and dendritic growth
629 by regulating the DGCR8/Drosha complex. *Dev. Cell* **28**, 547–560 (2014).
630 23. Georgel, P. T. *et al.* Chromatin compaction by human MeCP2. Assembly of novel secondary
631 chromatin structures in the absence of DNA methylation. *J. Biol. Chem.* **278**, 32181–32188
632 (2003).
633 24. Hansen, J. C., Ghosh, R. P. & Woodcock, C. L. Binding of the Rett syndrome protein,
634 MeCP2, to methylated and unmethylated DNA and chromatin. *IUBMB Life* **62**, 732–738
635 (2010).
636 25. Skene, P. J. *et al.* Neuronal MeCP2 is expressed at near histone-octamer levels and globally
637 alters the chromatin state. *Mol. Cell* **37**, 457–468 (2010).

638 26. Lyko, F., Ramsahoye, B. H. & Jaenisch, R. DNA methylation in Drosophila melanogaster.
639 *Nature* **408**, 538–540 (2000).
640 27. Buitrago, D. *et al.* Impact of DNA methylation on 3D genome structure. *Nat. Commun.* **12**,
641 3243 (2021).
642 28. Ganguly, P. *et al.* A scalable Drosophila assay for clinical interpretation of human PTEN
643 variants in suppression of PI3K/AKT induced cellular proliferation. *PLoS Genet.* **17**,
644 e1009774 (2021).
645 29. Post, K. L. *et al.* Multi-model functionalization of disease-associated PTEN missense
646 mutations identifies multiple molecular mechanisms underlying protein dysfunction. *Nat.*
647 *Commun.* **11**, 2073 (2020).
648 30. Young, B. P. *et al.* Sentinel interaction mapping - a generic approach for the functional
649 analysis of human disease gene variants using yeast. *Dis. Model. Mech.* **13**, dmm044560
650 (2020).
651 31. Her, Y., Pascual, D. M., Goldstone-Joubert, Z. & Marcogliese, P. C. Variant functional
652 assessment in Drosophila by overexpression: what can we learn? *Genome* **67**, 158–167
653 (2024).
654 32. Cukier, H. N. *et al.* Genetic Modifiers of MeCP2 Function in Drosophila. *PLoS Genet.* **4**,
655 e1000179 (2008).
656 33. Vonhoff, F., Williams, A., Ryglewski, S. & Duch, C. Drosophila as a model for MECP2 gain
657 of function in neurons. *PloS One* **7**, e31835 (2012).
658 34. Williams, A. A. *et al.* MECP2 impairs neuronal structure by regulating KIBRA. *Neurobiol.*
659 *Dis.* **91**, 284–291 (2016).

660 35. Williams, A. A. *et al.* Apoptotic Activity of MeCP2 Is Enhanced by C-Terminal Truncating
661 Mutations. *PLoS ONE* **11**, e0159632 (2016).
662 36. Brown, J. A. R., Ling, M. Y. M., Ausió, J. & Howe, L. J. Human MeCP2 binds to promoters
663 and inhibits transcription in an unmethylated yeast genome. *Genetics* iyaf043 (2025)
664 doi:10.1093/genetics/iyaf043.
665 37. Brnich, S. E. *et al.* Recommendations for application of the functional evidence PS3/BS3
666 criterion using the ACMG/AMP sequence variant interpretation framework. *Genome Med.*
667 **12**, 3 (2019).
668 38. McKnight, D. *et al.* Recommendations by the ClinGen Rett/Angelman-like expert panel for
669 gene-specific variant interpretation methods. *Hum. Mutat.* **43**, 1097–1113 (2022).
670 39. Rodriguez, J. M. *et al.* APPRIS: annotation of principal and alternative splice isoforms.
671 *Nucleic Acids Res.* **41**, D110–D117 (2013).
672 40. Belmadani, M. *et al.* VariCarta: A Comprehensive Database of Harmonized Genomic
673 Variants Found in Autism Spectrum Disorder Sequencing Studies. *Autism Res. Off. J. Int.*
674 *Soc. Autism Res.* **12**, 1728–1736 (2019).
675 41. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).
676 42. Good, K. V., Vincent, J. B. & Ausió, J. MeCP2: The Genetic Driver of Rett Syndrome
677 Epigenetics. *Front. Genet.* **12**, 620859 (2021).
678 43. Bischof, J., Maeda, R. K., Hediger, M., Karch, F. & Basler, K. An optimized transgenesis
679 system for Drosophila using germ-line-specific phiC31 integrases. *Proc. Natl. Acad. Sci. U.*
680 *S. A.* **104**, 3312–3317 (2007).
681 44. Tong, A. H. *et al.* Systematic genetic analysis with ordered arrays of yeast deletion mutants.
682 *Science* **294**, 2364–2368 (2001).

683 45. Young, B. P. & Loewen, C. J. R. Balony: a software package for analysis of data generated
684 by synthetic genetic array experiments. *BMC Bioinformatics* **14**, 354 (2013).
685 46. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest Package: Tests in Linear
686 Mixed Effects Models. *J. Stat. Softw.* **82**, 1–26 (2017).
687 47. Symington, L. S. Role of *RAD52* Epistasis Group Genes in Homologous Recombination and
688 Double-Strand Break Repair. *Microbiol. Mol. Biol. Rev.* **66**, 630–670 (2002).
689 48. Syed, A. & Tainer, J. A. The MRE11–RAD50–NBS1 Complex Conducts the Orchestration
690 of Damage Signaling and Outcomes to Stress in DNA Replication and Repair. *Annu. Rev.*
691 *Biochem.* **87**, 263–294 (2018).
692 49. Bähler, J. Cell-Cycle Control of Gene Expression in Budding and Fission Yeast. *Annu. Rev.*
693 *Genet.* **39**, 69–94 (2005).
694 50. Jiménez, J., Ricco, N., Grijota-Martínez, C., Fadó, R. & Clotet, J. Redundancy or specificity?
695 The role of the CDK Pho85 in cell cycle control. *Int. J. Biochem. Mol. Biol.* **4**, 140–149
696 (2013).
697 51. Tillotson, R. & Bird, A. The Molecular Basis of MeCP2 Function in the Brain. *J. Mol. Biol.*
698 **432**, 1602–1623 (2020).
699 52. Free, A. *et al.* DNA Recognition by the Methyl-CpG Binding Domain of MeCP2 *. *J. Biol.*
700 *Chem.* **276**, 3353–3360 (2001).
701 53. Ghosh, R. P. *et al.* Unique physical properties and interactions of the domains of methylated
702 DNA binding protein 2. *Biochemistry* **49**, 4395–4410 (2010).
703 54. Lyst, M. J. *et al.* Rett syndrome mutations abolish the interaction of MeCP2 with the
704 NCoR/SMRT co-repressor. *Nat. Neurosci.* **16**, 898–902 (2013).

705 55. Yang, Y., Kucukkal, T. G., Li, J., Alexov, E. & Cao, W. Binding Analysis of Methyl-CpG
706 Binding Domain of MeCP2 and Rett Syndrome Mutations. *ACS Chem. Biol.* **11**, 2706–2715
707 (2016).
708 56. Kucukkal, T. G., Yang, Y., Uvarov, O., Cao, W. & Alexov, E. Impact of Rett Syndrome
709 Mutations on MeCP2 MBD Stability. *Biochemistry* **54**, 6357–6368 (2015).
710 57. Sperlazza, M. J., Bilinovich, S. M., Sinanan, L. M., Javier, F. R. & Williams, D. C.
711 Structural Basis of MeCP2 Distribution on Non-CpG Methylated and Hydroxymethylated
712 DNA. *J. Mol. Biol.* **429**, 1581–1594 (2017).
713 58. Wang, Y. RING finger protein 4 (RNF4) derepresses gene expression from DNA
714 methylation. *J. Biol. Chem.* **289**, 33808–33813 (2014).
715

716 **ACKNOWLEDGEMENTS**

717 We thank Dr. Janel Kopp, UBC for access to the Panoramic Midi II digital slide scanner. Susan
718 Lindquist for the generous gift of the *pAG426GAL-ccdB* vector. Stocks obtained from the
719 Bloomington *Drosophila* Stock Center (NIH P40OD018537) were used in this study. We thank
720 *Drosophila* Genomics Resource Center, supported by NIH grant 2P40OD010949 for the esc
721 cDNA. We used FlyBase (releases 2020 through 2024) to find information on
722 phenotypes/function/stocks/gene expression (etc). We would like to thank Gino Laberge at
723 Genome ProLab for the kind gift of balancer stocks and also their transgenic services (Genome
724 ProLab Inc, Quebec, Canada). We thank Libby Natola and Guillaume Poirier-Morency (Pavlidis
725 Lab) for contributions to the variant annotation pipeline. This work was funded by an operating
726 grant from The Simons Foundation Autism Research Initiative (2021 Genomics of ASD:
727 Pathway to Genetic Therapies: AWD-021951. SIMOFOUN 2021).

728

729 **AUTHOR CONTRIBUTIONS**

730 DWA, CJRL, BPY, SMG, JTC, MD, SR, MD, PP contributed to conception of the work.
731 DWA, CJRL, GM, EC, BPY, SI, TL, JL, JBL, SMG, BB, MD, JS, SR, PP contributed to design
732 of the work.
733 DWA, CJRL, GM, EC, BPY, JBL, LM, JS, JBL, SR, PP contributed to acquisition, analysis, or
734 interpretation of data.
735 DWA, SR, PP, GM, EC, JBL, MD, BB, BW contributed to creation of new software used.
736 DWA, CJRL, JS, BPY, JTC, JBL, GM, EC, LM, JL, JBL, JTC, SR, PP contributed to drafting
737 the paper or substantively revised it.

738

739 **FIGURE CAPTIONS**

740 **Figure 1:** The forty-six MECP2 variants examined in this study. Domain locations of variants as
741 well as variant class, variant inheritance, and whether they are found in Rett Syndrome and/or
742 ASD. NTD – N-terminal domain, MBD – Methyl-CpG-binding domain, ID – Interdomain, TRD
743 – Transcription repressor domain, CTD – C-terminal domain.

744

745 **Figure 2:** MECP2 expression and effect on growth in yeast. (A) Untagged MECP2-ref expressed
746 from a 2-micron plasmid utilizing the estradiol-responsive promoter at various estradiol
747 concentrations in WT yeast. Equivalent amounts of yeast were loaded into each lane and western
748 blotted with anti-MECP2 antibody. M – molecular weight markers, arrowhead – full length
749 MECP2. (B) Effect of MECP2-ref expression on growth of WT yeast from a plasmid utilizing
750 the estradiol-responsive promoter at various estradiol concentrations. +EV – empty vector in

751 wild type yeast, +MECP2-ref – MECP2-ref plasmid in wild type yeast. (C) Similar to B except
752 using a 2-micron plasmid containing a galactose-responsive promoter and grown in galactose.
753 **** – indicates p value < 0.0001. ns – not significant.
754
755 **Figure 3:** Results of SIM genetic interaction screen with MECP2-ref. (A) Functional enrichment
756 for genetic interactions is shown where nodes indicate enriched GO terms and edges indicate
757 statistically significant associations between terms. Node size and edge thickness indicate
758 statistical significance (minimum $p < 0.05$, Bonferonni corrected). Highlighted labels indicate
759 most significant terms within networks. Network generated using Cytoscape and the ClueGO
760 plug-in. (B) Effect of MECP2-ref on yeast growth in mutants from the GO categories "DNA
761 damage response", "Chromatin organization" and "DNA binding" that were identified in the
762 SIM screen. Normalized colony size indicates the size of the colonies on the array plates after
763 accounting for the effect of MECP2-ref expression in WT yeast, in non-inducing (0 nM
764 estradiol) and inducing (10 nM estradiol) conditions.
765
766 **Figure 4:** Analysis of 46 MECP2 variants using the yeast assay. (A) Estimated functional effects
767 of MECP2 variants expressed in WT yeast. Data is normalized to MECP2-ref (=1) and empty
768 vector (=0). Error bars indicate 95% confidence intervals. (B) Protein stability of MECP2
769 variants in WT yeast. Protein levels were normalized to MECP2-ref and plotted on a log scale.
770 Error bars indicate S.D. ($n = 3$). * indicates $p < 0.05$, ** indicates $p < 0.01$. Colour coding same
771 as in (A). All variants are located in MBD except T228S (TRD), S356G (CTD) and P388S
772 (CTD).
773

774 **Figure 5:** *Drosophila* wing phenotypes induced by MECP2 variants. Representative images of
775 wings are shown expressing the indicated controls and variants. *attp40* – no MECP2, *MECP2-ref*
776 – WT MECP. All images are the same scale. Arrows indicate ectopic vein tissue for *UAS-*
777 *MECP2-ref* only.
778
779 **Figure 6:** Analysis of MECP2 variant wing vein phenotypes. We show our quantitation of total
780 wing vein per wing area for each genotype, normalized to *MECP2-ref* (WT MECP2) = 1 and
781 *attp40* (no MECP2) = 0. Error bars indicate 95% confidence intervals.
782
783 **SUPPLEMENTAL FIGURE CAPTIONS**
784 **Supplemental Figure 1**: Representative MECP2 Western blots. Arrow indicates the full-length
785 MECP2 protein. Antibody labelling is as follows: red - anti-MECP2, green - anti-PGK1.
786
787 **Supplemental Figures 2A and B**: Representative wing phenotypes for all MECP2 variants
788 tested in *Drosophila*. For each variant, raw images are on the left and the masked image used to
789 quantitate total wing vein per wing is on the right.
790
791 **SUPPLEMENTAL METHODS CAPTIONS**
792 **Supplemental Methods 1:** Detailed outline of wing vein quantitation methods.
793 **Supplemental Methods 2:** Yeast statistical model R script.
794 **Supplemental Methods 3:** Fly statistical model R script.
795
796 **SUPPLEMENTAL TABLE CAPTIONS**

797 **Supplemental Table 1:** MECP2 protein stability summary table.

798

799 **SUPPLEMENTAL DATA CAPTIONS**

800 **Supplemental Data 1:** Detailed information for each variant, primers used in cloning, the

801 molecular genetic reagents made for each organism, and the *Drosophila* stocks generated.

802 **Supplemental Data 2:** Yeast SIM screen raw data.

803 **Supplemental Data 3:** GO analysis results for yeast SIM screen listing genes in each GO term.

804 **Supplemental Data 4:** Raw data for yeast variant assays used in statistical model.

805 **Supplemental Data 5:** Raw data for *Drosophila* variant assays used in statistical model.

806

Figure 1

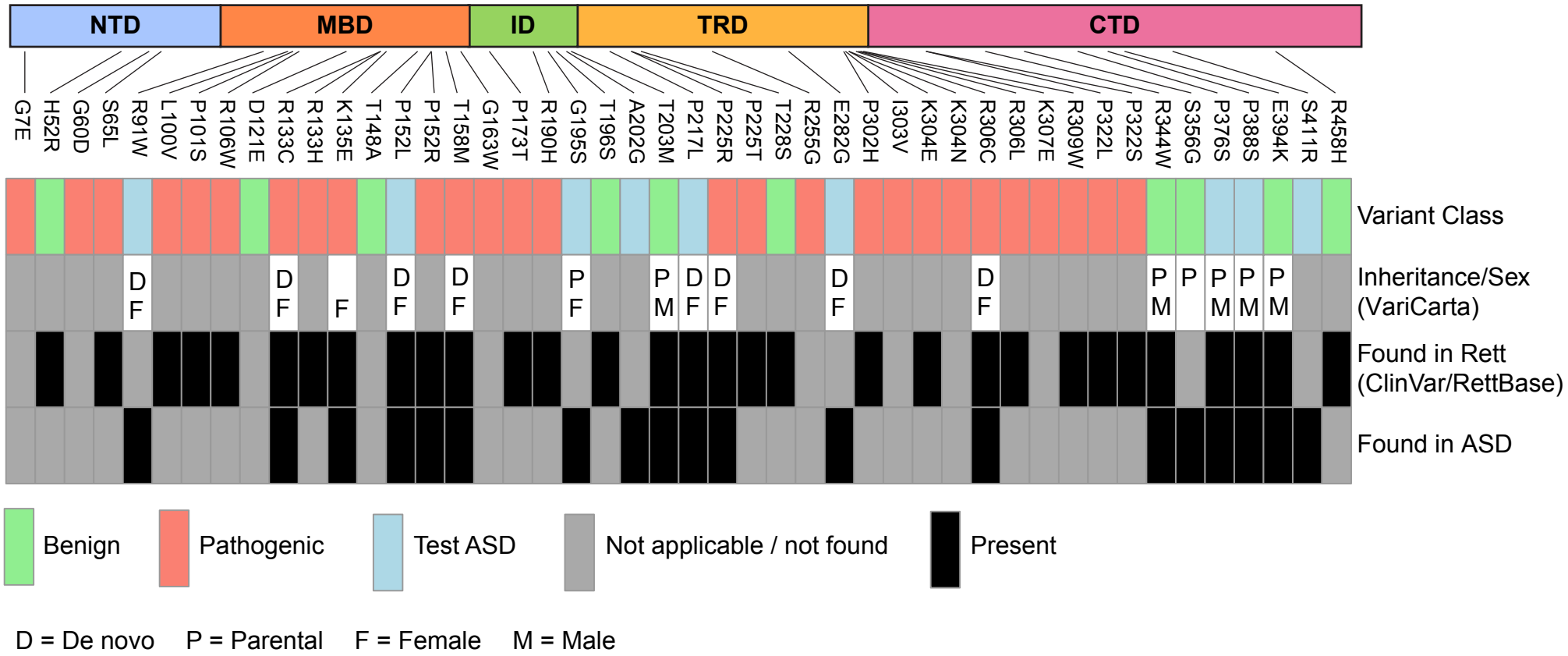

NTD
MBD
ID
TRD
CTD
G7E
H52R
G60D
S65L
R91W
L100V
P101S
R106W
D121E
R133C
R133H
K135E
T148A
P152L
P152R
T158M
G163W
P173T
R190H
G195S
T196S
A202G
T203M
P217L
P225R
P225T
T228S
R255G
E282G
P302H
I303V
K304E
K304N
R306C
R306L
K307E
R309W
P322L
P322S
R344W
S356G
P376S
P388S
E394K
S411R
R458H
Variant Class
Inheritance/Sex (VariCarta)
Found in Rett (ClinVar/RettBase)
Found in ASD
Benign
Pathogenic
Test ASD
Not applicable / not found
Present
D = De novo
P = Parental
F = Female
M = Male

Figure 2

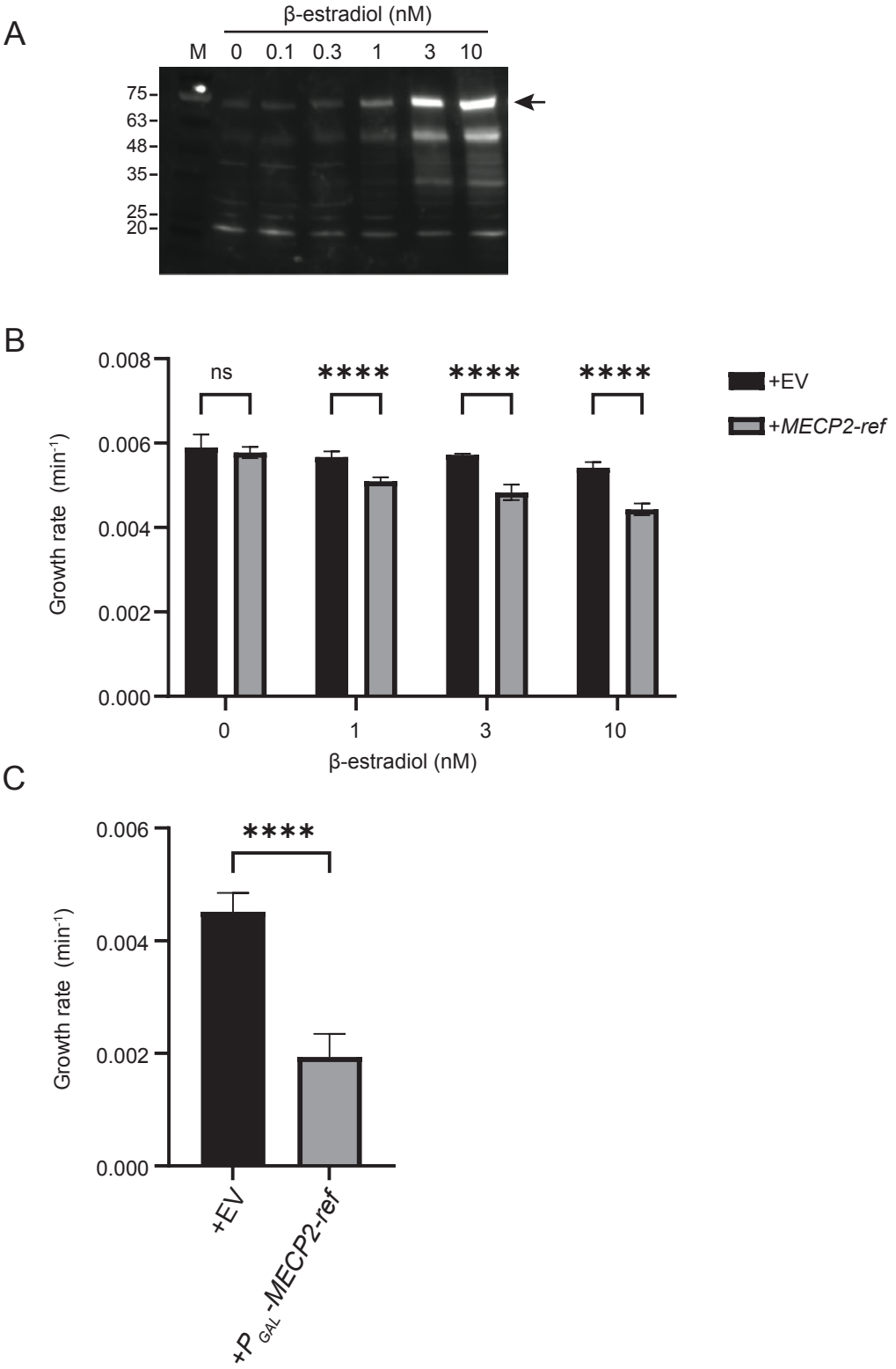
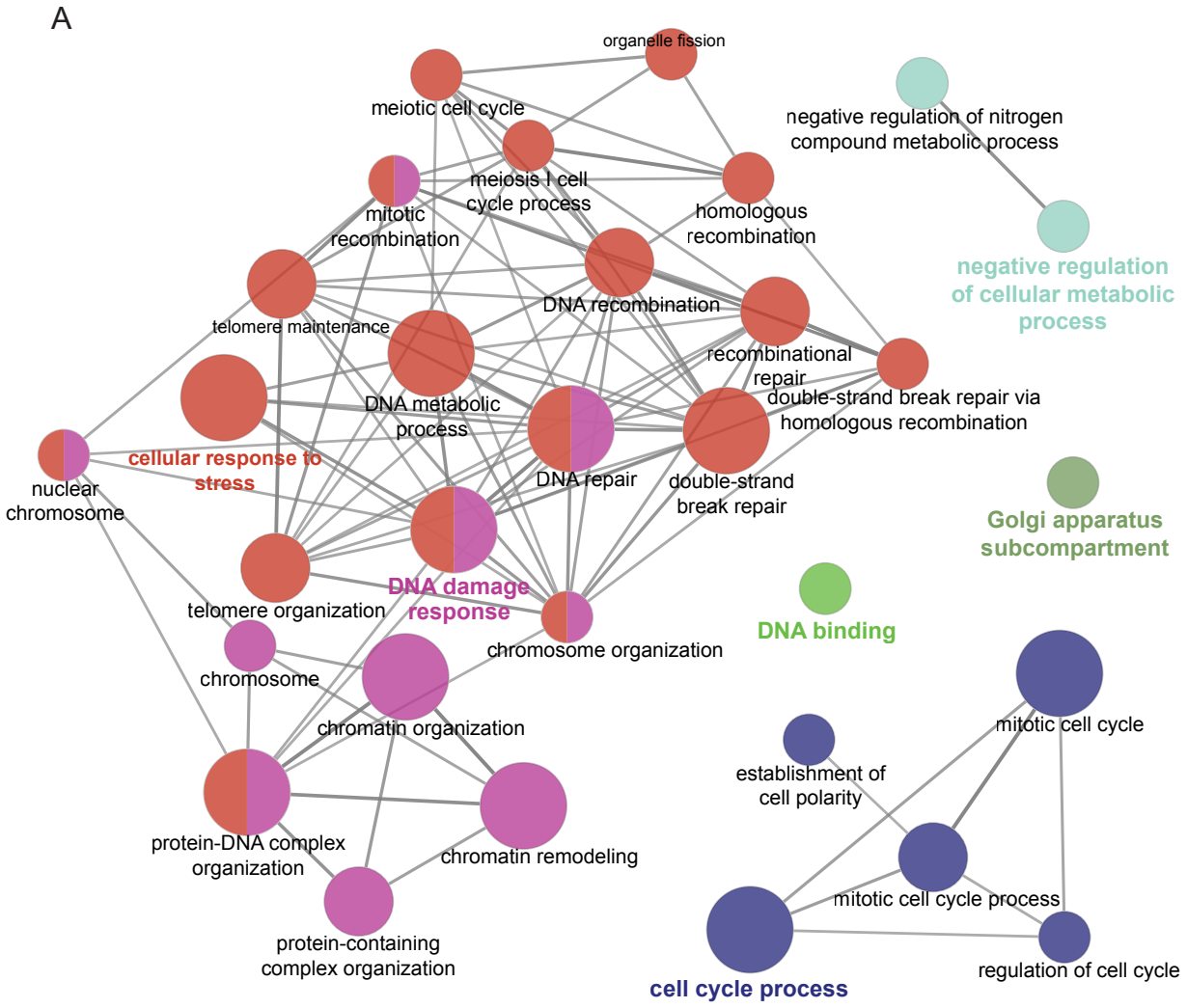

A
β-estradiol (nM)
M 0 0.1 0.3 1 3 10
75
63
48
35
25
20
B
0.008
0.006
0.004
0.002
0.000
Growth rate (min-1)
ns
****
****
****
+EV
+MECP2-ref
0
1
3
10
β-estradiol (nM)
C
0.006
0.004
0.002
0.000
Growth rate (min-1)
****
+EV
+P GAL -MECP2-ref

Figure 3

A


organelle fission
meiotic cell cycle
negative regulation of nitrogen compound metabolic process
meiosis I cell cycle process
mitotic recombination
homologous recombination
negative regulation of cellular metabolic process
DNA recombination
telomere maintenance
recombinational repair
double-strand break repair via homologous recombination
DNA metabolic process
cellular response to stress
nuclear chromosome
DNA repair
double-strand break repair
Golgi apparatus subcompartment
DNA damage response
telomere organization
DNA binding
chromosome organization
chromosome
chromatin organization
mitotic cell cycle
establishment of cell polarity
protein-DNA complex organization
chromatin remodeling
mitotic cell cycle process
protein-containing complex organization
regulation of cell cycle
cell cycle process

B

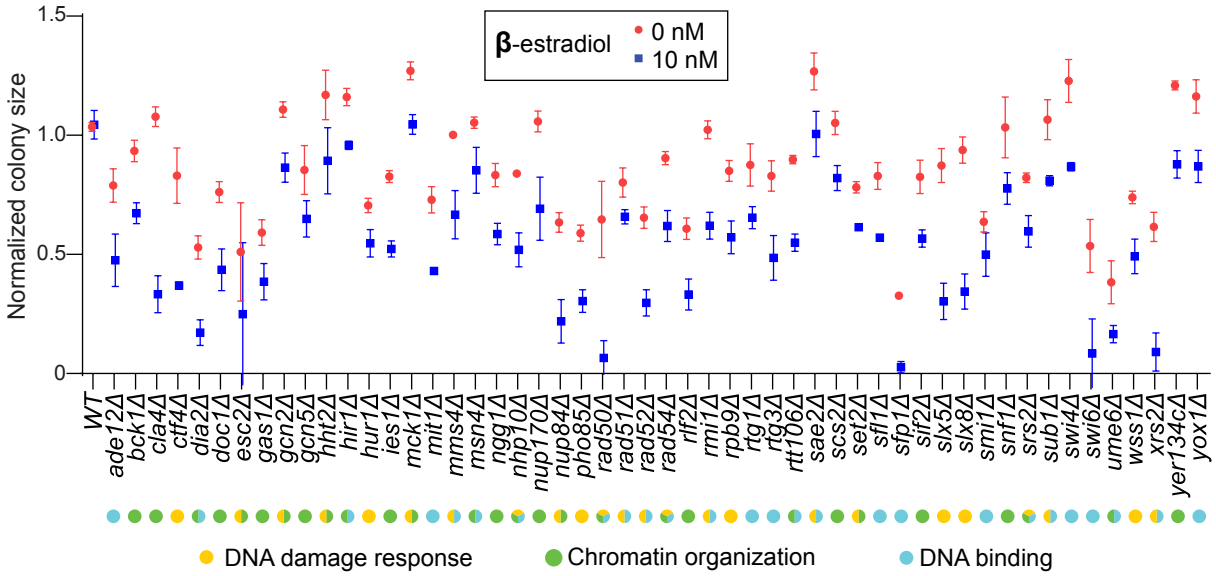

β-estradiol
0 nM
10 nM
Normalized colony size
1.5
1.0
0.5
0
WT
ade12Δ
bck1Δ
cla4Δ
ctf4Δ
dia2Δ
doc1Δ
esc2Δ
gas1Δ
gcn2Δ
gcn5Δ
hht2Δ
hir1Δ
hur1Δ
ies1Δ
mck1Δ
mit1Δ
mms4Δ
msn4Δ
ngg1Δ
nhp10Δ
nup170Δ
nup84Δ
pho85Δ
rad50Δ
rad51Δ
rad52Δ
rad54Δ
rif2Δ
rmi1Δ
rpb9Δ
rtg1Δ
rtg3Δ
rtt106Δ
sae2Δ
scs2Δ
set2Δ
sfi1Δ
sfp1Δ
sir2Δ
slx5Δ
slx8Δ
smi1Δ
snf1Δ
srs2Δ
sub1Δ
swi4Δ
swi6Δ
ume6Δ
wss1Δ
xrs2Δ
yer134cΔ
yox1Δ
DNA damage response
Chromatin organization
DNA binding

Figure 4

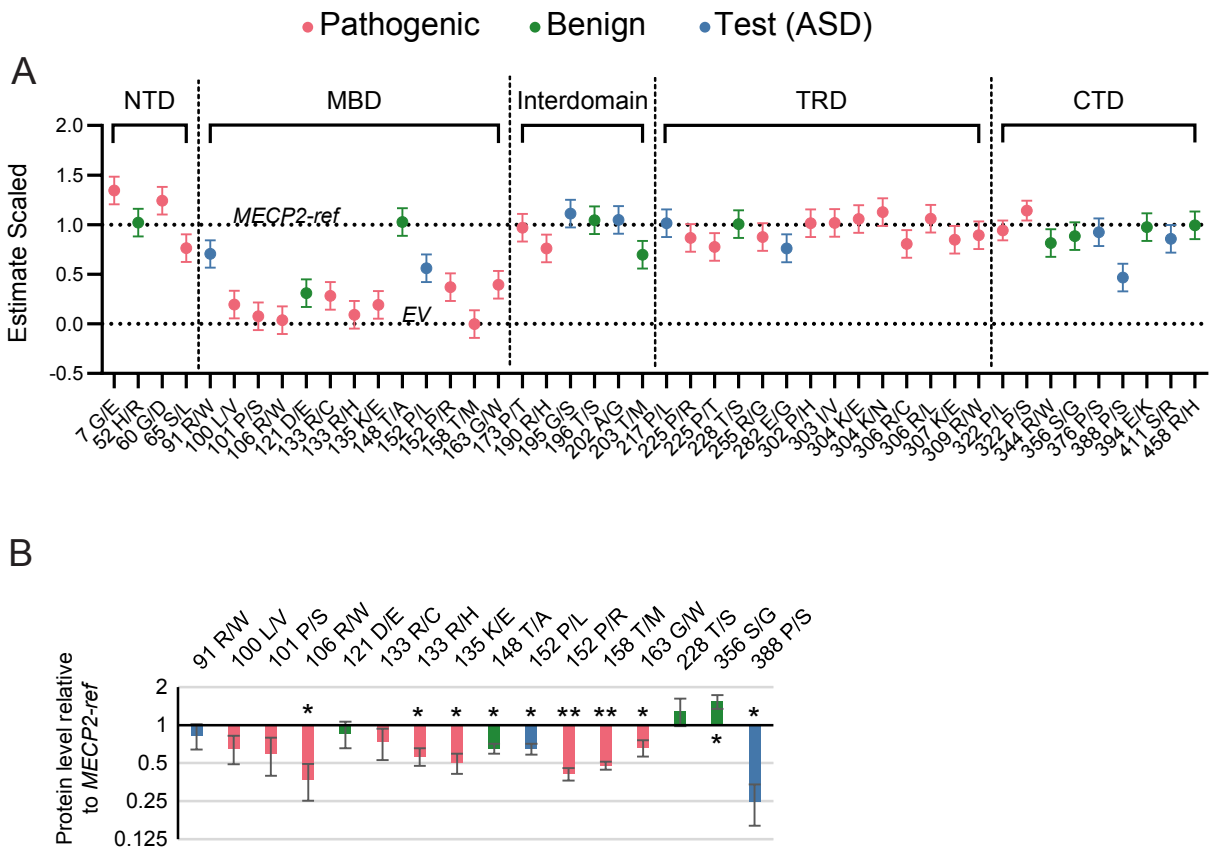

Pathogenic
Benign
Test (ASD)
A
NTD
MBD
Interdomain
TRD
CTD
Estimate Scaled
2.0
1.5
1.0
0.5
0.0
-0.5
MECP2-ref
EV
7 G/E
52 H/R
60 G/D
65 S/L
91 R/W
100 L/V
101 P/S
106 R/W
121 D/E
133 R/C
133 R/H
135 K/E
148 T/A
152 P/L
152 P/R
158 T/M
163 G/W
173 P/T
190 R/H
195 G/S
196 T/S
202 A/G
203 T/M
217 P/L
225 P/R
225 P/T
228 T/S
255 R/G
282 E/G
302 P/H
303 I/V
304 K/E
304 K/N
306 R/C
306 R/L
307 K/E
309 R/W
322 P/L
322 P/S
344 R/W
356 S/G
376 P/S
388 P/S
394 E/K
411 S/R
458 R/H
B
91 R/W
100 L/V
101 P/S
106 R/W
121 D/E
133 R/C
133 R/H
135 K/E
148 T/A
152 P/L
152 P/R
158 T/M
163 G/W
228 T/S
356 S/G
388 P/S
Protein level relative to MECP2-ref
2
1
0.5
0.25
0.125
*
*
*
*
*
**
**
*
*
*

Figure 5

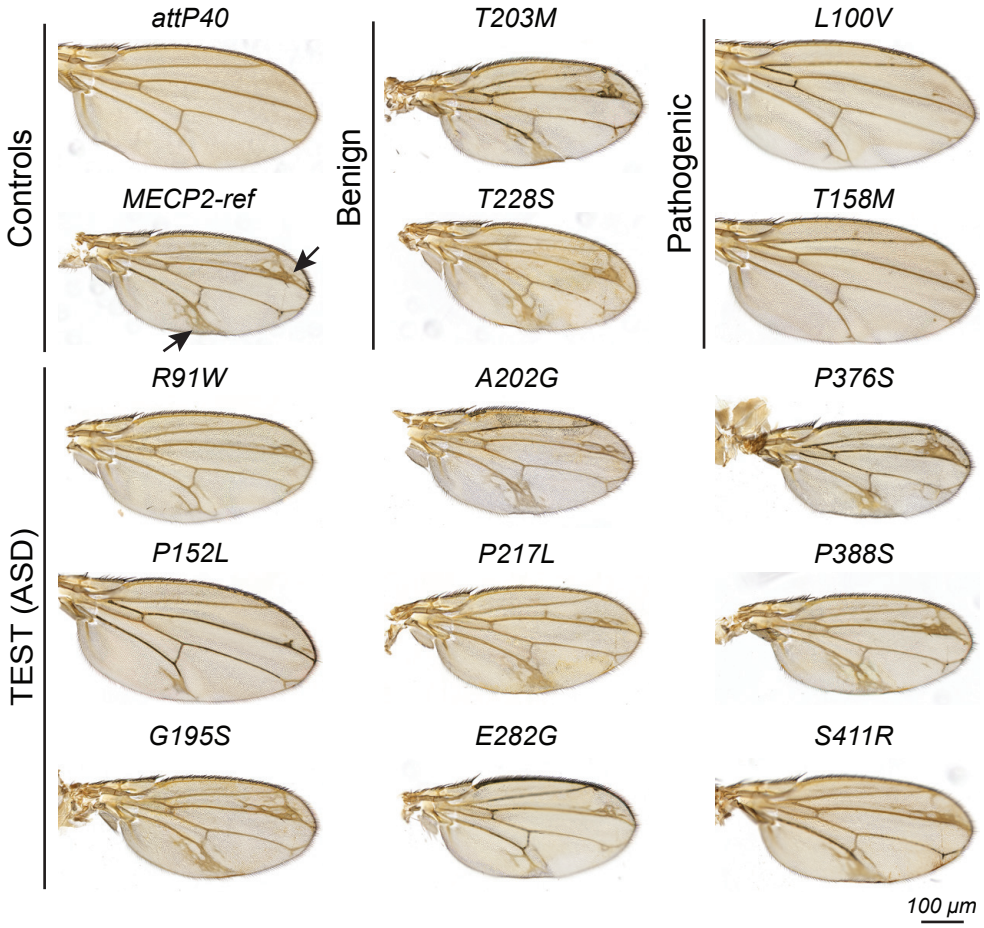

Controls
attP40
MECP2-ref
Benign
T203M
T228S
Pathogenic
L100V
T158M
TEST (ASD)
R91W
A202G
P376S
P152L
P217L
P388S
G195S
E282G
S411R
100 µm

Figure 6

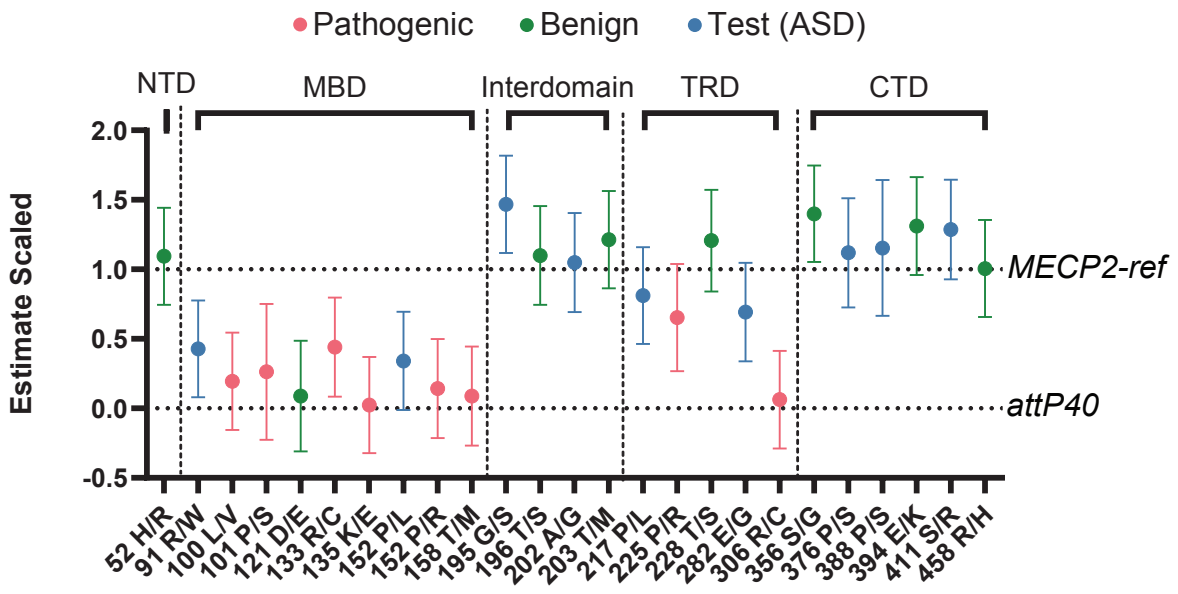

Pathogenic
Benign
Test (ASD)
NTD
MBD
Interdomain
TRD
CTD
Estimate Scaled
2.0
1.5
1.0
0.5
0.0
-0.5
MECP2-ref
attP40
52 H/R
91 R/W
100 L/V
101 P/S
121 D/E
133 R/C
135 K/E
152 P/L
152 P/R
158 T/M
195 G/S
196 T/S
202 A/G
203 T/M
217 P/L
225 P/R
228 T/S
282 E/G
306 R/C
356 S/G
376 P/S
388 P/S
394 E/K
411 S/R
458 R/H

Figure S1

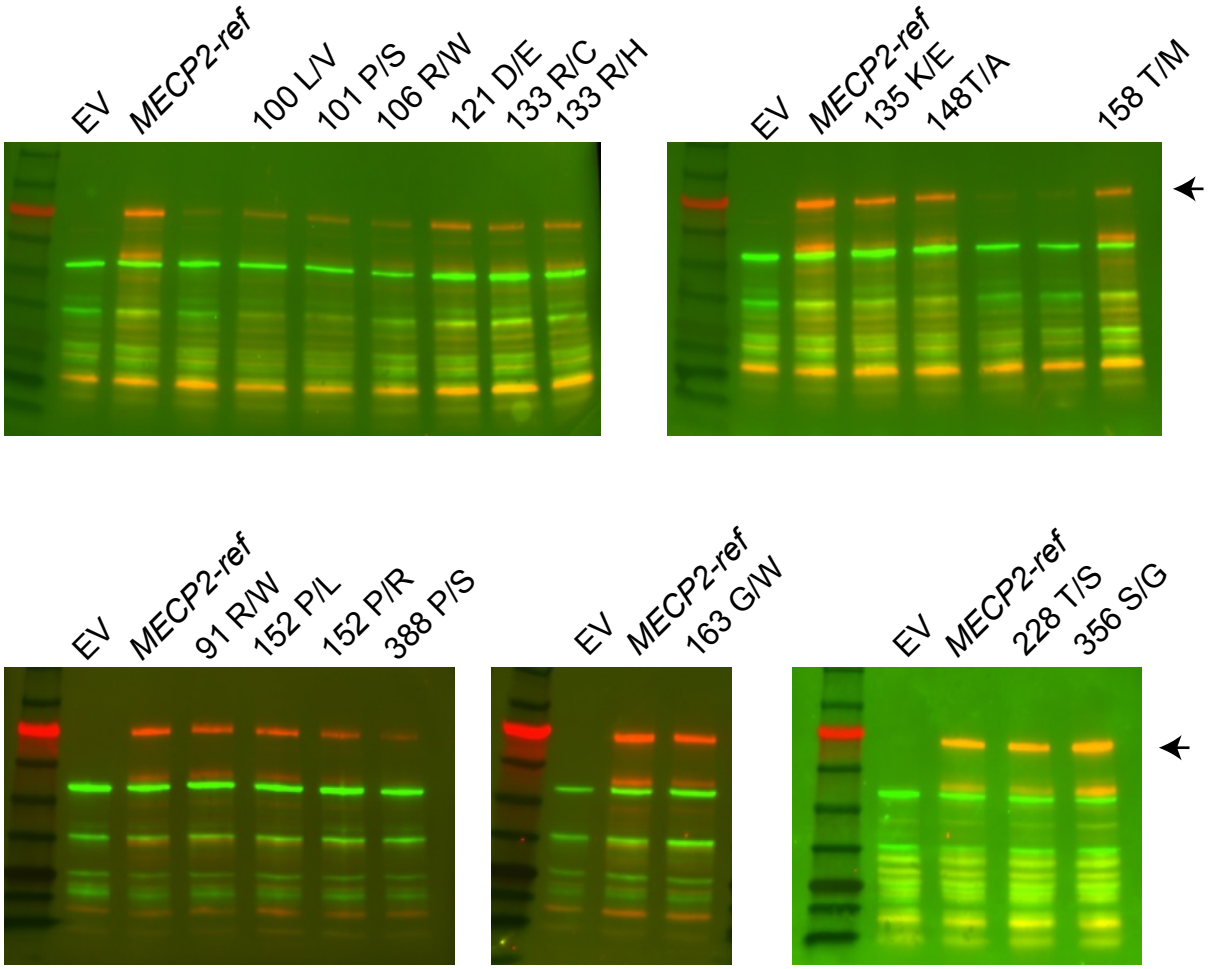

EV
MECP2-ref
100 L/V
101 P/S
106 R/W
121 D/E
133 R/C
133 R/H
EV
MECP2-ref
135 K/E
148T/A
158 T/M
EV
MECP2-ref
91 R/W
152 P/L
152 P/R
388 P/S
EV
MECP2-ref
163 G/W
EV
MECP2-ref
228 T/S
356 S/G

Figure S2A

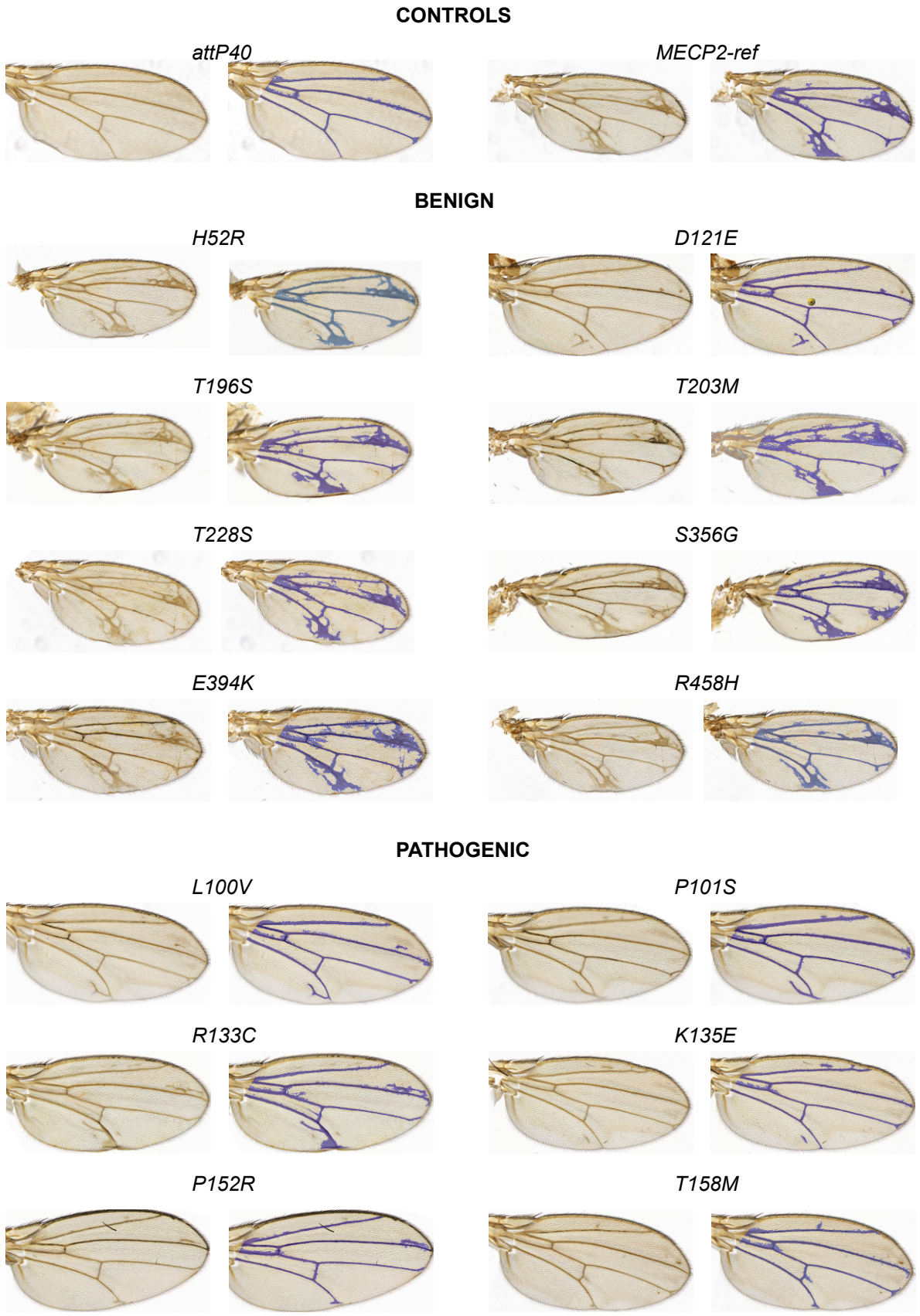
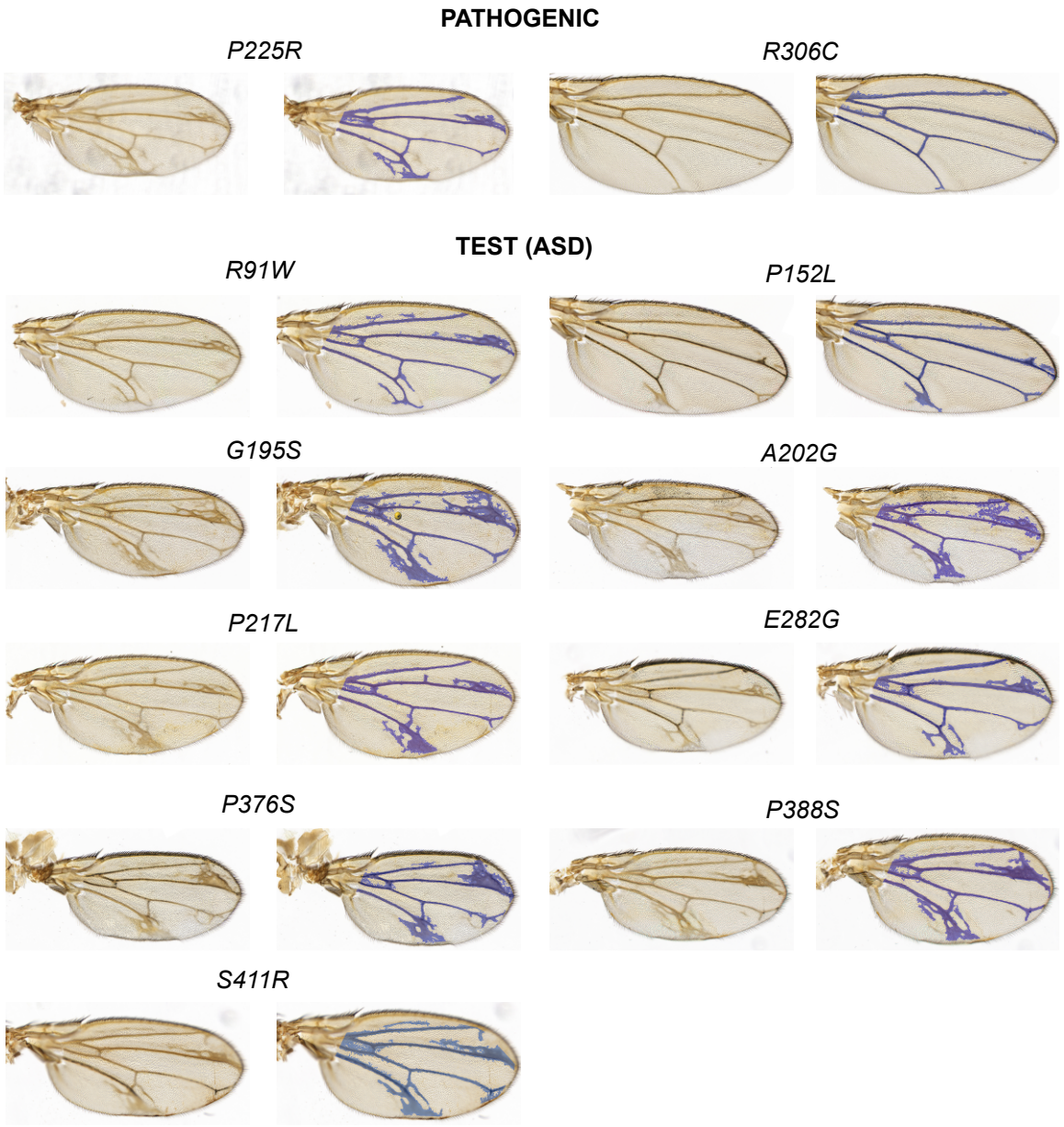

CONTROLS
attP40
MECP2-ref
BENIGN
H52R
D121E
T196S
T203M
T228S
S356G
E394K
R458H
PATHOGENIC
L100V
P101S
R133C
K135E
P152R
T158M

Figure S2B


PATHOGENIC
P225R
R306C
TEST (ASD)
R91W
P152L
G195S
A202G
P217L
E282G
P376S
P388S
S411R

Supplementary Table 1: Summary of MECP2 protein stability results

Variant	Yeast* (this study)	Yang [1] (in vitro, WT = 1)	Kucukkal [2] (in vitro) (ΔΔG)	Sperlazza [3] (NMR) (+ slight, ++ more destabilized)	Tillotson [4] (mouse models, WT = 1)
R91W	1				
L100V	1	0.5	+1.25		
P101S	1			+	
R106W	0.37	1	+0.1	++	0.1 (TAVI)
D121E	1				
R133C	1	1	+0.4		0.6 (eGFP)
R133H	0.56	1	+0.25		
K135E	0.50				
T148A	0.65				
P152L	0.65				
P152R	0.41	0.25	+1.5		
T158M	0.48	1	+0.25	+	0.3 (eGFP), 0.25 (TAVI)
G163W	0.66				
P225R					0.22
T228S	1				
R306C					1.0
P322L					0.03
S356G	1.54				
P388S	0.25				

* Values not significantly different than MECP2-Ref are set to 1, all others are significant ($p<0.05$)

References:

1. Yang, Y., Kucukkal, T. G., Li, J., Alexov, E. & Cao, W. Binding Analysis of Methyl-CpG Binding Domain of MeCP2 and Rett Syndrome Mutations. *ACS Chem. Biol.* **11**, 2706–2715 (2016).
2. Kucukkal, T. G., Yang, Y., Uvarov, O., Cao, W. & Alexov, E. Impact of Rett Syndrome Mutations on MeCP2 MBD Stability. *Biochemistry* **54**, 6357–6368 (2015).
3. Sperlazza, M. J., Bilinovich, S. M., Sinanan, L. M., Javier, F. R. & Williams, D. C. Structural Basis of MeCP2 Distribution on Non-CpG Methylated and Hydroxymethylated DNA. *J. Mol. Biol.* **429**, 1581–1594 (2017).
4. Tillotson, R. & Bird, A. The Molecular Basis of MeCP2 Function in the Brain. *J. Mol. Biol.* **432**, 1602–1623 (2020).