Molecular Therapy
Nucleic Acids
Brief report

Multi-gene duplication removal in an engineered human cellular MECP2 duplication syndrome model with an *IRAK1-MECP2* duplication

Samar Z. Rizvi,[1,3,7] Wing Suen Chan,[1,7] Eleonora Maino,[2] Sydney Steiman,[1,4] Georgiana Forguson,[1] Maya Klepfish,[1] Ronald D. Cohn,[1,3,5,6,8] and Evgueni A. Ivakine[1,3,8]

[1]Program in Genetics and Genome Biology, The Hospital for Sick Children Research Institute, Toronto, ON M5G 0A4, Canada; [2]Biozentrum, The Center for Molecular Life Sciences, University of Basel, Basel 4056, Switzerland; [3]Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada; [4]Department of Physiology, University of Toronto, Toronto, ON M5S 1A8, Canada; [5]Institute of Medical Science, University of Toronto, Toronto, ON M5S 1A8, Canada; [6]Department of Pediatrics, The Hospital for Sick Children, Toronto, ON M5G 1E8, Canada

Recent progress in genome editing technologies has catalyzed the generation of sophisticated cell models; however, the precise modeling of copy-number variation (CNV) diseases remains a significant challenge despite their substantial prevalence in the human population. To overcome this barrier, we have explored the utility of HAP1 cells for the accurate modeling of disease genomes with large structural variants. As an example, this study details the strategy to generate a novel cell line that serves as a model for the neurological disorder methyl CpG binding protein 2 (MECP2) duplication syndrome (MDS), featuring the critical duplication of both the *MECP2* and *IRAK1* genes. This model faithfully recapitulates MDS genomic rearrangement, allowing for the mechanistic study of gene overexpression and the development of therapeutic interventions. Employing a single-guide RNA (gRNA) CRISPR-Cas9 strategy, we successfully excised the duplicated genomic segment, notably halving both *MECP2* and *IRAK1* expression levels. The evidence establishes our model as a crucial tool for research into MDS. Furthermore, the outlined workflow is readily adaptable to model other CNV disorders and subsequently test genomic and pharmacological interventions.

INTRODUCTION

HAP1 cells are routinely used to generate disease models across the fields of immunology,[1,2] oncology,[3,4] and metabolism[5,6] due to their amenability to genetic manipulation and their rapid growth rate. This cell line can spontaneously diploidize in culture, enabling us to understand the mechanisms of pathogenicity and test different therapeutic genome editing strategies. Nonetheless, creating copy-number variations (CNVs) is significantly challenging due to the involvement of large DNA fragments. In this study, we evaluated an opportunity to utilize diploidized HAP1 cells for modeling CNV syndromes, focusing on methyl CpG binding protein 2 (*MECP2*) duplication syndrome (MDS).

MDS is currently an incurable disease[7] that results from a duplication of the *MECP2* locus on the X chromosome.[8] In MDS, duplications of the *MECP2* locus lead to neurodevelopmental and neurodegenerative symptoms including early developmental delay, intellectual disability, speech difficulties, seizures, and features of autism.[8] Importantly, the adjacent *IRAK1* gene is always duplicated with *MECP2* across patients with MDS, and the phenotypic contributions of this duplicated immune-regulatory gene are strikingly understudied.[9] IRAK1 plays a crucial role in adaptive immunity.[10] It is controversial whether *IRAK1* duplication contributes to peripheral immunologic phenotypes such as fatal respiratory tract infections, consistently reported in 70%–75% of patients with MDS.[11] Data from multi-omics analysis[12] and cytokines/chemokines analysis[13] from patients' samples showed contradictive results. The limited sample size and variability in human subjects and the critical lack of an *IRAK1* overexpression or duplication model together hinder the study of its functional role in MDS pathogenesis. Hence, a representative MDS model should necessarily harbor an *IRAK1-MECP2* duplication.

Here, we generated the disease model on diploidized HAP1 cells consisting of two X chromosomes. The *IRAK1-MECP2* fragment was duplicated on one chromosome and deleted on the other (i.e., a [Dup+Del] model). After establishing the disease model, we also used the single-guide RNA (gRNA) CRISPR-Cas9 approach to correct the duplication, demonstrating that the disease model is a powerful tool to test therapeutic strategies.

RESULTS

A dual-gRNA CRISPR-Cas9 system generates target-specific deletions, inversions, and duplications in HAP1 cells

To generate structural variants (SVs), a pair of gRNAs was required to flank the region to be modified (Figure S1A). In

Received 1 December 2023; accepted 4 October 2024;
https://doi.org/10.1016/j.omtn.2024.102356.

[7]These authors contributed equally

[8]Senior author

Correspondence: Evgueni A. Ivakine, Program in Genetics and Genome Biology, The Hospital for Sick Children Research Institute, Toronto, ON M5G 0A4, Canada.
E-mail: zhenya.ivakine@sickkids.ca

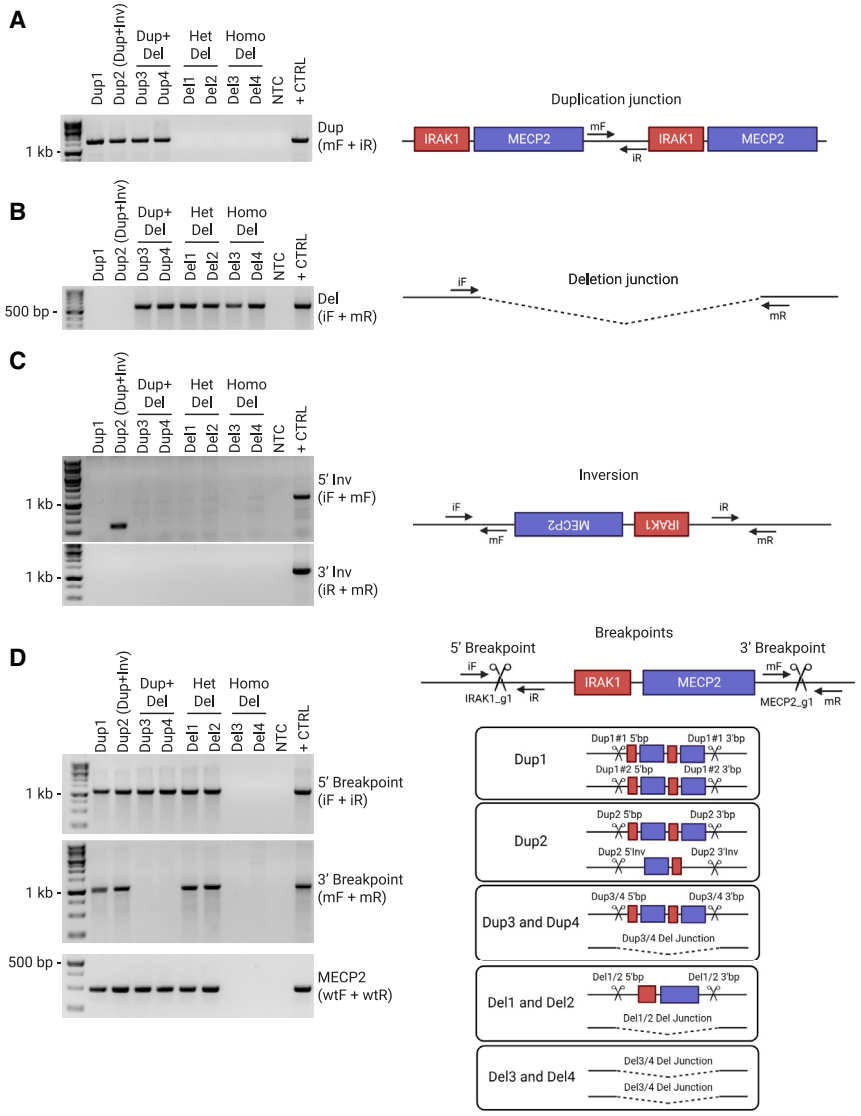


Figure 1. HAP1 MDS model was generated by CRISPR-Cas9 method and identified with a PCR-based strategy

(A) Duplication junction amplified with mF + iR primers.
(B) Deletion junction amplified with iF + mR primers.
(C) 5′ inversion junction amplified with iF + mF primers and 3′ inversion junction amplified with iR + mR primers.
(D) 5′ breakpoints amplified with iF + iR primers and 3′ breakpoints amplified with mF + mR primers. Part of the *MECP2* gene was amplified with wtF + wtR primers. The genotype of each duplication clone is shown on the bottom right.

HEK293T, IRAK1_g1 and MECP2_g1 displayed the highest insertion or deletion (indel) formation (Figure S1B) and were co-transfected into HAP1 cells to generate the disease model. The bulk transfected population was single-cell sorted into individual clones. We isolated gDNA from each clone and utilized a PCR-based method to screen for duplications, deletions, and inversions.

The PCR screening strategy for four potential duplication clones is exemplified in Figure 1. Clones Dup1, Dup2, Dup3, and Dup4 all contained a duplication junction (Figure 1A). Dup3 and Dup4 also harbored a deletion junction on the other chromosome (Figure 1B). For Dup2, an inversion was generated beside the duplication, giving a [Dup+Inv] model (Figure 1C). The genotypes for the clones are illustrated in Figure 1D. The 5′ and 3′ breakpoints were also amplified and sequenced to check whether the Cas9-targeted sites were intact. Dup1 harbored a duplication on both chromosomes so that it had two sets of 5′ and 3′ breakpoints (i.e., Dup1#1 and Dup1#2) (Figure 1D). Dup2 had a wild-type (WT) allele with intact 5′ and 3′ breakpoints (Figure 1D). The other allele had an inversion of the *IRAK1-MECP2* fragment (Figure 1C) so that only one set of the 5′ and 3′ breakpoints was sequenced (Table S1). Similar to Dup2, only one 5′ breakpoint was sequenced in Dup3 and Dup4 because the fragment was deleted on one of the chromosomes (Figure 1B). They both also possessed an intact 5′ breakpoint, but the 3′ breakpoint could not be amplified (Figure 1D). We noticed indel formation on all sequenced breakpoints, including those on the WT allele (Table S1). This demonstrated an instance where the Cas9 nuclease created a double-strand break (DSB) and generated indels at the site without any SVs generated. Occasionally, some breakpoints such as the 3′ end of the inverted fragment in Dup2 (Figure 1C), the 5′ breakpoint in Dup1#2, and the 3′ breakpoint in Dup3 and Dup4 (Figure 1D) were undetectable, indicating large rearrangements that disrupted the primer annealing sites.

We also screened for deletion clones that could serve as controls with none or one copy of *MECP2* and *IRAK1*. Clones Del1, Del2, Del3, and Del4 contained only a deletion junction (Figure 1B) and no duplication or inversion junction (Figures 1A and 1C). Del1 and Del2 had detectable 5′ and 3′ breakpoints and the WT fragment, indicative of a heterozygous deletion (Het Del) clone (Figure 1D), where the deletion only occurred on one chromosome. Del3 and Del4 were homozygous deletion (Homo Del) clones, as indicated by the absence of 5′ and 3′ breakpoints and the WT fragment (Figure 1D).

Out of 307 clones that were screened, deletions occurred most frequently (34%), followed by inversions (21%) and duplications (1%).

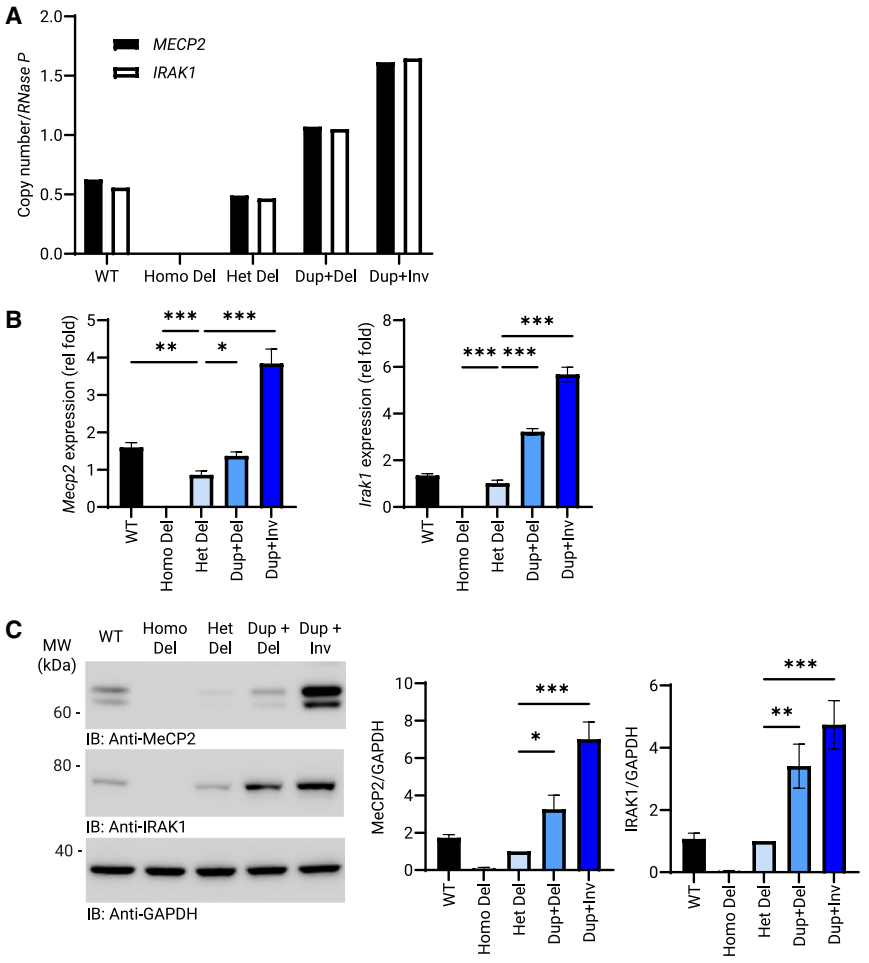


Figure 2. *MECP2* and *IRAK1* expression and protein levels aligned with their copy numbers

(A) Copy-number ratios of *MECP2* (black bar) or *IRAK1* (white bar)/*RNase P* genes in parental WT HAP1, Homo Del, Het Del, [Dup+Del], and [Dup+Inv] clones determined by ddPCR.

(B) *MECP2* and *IRAK1* expressions in parental WT HAP1, Homo Del, Het Del, [Dup+Del], and [Dup+Inv] clones determined by real-time qPCR and normalized to the Het Del clone. WT HAP1 was included for reference (one-way ANOVA, $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$, $n = 4$).

(C) MeCP2 and IRAK1 protein levels in WT HAP1, Homo Del, Het Del, [Dup+Del], and [Dup+Inv] clones determined by western blot and normalized to the Het Del clone. WT HAP1 was included for reference (one-way ANOVA, $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$, $n = 3$).

Copy numbers of MECP2 and IRAK1 in HAP1 clones were confirmed by ddPCR

To validate the genotypes of these clones, droplet digital PCR (ddPCR) was employed to confirm the copy numbers of *MECP2* and *IRAK1* across the entire genome. We confirmed that Del3 and Del4 were Homo Del clones with no copies of *MECP2* and *IRAK1*, Del1 and Del2 were Het del clones with one copy of each gene, Dup3 and Dup4 were [Dup+Del] clones with 2 copies of each gene, and Dup2 was [Dup+Inv] with 3 copies of each gene (Figure 2A).

MECP2 and IRAK1 expression and protein levels correlated with their copy numbers

Given the change in *MECP2* and *IRAK1* copy numbers, their gene expression levels were expected to be altered accordingly. The X chromosome inactivation (XCI) assay revealed that XCI did not occur during diploidization (Figure S2), so *MECP2* and *IRAK1* genes from both alleles could be expressed. We quantified *MECP2* and *IRAK1* gene expression in the generated cell lines and normalized these values to those observed in the diploid Het Del clone. As expected, the expression (Figure 2B) and protein levels (Figure 2C) of *MECP2* and *IRAK1* in these clones were directly proportional to the copy numbers.

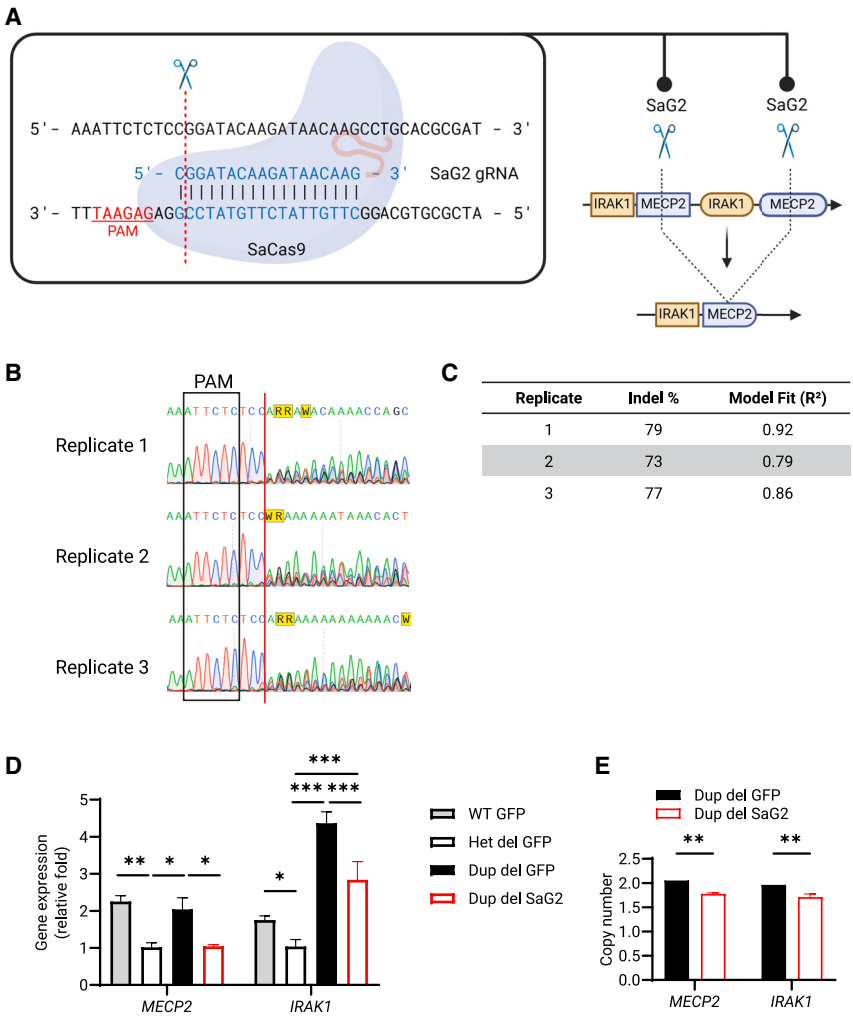
A single-gRNA CRISPR-Cas9 strategy corrected the duplication in the Dup+Del clone

With the new model, we could test a CRISPR-Cas9 method using a single-gRNA[14] to correct the *IRAK1-MECP2* duplication in the [Dup+Del] clone. We designed a gRNA paired with a SaCas9 (SaG2) targeting the intronic region between exons 2 and 3 of *MECP2* to induce deletion by non-homologous end joining (NHEJ) (Figure 3A). SaG2 was transfected into the HAP1 [Dup+Del] clone. Using Inference of CRISPR Edits (ICE) analysis, we determined that SaG2 had an average indel formation efficiency of 76.3% (Figures 3B and 3C).

Real-time qPCR showed that the SaG2 treatment lowered *MECP2* and *IRAK1* expression levels by 49% and 35%, respectively, compared to the [Dup+Del] clones (Figure 3D). Notably, the *IRAK1* expression level was not restored completely to the Het Del level. ddPCR revealed a 13% reduction in the copy numbers of *MECP2* and *IRAK1* (Figure 3E).

DISCUSSION

It has been over a decade since MDS was defined as a neurodevelopmental disease,[8] but we still lack a disease model that faithfully recapitulates the minimal *IRAK1-MECP2* duplication underlying the disorder. For the development of human-sequence-specific genome editing strategies, patient-derived primary fibroblasts[15] and induced pluripotent stem cells (iPSCs)[16] are sources of *in vitro* models. Nonetheless, the variable genomic makeup among patients and the limited availability of patient cells hinder a comprehensive understanding of general disease pathology. In this study, we utilized



Replicate	Indel %	Model Fit (R^2)
1	79	0.92
2	73	0.79
3	77	0.86

Figure 3. A single-gRNA CRISPR-Cas9 approach corrected the duplication in the Dup+Del clone

(A) Schematic diagram showing the single-gRNA CRISPR-Cas9 strategy used for duplication removal in [Dup+Del] clone. Cutting sites are indicated with scissors and dotted lines. The protospacer adjacent motif (PAM) site is underlined and highlighted in red.

(B) Sequencing chromatogram for [Dup+Del] clone treated with SaG2. The PAM site is in a black rectangle, and the cutting site is in a red line.

(C) Indel percentage at SaG2 cutting site in the [Dup+Del] clone treated with SaG2 determined by ICE analysis.

(D) Gene expression levels of *MECP2* and *IRAK1* in parental WT HAP1 cells treated with GFP, Het Del clone treated with GFP, and [Dup+Del] clone treated with GFP or SaG2 determined by real-time qPCR (two-way ANOVA, $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$, $n = 4$).

(E) Copy numbers of *MECP2* and *IRAK1* in [Dup+Del] clone determined by ddPCR (Student's t test, $**p < 0.01$, $n = 4$). Copy numbers were normalized to the *RNase P* gene.

the CRISPR-Cas9 strategy to generate the first genetically engineered human cellular model that harbors the patient-found *IRAK1-MECP2* duplication in diploidized HAP1 cells. HAP1 cells can be an ideal platform to generate models for other diseases involving CNVs, such as proteolipid protein 1 (*PLP1*) duplication in Pelizaeus-Merzbacher disease (PMD)[17] and Duchenne muscular dystrophy (DMD),[14] using the same workflow described in this study. These engineered cell models are easily accessible and permit the characterization of disease pathology and therapy development.

Using a HAP1 MDS model with an *IRAK1-MECP2* tandem head-to-tail duplication, we were able to develop a Cas9 genome editing strategy for duplication removal. Our results showed that the treatment resulted in only a 13% reduction in copy numbers of *MECP2* but also a 49% reduction in *MECP2* expression. One explanation for the lower-than-expected reduction in copy numbers could be the re-insertion of deleted fragments into the genome. SaG2 cuts at the intronic region between exons 2 and 3 of *MECP2* so that even when the fragment is re-integrated, the gene will not be expressed. This may explain the milder reduction in *IRAK1* expression, as the *IRAK1* locus remained intact within the displaced fragment.

Achieving successful duplication removal and correction of transcriptional changes in the HAP1 MDS model paves the way for translating this genome editing strategy to rescue phenotypes in disease animal models. Yu et al. previously demonstrated the feasibility of using CRISPR-Cas9-mediated genome editing to restore normal *Mecp2* expression in the prefrontal cortex, hence partially rescuing the disease phenotypes.[18] Importantly, we utilized the SaCas9 system, which can be packaged into a single adeno-associated virus (AAV) particle along with the gRNA sequence. This strategy holds promising *in vivo* translational potential to target both *MECP2* and *IRAK1* simultaneously.

Alternatively, the protocol can be further optimized to generate disease models in iPSCs, which can be differentiated into various cell types for physiologically relevant, cell-type-specific studies. Hence, this strategy has great potential in the study of disease pathophysiology and the development of therapeutic interventions.

MATERIALS AND METHODS

Materials and methods can be found in the supplemental information.

DATA AND CODE AVAILABILITY

All original data are available from the authors without any restrictions.

ACKNOWLEDGMENTS

We thank the members of the Cohn lab and the Ivakine lab for their input in this study. We thank Dr. Tara Paton and Guillermo Casallo (The Centre for Applied Genomics) for their support with ddPCR. We are thankful to Dr. James Ellis for providing a protocol for the XCI assay and Dr. Stephen Scherer for providing the female iPSC gDNA sample. This work was funded by the Reverse Rett Research Fund. The figures were created in BioRender.com.

AUTHOR CONTRIBUTIONS

Conceptualization, S.Z.R. and E.A.I.; methodology, S.Z.R., W.S.C., E.M., S.S., G.F., and M.K.; investigation, S.Z.R., W.S.C., E.M., S.S., and G.F.; resources, R.D.C. and E.A.I.; data curation, S.Z.R., W.S.C., and E.A.I.; writing – original draft, W.S.C.; writing – reviewing & editing, S.Z.R., W.S.C., G.F., M.K., and E.A.I.; supervision, R.D.C. and E.A.I.; funding acquisition, S.Z.R., R.D.C., and E.A.I. All authors reviewed the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtn.2024.102356.

REFERENCES

1. Sonnenschein, H.A., Lawrence, K.F., Wittenberg, K.A., Slykas, F.A., Dohleman, E.L., Knoublauch, J.B., Fahey, S.M., Marshall, T.M., Marion, J.D., and Bell, J.K. (2018). Suppressor of IKKepsilon forms direct interactions with cytoskeletal proteins, tubulin and α-actinin, linking innate immunity to the cytoskeleton. FEBS Open Bio *8*, 1064–1082. https://doi.org/10.1002/2211-5463.12454.

2. Thielen, A.J.F., van Baarsen, I.M., Jongsma, M.L., Zeerleder, S., Spaapen, R.M., and Wouters, D. (2018). CRISPR/Cas9 generated human CD46, CD55 and CD59 knockout cell lines as a tool for complement research. J. Immunol. Methods *456*, 15–22. https://doi.org/10.1016/j.jim.2018.02.004.

3. Findlay, G.M., Daza, R.M., Martin, B., Zhang, M.D., Leith, A.P., Gasperini, M., Janizek, J.D., Huang, X., Starita, L.M., and Shendure, J. (2018). Accurate classification of BRCA1 variants with saturation genome editing. Nature *562*, 217–222. https://doi.org/10.1038/s41586-018-0461-z.

4. Jia, X., Burugula, B.B., Chen, V., Lemons, R.M., Jayakody, S., Maksutova, M., and Kitzman, J.O. (2021). Massively parallel functional testing of MSH2 missense variants conferring Lynch syndrome risk. Am. J. Hum. Genet. *108*, 163–175. https://doi.org/10.1016/j.ajhg.2020.12.003.

5. Karakus, E., Wannowius, M., Müller, S.F., Leiting, S., Leidolf, R., Noppes, S., Oswald, S., Diener, M., and Geyer, J. (2020). The orphan solute carrier SLC10A7 is a novel negative regulator of intracellular calcium signaling. Sci. Rep. *10*, 7248. https://doi.org/10.1038/s41598-020-64006-3.

6. Llargués-Sistac, G., Bonjoch, L., and Castellvi-Bel, S. (2023). HAP1, a new revolutionary cell model for gene editing using CRISPR-Cas9. Front. Cell Dev. Biol. *11*, 1111488. https://doi.org/10.3389/fcell.2023.1111488.

7. Van Esch, H. (1993). MECP2 Duplication Syndrome. In GeneReviews®, M.P. Adam, J. Feldman, G.M. Mirzaa, R.A. Pagon, S.E. Wallace, L.J. Bean, K.W. Gripp, and A. Amemiya, eds. (University of Washington, Seattle).

8. Ta, D., Downs, J., Baynam, G., Wilson, A., Richmond, P., and Leonard, H. (2022). A brief history of MECP2 duplication syndrome: 20-years of clinical understanding. Orphanet J. Rare Dis. *17*, 131. https://doi.org/10.1186/s13023-022-02278-w.

9. Pascual-Alonso, A., Blasco, L., Vidal, S., Gean, E., Rubio, P., O'Callaghan, M., Martínez-Monseny, A.F., Castells, A.A., Xiol, C., Català, V., et al. (2020). Molecular characterization of Spanish patients with MECP2 duplication syndrome. Clin. Genet. *97*, 610–620. https://doi.org/10.1111/cge.13718.

10. Janssens, S., and Beyaert, R. (2003). Functional diversity and regulation of different interleukin-1 receptor-associated kinase (IRAK) family members. Mol. Cell *11*, 293–302. https://doi.org/10.1016/s1097-2765(03)00053-4.

11. Bauer, M., Kölsch, U., Krüger, R., Unterwalder, N., Hameister, K., Kaiser, F.M., Vignoli, A., Rossi, R., Botella, M.P., Budisteanu, M., et al. (2015). Infectious and immunologic phenotype of MECP2 duplication syndrome. J. Clin. Immunol. *35*, 168–181. https://doi.org/10.1007/s10875-015-0129-5.

12. Pascual-Alonso, A., Xiol, C., Smirnov, D., Kopajtich, R., Prokisch, H., and Armstrong, J. (2024). Multi-omics in MECP2 duplication syndrome patients and carriers. Eur. J. Neurosci. *60*, 4004–4018. https://doi.org/10.22541/au.171276288.82845674/v1.

13. Gottschalk, I., Kölsch, U., Wagner, D.L., Kath, J., Martini, S., Krüger, R., Puel, A., Casanova, J.-L., Jezela-Stanek, A., Rossi, R., et al. (2023). IRAK1 Duplication in MECP2 Duplication Syndrome Does Not Increase Canonical NF-κB–Induced Inflammation. J. Clin. Immunol. *43*, 421–439. https://doi.org/10.1007/s10875-022-01390-7.

14. Maino, E., Wojtal, D., Evagelou, S.L., Farheen, A., Wong, T.W.Y., Lindsay, K., Scott, O., Rizvi, S.Z., Hyatt, E., Rok, M., et al. (2021). Targeted genome editing *in vivo* corrects a Dmd duplication restoring wild-type dystrophin expression. EMBO Mol. Med. *13*, e13228. https://doi.org/10.15252/emmm.202013228.

15. Wojtal, D., Kemaladewi, D.U., Malam, Z., Abdullah, S., Wong, T.W.Y., Hyatt, E., Baghestani, Z., Pereira, S., Stavropoulos, J., Mouly, V., et al. (2016). Spell Checking Nature: Versatility of CRISPR/Cas9 for Developing Treatments for Inherited Disorders. Am. J. Hum. Genet. *98*, 90–101. https://doi.org/10.1016/j.ajhg.2015.11.012.

16. Nageshappa, S., Carromeu, C., Trujillo, C.A., Mesci, P., Espuny-Camacho, I., Pasciuto, E., Vanderhaeghen, P., Verfaillie, C.M., Raitano, S., Kumar, A., et al. (2016). Altered neuronal network and rescue in a human MECP2 duplication model. Mol. Psychiatry *21*, 178–188. https://doi.org/10.1038/mp.2015.128.

17. Hodes, M.E., Woodward, K., Spinner, N.B., Emanuel, B.S., Enrico-Simon, A., Kamholz, J., Stambolian, D., Zackai, E.H., Pratt, V.M., Thomas, I.T., et al. (2000). Additional copies of the proteolipid protein gene causing Pelizaeus-Merzbacher disease arise by separate integration into the X chromosome. Am. J. Hum. Genet. *67*, 14–22. https://doi.org/10.1086/302965.

18. Yu, B., Yuan, B., Dai, J.-K., Cheng, T.L., Xia, S.-N., He, L.-J., Yuan, Y.-T., Zhang, Y.-F., Xu, H.-T., Xu, F.-Q., et al. (2020). Reversal of Social Recognition Deficit in Adult Mice with MECP2 Duplication via Normalization of MeCP2 in the Medial Prefrontal Cortex. Neurosci. Bull. *36*, 570–584. https://doi.org/10.1007/s12264-020-00467-w.