

Mutations of the Histone Linker H1-4: An Expanded Cohort and Functional Characterization of Frameshift Mutant H1.4 in Neurons

Martine W Tremblay

Duke University <https://orcid.org/0000-0002-0066-964X>

Matthew V Green

Duke University

Jill A Rosenfeld

Baylor College of Medicine

Haley Streff

Baylor College of Medicine

William Craigen

Baylor College of Medicine

Nasim Bekheirnia

Baylor College of Medicine

Saeed Al Tala

Armed Forces Hospital SR, Khamis Mushayt

Wendy D Tan

Duke University

Anne E West

Duke University

Yong-hui Jiang (✉ yong-hui.jiang@yale.edu)

Department of Genetics <https://orcid.org/0000-0001-6188-8806>

Research

Keywords: Rahman syndrome (RMNS), HIST1H1E, H1-4, H1.4, linker histone, intellectual disability (ID)

Posted Date: December 31st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-135774/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Rahman syndrome (RMNS) is a rare genetic disorder characterized by mild to severe intellectual disability, hypotonia, anxiety, autism spectrum disorder, vision problems, brittle bones, and dysmorphic facies. *De novo* heterozygous mutations in *H1-4* (*HIST1H1E*) encoding the linker histone H1.4 are found in patients with RMNS; however, the underlying mechanisms causing the pronounced neurological manifestations are not understood. The majority of reported mutations in *H1-4* are small insertions or deletions that create a shared frameshift, resulting in an H1.4 protein that is both truncated and possessing an abnormal C-terminal tail.

Methods: Seven Rahman syndrome subjects with C-terminal frameshift mutations as well as three patients with heterozygous null mutations in *H1-4* were described in detail. Lymphoblastoid cells from these patients were used to identify transcriptional abnormalities in RMNS. Wildtype or mutant frameshifted human H1.4 protein was exogenously expressed in primary rat hippocampal neurons, and neuronal structure and function were assessed using immunohistochemistry and multi-electrode array recordings.

Results: Individuals with heterozygous null variants in the *H1-4* gene lack several key RMNS phenotypes, supporting the hypothesis that RMNS is due to a gain-of-function of the frameshift mutant H1.4 protein. In cultured rat hippocampal neurons, H1.4 was localized to the nucleus, though the frameshift mutant H1.4 had a distinct subnuclear distribution and enlarged nuclei. Overexpression of frameshift mutant H1.4 had minimal effects on dendritic morphology; however, it significantly reduced neuronal firing rate relative to neurons overexpressing wildtype human H1.4.

Limitations: Given the small number of RMNS cases worldwide, the true breadth of phenotypes remains unknown. Though our data do not show robust differential expression of any genes, larger cohorts for clinical and molecular studies will be needed to gain reliable data.

Conclusions: These data are the first to characterize the consequence of frameshift mutant H1.4 in neurons. These data provide new insights into the breadth of phenotypes and causes of neurological dysfunction in RMNS and highlight the need for future studies on the function of histone H1.4 in neurons.

Background

Rahman syndrome (RMNS, also referred to as HIST1H1E syndrome, MIM #617537) is a syndromic form of intellectual disability (ID). Symptoms of RMNS can vary but include mild to severe ID, hypotonia in newborns and infants, anxiety, autism spectrum disorder (ASD), vision problems, brittle bones and teeth, and dysmorphic facial features (ie: high hairline, wide-set eyes, and low-set ears) (1-7). A total of about 40 patients have been reported worldwide since the disorder was originally characterized (1). The full spectrum of clinical presentations at different ages, specifically in adults, remains to be delineated.

RMNS is caused by *de novo* pathogenic variants in the *H1-4* gene (previous gene symbol *HIST1H1E*, OMIM:142220, HGNC: 4718), which encodes the linker histone H1.4. Most pathogenic variants identified are insertions or deletions (indels) that alter the codon reading frame by the addition of one base pair, creating truncated proteins that not only lack the normal C-terminal domain but also contain a shared, novel C-terminus of approximately 40 amino acids (1, 3, 5). These data raise the possibility that cellular pathophysiology in RMNS could arise either from loss-of- WT H1.4 function and/or from dominant negative or gain-of-function effects of the mutant protein.

Histones function in the nucleus to organize genomic DNA, with DNA wrapping around octamers of the core histones (H2A, H2B H3, and H4) in a repeating structure called the nucleosome. The H1 histone linker proteins flank nucleosomes at the DNA entry and exit sites, where they are thought to play a role in chromatin compaction (8). The positive charge of the histone H1 allows the protein to neutralize and bind to negatively charged DNA (9). The shared frameshift of *H1-4* seen in RMNS creates an aberrant H1.4 tail that has a net negative charge, potentially disrupting this interaction (3). However, the functional importance of this charge shielding for chromatin structure or gene regulation is not understood.

H1-4 is just one member of the histone H1 family, which contains 11 members in humans. The histone H1 family includes both germ cell-specific and somatic subtypes, which are expressed in nearly all cells, including neurons (10). The somatic histone H1s are divided into replication-dependent (*H1-1* through *-5*) and replication-independent (*H1-0*, *H1-x*) subtypes based on their preferential expression and incorporation into chromatin during specific phases of the cell cycle (11). Histone H1.4 is a 219 amino acid protein that shares >80% amino acid identity with each of the four other replication-dependent histone H1s (9). The high homology among the H1 family suggests that functional compensation between H1 family members may explain why knockout mice lacking single H1s, including knockout of the mouse *H1f4* gene, fail to show apparent developmental phenotypes (12).

The detailed cellular functions of the individual H1 histone subtypes are poorly understood. Knockout studies have implicated the histone H1 proteins in a number of gene regulatory processes including DNA methylation, chromatin compaction, and regulation of repressive histone modifications (11, 13). Consistent with possible disruption of chromatin regulation in RMNS, one group reported a distinct methylation profile in peripheral blood samples of a cohort of affected individuals compared with healthy controls (14). However, whether frameshift mutations of *H1-4* result in select neurodevelopmental phenotypes including ASD and ID via effects on chromatin, and whether frameshift mutant H1.4 acts in neurons to disrupt brain function remain entirely unknown.

To better understand the pathophysiology of frameshift mutant H1.4, we gathered a cohort of 9 individuals (7 RMNS, 2 loss-of-function), established immortalized lymphoblastoid cell lines from 3 of these RMNS patients to assess regulation of gene expression, and investigated the cellular consequences of exogenous wildtype (WT) and frameshift mutant H1.4 expression in primary rat neurons. Our patient data expand the clinical understanding of the relationship between frameshifted mutant H1.4 and ID. Most importantly, despite the fact that prior studies have focused on identifying functions of H1.4 in

dividing cells, our data reveal that expression of frameshift mutant H1.4 has significant effects on the physiology of postmitotic neurons, raising the possibility that disturbed function of H1.4 in mature neurons of the brain may contribute to neurological deficits in patients with RMNS.

Methods

Subjects: This study was approved by the Duke University Institutional Review Board and the Yale University School of Medicine. Written consent was obtained from all subjects. Clinical exome sequencing (ES) reports and medical records (including neuropsychological reports) were collected from individuals. Families also completed a brief questionnaire regarding clinical features and medical issues associated with subjects. For each subject, mutations of *H1-4* reported by clinical genetics diagnostic laboratories were reviewed and confirmed by a clinical geneticist.

***In Silico* Frameshift Predictions in H1 Genes:** Human mRNA sequences for each H1 family member (*H1-1* through *-5*) were obtained from Gencode GRCh38/hg38 Assembly. R programming software was used to model single base pair duplications and deletions into the *H1-1* through *-5* coding regions. Please see <https://github.com/WestLabDuke/H1.4-repository> for source code and more information.

RMNS Patient Derived Lymphoblastoid Cell Lines: For subjects 2, 3, and 5, whole blood samples were collected, and lymphocytes were isolated. The immortalized lymphoblast cell were generated by a standard protocol using Epstein Barr Virus (EBV) (15).

Control Lymphoblastoid Cell Lines: Control EBV immortalized B lymphoblast cell lines (LCLs) ND23648, ND23960, AG09393 were obtained from Coriell Cell Repositories. These healthy control lines were sex, age, race, and country of origin matched as closely as possible to the subject derived LCLs. RMNS and control cell lines were cultured in RPMI with L-glut plus 10% FBS in parallel for 3 passages before collection of samples for molecular experiments.

Genotyping: All LCLs were Sanger sequenced (Eton Bioscience, Research Triangle Park, NC) to confirm genotype of the *H1-4* gene. Primers are listed in **Supplemental Table 2**.

Reverse Transcription and Quantitative PCR (qRT-PCR): RNA from cell pellets was isolated using Trizol Reagent (Invitrogen cat# 15596026). RNA was treated with DNase (NEB cat# M0303) prior to cDNA synthesis (Invitrogen SuperScript III cat#18080051) per manufacturers protocols. qRT-PCR was performed on a QuantStudio3 machine (ThermoFisher Scientific) using SYBR green master mix (ABI cat#4667659). Primers for qRT-PCR are listed in **Supplemental Table 1**. For all genes except *H1-2* through *-4* (single exon), primers were designed to be intron spanning and to encompass the most conserved exons (retained in most isoforms) per the UCSC human genome browser GRCh38/hg38 assembly (16). Presence of a single amplicon was confirmed by running qRT-PCR products on agarose gel. Primers used for investigating expression levels of additional H1 family member transcripts were validated by knocking down H1s by siRNA mediated knockdown (Sigma Mission esiRNA: (*H1-2*) EHU014531 (*H1-3*) EHU149861

(*H1-4*) EHU038891) in HEK293T cells using Lipofectamine RNAiMAX (ThermoFisher 13778100) per manufacturer's instructions.

Exogenous lentiviral expression of wildtype and frameshift mutant H1.4: We generated a dual FLAG-myc-tagged wildtype and c.430dupG frameshift mutant *H1-4* cloned into a modified FUGW vector (Addgene plasmid #14883 (17)). The FLAG-myc tag was placed at the N-terminus in frame with the ATG, consistent with constructs used in prior studies (18-20) as not to disrupt the C-terminus. A GFP expressing plasmid was used as a control (Addgene cat#53188). Lentivirus was produced in HEK293T cells using lentiviral packing plasmids p-CMV-VSV-G (Addgene plasmid #8454) and pCMV-dR8.2dvpr (Addgene plasmid #8455(21)). Briefly, lentiviral packing and overexpression plasmids were transfected into 293T cells using calcium phosphate transfection. Supernatant was collected three days post transfection and virus was concentrated by ultracentrifugation. For LCL experiments, cells were transduced using a spinoculation protocol (22). Transduced cells were maintained for two weeks before being collected for experiments. For neuron experiments, viruses were functionally titered in primary rat hippocampal neuron cultures. Rat neurons were infected on DIV1 at a multiplicity of infection of 1 in Basal Medium Eagle (BME) (Sigma B1522) with 0.4 µg/mL added polybrene.

Western blot: Cells were lysed in sodium dodecyl sulfate plus beta-mercaptoethanol, boiled for 10min to denature, and run on a 12% polyacrylamide gel. Blots were transferred for 1.5hrs at 300mA constant on to a nitrocellulose membrane. Membranes were blocked for 1hr in 5% milk/TBS+1% Tween, probed with primary antibody (mouse anti-Flag (M1), Sigma F3165, 1:1000 or mouse anti-actin, Millipore MAB1501, 1:10,000) overnight in blocking solution, rinsed, and incubated with infrared conjugated secondary antibodies (Biotium) for 2hr at room temp. Blots were imaged using the Li-Cor Odyssey (9120) system.

Primary rat hippocampal neuron culture: Hippocampi of CD IGS E18.5 rat embryos (Charles River Laboratories) were dissected as previously described (23). Both male and female embryos were used for all experiments, and all experiments were conducted in accordance with an animal protocol approved by the Duke University Institutional Animal Care and Use Committee. For cultures, briefly, hippocampi were removed and immediately placed in cold Ca²⁺ and Mg²⁺-free HEPES-buffered solution (ThermoFisher Cat:14175-095). Tissue was then trypsinized in 3 ml TrypLE Express (ThermoFisher Cat:12604013) for 15 min at 37°C. Trypsin-containing buffer was then replaced with Basal Medium Eagle (BME) (Sigma B1522) supplemented with Glutamax (ThermoFisher Cat:35050061), 10% fetal bovine serum, and penicillin/streptomycin, and triturated using flame-narrowed Pasteur pipettes. Cells were plated at a density of 120k/well on coverslips coated with poly-D-lysine (Sigma P7280) and laminin (Sigma L2020). On day in vitro (DIV) 1, the media was exchanged with Neurobasal medium (ThermoFisher Cat:21103049) supplemented with B-27 (ThermoFisher Cat:17504044), Glutamax, and penicillin/streptomycin.

Immunocytochemistry (ICC): Primary rat hippocampal neurons were cultured on coverslips coated with poly-d-lysine and laminin. At DIV 7 or DIV14, cells were fixed in 4% PFA for 10 min at room temperature (RT) and rinsed with PBS. Neurons were permeabilized in PBS+0.2% Triton-X and blocked in 10% Goat

serum prior to incubation with primary antibody (chicken anti-MAP2 1:3000 Millipore AB5543; mouse anti-Myc 1:200 Sigma M4439) overnight at 4C. Cells were then washed and incubated with fluorescent secondary antibody (Invitrogen) for one hour at RT. Coverslips were then counterstained with Hoechst and mounted on slides. Images were captured on a confocal microscope (Leica SP8).

Nuclear Morphology Analysis and Sholl Analysis: Microscopy images were blinded before all analysis. CellProfiler (24) was used for nuclear morphology analysis to determine overlap of myc and Hoechst signal in each nuclei and cell size. Representative images of nuclear distribution were produced using FIJI ImageJ (25) by quantifying intensity across the diameter of each nucleus. For Sholl analysis, neuron projections were traced using FIJI ImageJ and analyzed using the Sholl plugin.

Multi-electrode Array (MEA) Recordings: E18.5 CD IGS rat hippocampal neurons were plated in a 48-well MEA plate (Lumos 48, Axion Biosystems) at a density of 60,000 cells/well. Wells were coated with poly-d-lysine and laminin prior to plating. Cells were infected on DIV 1 and half-media changes were performed at DIV 10 and 13 with Neurobasal culture medium supplemented with B-27, Glutamax, and penicillin/streptomycin. Extracellular recordings were performed at 37C with 5% CO₂ using the Maestro MEA system with AxIS software (Axion Biosystems). Data were acquired at 12.5 KHz and was filtered with a Butterworth bandpass filter of 200 Hz-3 kHz. Spike detection was performed to detect action potentials with a 6x standard deviation. Electrodes were considered active if they displayed more than 5 spikes/min. Recordings were collected for 5 minutes every day from DIV 8-14 and data were analyzed for spike frequency and synchrony index using Axion Biosystem's Neural Metrics Tool.

Data and Statistical Analysis: All data are expressed as mean \pm SD. Analyses were performed with Prism 8 statistical package (Graphpad, San Diego, CA). Figure legends contain experiment numbers (N), group sizes (n), and statistical test performed for each experiment. The experimenters were blind to sample genotype and treatment prior to data collection and analysis.

Results

Genetic and clinical profiles of subjects with *H1-4* mutations

Seven individuals with RMNS C-terminal frameshift mutations and three individuals with loss-of-function of wild type (WT) *H1-4* were recruited for this study (**Figure 1**) (see **Supplemental Table 1** for histone H1 family gene nomenclature). Similar to other reports, most variants in the *H1-4* gene were clustered in a 100bp region between c.360 and c.450 in the C-terminal domain of the transcript and result in long frameshift mutations predicted to have a strongly negative net charge compared to WT H1.4 protein. Six of these mutations are recurrent and have been reported by other groups, but four mutations (c.392dupC in subject #4, c.1A>G in subject #8, c.100_101insT #9, c.265delA #10) are newly reported here. The duplication variant in subject #4 is consistent with a RMNS associated frameshift mutation. By contrast, the c.1A>G variant is novel and predicted to disrupt the ATG start codon of *H1-4* and result in loss-of-function, as the alternative protein product from an alternative ORF is a small 17 amino acid peptide.

Similarly, c.100_101insT results in a peptide of 34 WT amino acids followed by a frameshift of 13 amino acids (p.K34ifs*13), and c.265delA results in p.S89Afs*140, both mutations result in a loss-of-function of the WT C-terminal tail.

The clinical features we collected from a review of medical records and parent-completed questionnaires are summarized in **Figure 2A** and **Supplemental File 1**. RMNS subjects with *H1-4* mutations were originally characterized as having a syndrome of “Intellectual disability and overgrowth”(1). However, except for two subjects with macrocephaly in our cohort, we did not find a significant deviation above normal growth parameters for newborns nor in current measures of head circumference and height, per Centers of Disease Control and Prevention (CDC) clinical growth charts (26) (**Figure 2B**). The major features of RMNS subjects carrying C-terminal frameshift mutations included dysmorphic facies (7/7, 100%), ID (7/7, 100%); ASD (3/4, 75%); motor delay (7/7, 100%); vision issue (6/6, 100%). Overall, the features of these subjects in our cohort are consistent with previous reports with large RMNS cohorts (2, 5).

Summary of variants in *H1-4*

To better understand the array of *H1-4* mutations associated with an ASD or ID phenotype, we conducted a systemic review of the Baylor Genetics database of ~17,000 individuals undergoing clinical exome sequencing (ES). This revealed 15 variants in the UTR, 11 in frame amino acid deletions, 403 missense mutations, and 94 silent mutations in addition to the identified *H1-4* frameshift indels described in **Figure 1**. Among individuals undergoing clinical chromosomal microarray analysis, we also identified three individuals with phenotypes of developmental delay or autism who had large CNV gains of 565-602kb in regions including the *H1-4* gene.

The contribution of haploinsufficiency resulting from heterozygous *H1-4* mutations has not previously been described. The lack of *H1-4* CNV losses in this clinical database is consistent with the dominant negative/gain-of-function hypothesis for the frameshift mutant H1.4 protein in RMNS, rather than a loss-of-function.

Identification of novel loss-of-function *H1-4* alleles

The *de novo* c.1A>T variant in subject #8 from our cohort is predicted to disrupt H1.4 protein expression as it occurs in the start codon and there is no alternative ATG downstream within the same reading frame. In the unlikely event that an alternative reading frame is utilized, the only alternative methionine results in a small peptide of 17 amino acids (MTWRRTTAASSWVSRAW in the +1 codon reading frame shift). The clinical features of the subject with the c.1A>T variant are summarized in **Figure 2A (and Supplemental File 1)**. This individual has dysmorphic features including broad nasal tip, thick alae nasi, accentuated Cupid's bow, large and fleshy earlobes bilaterally, and small mouth leading to mild dental crowding. Together, this set of dysmorphic features is distinct from those in individuals carrying C-terminal frameshift mutations in *H1-4* and diagnosed with RMNS. Other minor anomalies include bilateral fifth finger clinodactyly and fetal finger pads. Her language development was delayed as she

spoke her first word at 18 months, two words together at 3 years, and short sentences at 4 years. She had a clinical diagnosis of ASD based on DSM-5 criteria due to her deficits in social communication and interaction, as well as repetitive behaviors. She has family history of ASD, with a sister who also has ASD and does not carry a *H1-4* variant. In addition, she has flat affect, impairments in social-emotional reciprocity and nonverbal communication, and hyperreactivity to sensory input. In contrast, she did not have history of hypotonia or motor delay, as she walked at 13 months. She also did not have vision problems, bone concerns, ADHD, or anxiety. While this individual is a single case, her features are overlapping but clearly distinct clinically from the subjects with frameshift mutations in the C-terminus of *H1-4*.

The clinical presentation of subject #9 with a maternally inherited c.100_101insT mutation in *H1-4* is also distinct from RMNS phenotypes. In the event that this transcript eludes nonsense mediated decay, the protein product is a peptide of 47 amino acids (p.K34Ifs*13), likely resulting in a loss-of-function of WT H1.4 C-terminus. This individual presents with chronic kidney disease, of which the cause is homozygous *NPHP1* deletion [OMIM: 607100] identified by CMA comprehensive. However, this individual also presents with a history of delayed speech and language development (he did not speak until he was 4 years old) for which he received speech therapy. He did not have any other delays in childhood and was formally screened for developmental delay, of which he did not meet diagnostic criteria for any other delays or disorders. This subject does not have ID and has reached academic milestones as expected with typical development. As this variant is maternally inherited, we briefly investigated the phenotype of the mother. The mother of subject #9 did not have any delays during her development and of note, was enrolled in gifted programs. Neither subject #9 nor his mother have dysmorphic facial features.

We also identified an individual with a C-terminal tail distinct from the tail found in RMNS subjects. Subject #10 carries a *de novo* c.265delA which results in p.S89Afs140* if the transcript escapes nonsense mediated decay. While the predicted protein product contains a long tail, this frameshift tail is in the alternative reading frame from that of RMNS subjects. This individual presents in the clinic with short stature, ankylosis of both jaw joints (resolved by surgical release), retromicrognathia, and bilateral conductive hearing loss. With regards to neurological development, he had a speech delay early in life. However, he currently speaks in broken sentences, and has a below average IQ test score of 86. However, he has not met diagnostic criteria for ID or ASD. He did not have infantile hypotonia and was able to walk without support at 12 months of age. In addition, his facies are distinct from RMNS individuals in that the sole dysmorphic feature is the locked jaw. He has a family history of global developmental delay, speech delay, motor delay, and depression presenting in two male siblings who we previously reported to have a different disease caused by homozygosity for a *THUMPD1* variant (27), which was not identified in the patient presented here.

The phenotype of these individuals further supports the hypothesis that the frameshift mutations clustered in the C-terminus of *H1-4* operate via a gain-of-function mechanism, rather than through *H1-4* haploinsufficiency or loss-of-function of the H1.4 C-terminal tail.

Conserved frameshifts and net charge changes in H1 family

The cluster of RMNS associated C-terminal frameshift mutations in *H1-4* raises the question of why similar disease-associated mutations are not found in other Histone H1 family members. To investigate the consequence of indels in *H1-1* through *-5*, we computationally introduced 1 base pair indels throughout the length of the *H1-1*, *H1-2*, *H1-3*, *H1-4*, and *H1-5* gene sequences and determined the consequence of these frameshift mutations on the both the length of the protein translated and the change in charge of the resulting peptide compared with the WT protein (**Figure 3**). We found that all histone H1 family members could produce out-of-frame proteins of various sizes if indels were introduced at any one of a number of positions (**Supplemental Figure 3**). However, the frameshift mutations in the *H1-4* RMNS hotspot were unique for their ability to result in peptides with a large negative change in the charge of the C-terminal tail (**Figure 3**). We also observed that insertions in a similar region of *H1-5* could have a smaller but potentially significant effect on protein charge. However, when we queried the clinical ES database at Baylor College of Medicine (one of the largest and non-public clinical ES databases) for frameshift mutations in *H1-5* which were associated with an ID phenotype, we found none. There were 9 individuals with frameshifts in *H1-1* (3 of these mutations were inherited, 5 were not tested for origin), and two individuals with frameshift mutations in *H1-2* (**Supplemental File 2**), however these mutations had minimal effects on net protein charge. Taken together these data are consistent with the hypothesis first proposed by Tatton-Brown et al. (2017) that pathogenic mutations in *H1-4* are those that disrupt the positively charged C-terminal tail, impairing the ability of this H1 histone to interact with negatively charged DNA (11, 13).

Transcriptional characterization of RMNS patient-derived cells

Although H1.4 is a histone linker protein, its exact transcriptional functions remain poorly understood. To determine whether there were gene expression changes in our *H1-4* frameshift mutant cohort, we generated immortalized LCLs from three subjects and compared RNA expression in these cells to LCLs from age, sex, and ethnically matched controls (**Figure 4A, B**). Prior studies from knockout mice have shown that loss of a single histone H1 gene can be associated with compensatory upregulation of other histone H1 family members (12), thus we first assessed relative expression of H1 family genes between the LCL lines. We validated primers to detect *H1-2*, *H1-3*, and *H1-4* in human cells (**Supplemental Figure 2**) and used these to measure H1 family gene expression. We did not find any significant differences in *H1-2*, *H1-3* or *H1-4* expression in *H1-4* mutant LCLs compared with control (**Figure 4C**). These data show that the RMNS-associated *H1-4* frameshift mutations in our cohort do not induce compensation via transcriptional induction of other H1 family members.

We then took an approach to examine the expression of candidate genes suggested from previous studies to be associated with H1.4 function. These candidate genes fall into four classes. First, we assessed mRNA encoding proteins that have been shown to physically interact with H1.4 in mouse cells (18-20, 28-30) (**Figure 5A**). Second, we examined the expression of epigenetic readers, writers, or erasers of DNA methylation, which is a process that has been reported to be regulated by H1 and aberrant in *H1-4*

frameshift mutant fibroblasts obtained from RMNS individuals (14, 31) (**Figure 5B**). Third, we looked at the expression of components of the complement system, because one complement factor, C3, was previously identified as being differentially methylated in *H1-4* mutant skin fibroblasts (14) (**Figure 5C**). Alterations in the complement system would be of particular interest in ID because this pathway has been implicated in synaptic pruning in neurons (32) and linked to neurodevelopmental disorders such as ASD (33-35), thus finally we also examined the expression of genes that link complement cascades to synapses (**Figure 5D**).

The only genes we observed to be significantly different between patient and control LCLs were the complement components *C3*, *CR1*, and *CR2* ($p=0.0439$, 0.0021 , 0.0130 respectively) (**Figure 5C**). However, because LCLs are oligo-clonal cell lines derived from cells of the immune system, it is possible that differences in complement component expression between these lines could reflect clonal responses to serum factors in the media rather than a specific effect of the frameshift mutant H1.4 expression. We reasoned that if complement C3 was dysregulated in the patient LCLs due to the expression of mutant H1.4, then if we recreated the same H1.4 expression pattern in control cells by expressing either WT or frameshift mutant H1.4, we should generate the same difference in complement gene regulation. Similar to constructs used in prior studies(18-20) we generated N-terminal dual FLAG-myc-tagged lentiviral constructs to exogenously express either WT or frameshift mutant (c.430dupG) H1.4 in the WT control LCL cells. Western blotting against the FLAG tag in LCL lysates verified exogenous expression of the WT and mutant FLAG-myc-tagged H1.4 protein in the control lines (**Supplemental Figure 3**). Nonetheless, we found no difference in transcript abundance of *C3*, *CR1*, or *CR2* between any of the conditions in the three control lines ($p=0.449$, 0.265 , 0.286 respectively) (**Figure 5E**). These data suggest that exogenous expression of frameshift mutant H1.4 alone is not sufficient to drive upregulation of complement expression. Future studies will be required to determine whether other genetic factors interact with mutant H1.4 to alter complement component expression in LCLs from RMNS patients.

Altered distribution of H1.4 mutant protein in nuclei of rat hippocampal neurons

Frameshift mutations in *H1-4* are strongly associated with neurodevelopmental impairments in humans. However, the consequences of frameshifted mutant H1.4 expression on the structure or function of neurons is unknown. Importantly, there is limited evidence of consistently prevalent and conserved gross structural malformations of brains of these individuals (**Supplemental File 1** and (2, 5)), which raises the possibility that expression of frameshift mutant H1.4 may disrupt functional aspects of neuronal physiology. To elucidate the possible impact of frameshift mutant H1.4 protein on neurons, we compared the consequences of overexpressing either WT or mutant H1.4 protein in cultured primary embryonic rat neurons (**Figure 6A**). Immunostaining of the Myc-tag of the dual FLAG-myc-H1.4 constructs showed that most cells expressed the exogenous human proteins (**Figure 6B**). Thus, our model mimics the germline heterozygous expression of *H1-4* frameshift mutant protein in human RMNS. Consistent with previous experiments in other cell types (5), we found that both WT H1.4 and frameshift mutant H1.4 protein were detected exclusively in the nucleus. (**Figure 6B**). However, the percent overlap between H1.4 and Hoescht was lower in neurons expressing the frameshift mutant H1.4 (**Figure 6C**) suggesting there was a

difference either in nuclear morphology or in the distribution of mutant histone H1.4 compared with WT within the nucleus. Indeed, the nuclear size of neurons expressing H1.4 mutant protein is significantly larger than that of WT ($p=0.0015$) (**Figure 6D**). Yet even controlling for size, we found that the H1.4 frameshift mutant protein has a greater localization towards the periphery of the nucleus compared with the distribution of the WT protein (**Figure 6E**). These results indicate a potential large-scale change in chromatin or nuclear organization associated with the expression of mutant H1.4 protein.

Frameshift mutant H1.4 disturbs action potential frequency and synchrony in hippocampal neurons

To determine if the expression of H1.4 frameshift mutant protein results in functional changes in neuronal physiology, we first assessed dendritic morphology, which determines key aspects of synapse formation between neurons. We did not find any difference in length and number of primary or secondary neurites at DIV7 between neurons expressing either exogenous WT or frameshift mutant H1.4 protein (**Figure 6F**). In addition, Sholl analysis performed on this data did not note a difference in number of crossings of cell neurites for the neurons expressing H1.4 mutant protein compared with neurons expressing WT protein ($p=0.2344$) (**Figure 6G**).

To determine if there is a functional consequence of H1.4 frameshift mutant protein for neural network activity, we measured action potential spike frequency and synchrony in cultured hippocampal neurons plated on 16 electrode multielectrode arrays (**Figure 7A**). We detected spiking with similar waveforms regardless of whether neurons were infected with the WT H1.4, frameshift mutant H1.4, or GFP (**Figure 7B, C**). For neurons expressing either WT H1.4 or GFP as control, both firing rate detected at a single electrode and the synchrony of firing between electrodes increased from DIV 8-13, which are consistent with increased connectivity and synaptic maturation over this period (**Figure 7D**). However, neurons expressing H1.4 frameshift mutant protein showed lower levels of firing (mixed effects, $p=0.0006$) with a delayed rise to later days in culture and significantly reduced synchrony (mixed effects, $p=0.0017$) when compared to WT H1.4 and GFP groups (**Figure 7D**). Thus, these data provide the first evidence that frameshift mutant H1.4 protein disrupts neuronal activity

Discussion

In this study we report *H1-4* mutations and clinical phenotypes for several new subjects, furthering understanding of the genetics of RMNS. We characterize gene expression in LCLs derived from these subjects and use them to characterize the expression of genes that could contribute to disease pathophysiology. Finally, we show that expression of frameshift mutant H1.4 in neurons disrupts neuronal firing, providing a novel insight into the underlying causes of ID in RMNS. Taken together these data significantly expand our understanding of how mutations in *H1-4* lead to RMNS.

The *H1-4* mutations identified in RMNS are frameshift mutations, all of which encode a similar long novel C-terminal tail before terminating. Whereas the WT C-terminal domain of H1.4 contains a high density of positively charged amino acids, the frameshift significantly reduces the overall charge of the truncated protein (1). Our systematic modeling of frameshift mutations across the closely related *H1-1* through -5

genes confirmed that the RMNS associated frameshift mutations in *H1-4* are unique in their impact on the charge of the resulting truncated proteins (**Figure 3**). This may also explain why similar frameshift variants in other histone H1 genes have not been implicated as causative in other human genetic disorders, despite the evidence from mouse knockout studies that suggest functions of this sub-family of histone H1 proteins are substantially overlapping (15). The positive charge of the histone H1 proteins is thought to be important for shielding the negative charge of the linker DNA at the entry and exit points of nucleosomes (11). Our data further support the hypothesis that impairment of this charge shielding contributes to the pathophysiology of *H1-4* frameshift mutations, though the mechanisms by which linker DNA regulation contributes to cellular physiology remain poorly understood.

The mutant frameshift H1.4 is hypothesized to act via a gain-of-function mechanism given that the mutant transcripts escape nonsense mediated decay and encode an aberrant protein that disrupts cellular physiology even in the presence of WT H1.4 expressed from the other allele. In this study, we report three loss-of-function *H1-4* mutations from a clinical ES cohort. These subjects (#8, #9, #10, **Figure 2A**) carry a *de novo* c.1A>T variant, a maternally inherited c.100insT variant, and a *de novo* c.265delA variant, respectfully, that are predicted to lead to haploinsufficiency. Subjects #9 and #10 do not have ID, ASD, RMNS related dysmorphic facial features, motor delay, or hypotonia in infancy. Although subject #8 was notable for the fact that she met clinical criteria for diagnosis with ASD and ID, it should be noted that she also has a sister also affected with ASD. Similarly, subject #10 also had a family history of depression, global developmental delay, speech delay, and motor delay. Therefore, it is possible that the conditions found in these individuals is due to another undetected or characterized pathogenic variant elsewhere in the genome. Interestingly we observed that the phenotype of these three loss-of-function subjects lack some common features of individuals carrying RMNS associated *H1-4* frameshift mutations, suggesting that haploinsufficiency alone and loss of the WT C-terminal tail is not sufficient to lead to the spectrum of full RMNS phenotypes. This data is supported by existing mouse data, in which knockout of the mouse *H1f4* gene does not result in abnormal phenotype during development or adult animals (12). Identification of other loss-of-function mutations in *H1-4* and their correlation with clinical phenotypes will be needed to further explore this possibility.

Despite strong genetic evidence linking frameshift mutant *H1-4* with ID, little is known about how disturbing the normal function of this protein would lead to cellular dysfunction and disease. The histone H1 proteins organize chromatin structure in genomic DNA and are presumed to play a role in coordinating transcription (36). Consistent with the frameshift mutations in H1.4 disrupting chromatin structure, we found that exogenous expression of frameshift mutant H1.4 in primary rat hippocampal neurons increased nuclear size (**Figure 6D**). Furthermore, we saw a distinction in the distribution of WT and frameshift mutant H1.4 in the nuclei of these neurons, with WT H1.4 showing a uniform distribution, whereas frameshift mutant H1.4 was distributed toward the outer edges of the nucleus (**Figure 6E**). We noticed a similar distinction in published images of the nuclear distribution of wildtype and frameshift mutant H1.4 in a study that used exogenous expression in HeLa cells (5), raising the possibility this could be a general effect of the frameshift mutations. Chromatin analyses are only beginning to be performed

on cells carrying the frameshift mutations (5, 14) and future genome-wide studies of chromatin structure and function in these cells are likely to be informative.

Despite evidence for histone H1.4 function in chromatin regulation, how expression of the frameshift mutant disrupts gene expression has remained obscure. Only one group has characterized cells from patients with *H1-4* frameshift mutations and although these results found evidence for changes in DNA methylation, these studies did not report gene expression changes (5, 14). To address this gap in knowledge, we generated LCLs from three of our subjects and compared the expression of candidate genes between these cells and control LCLs with WT or frameshift mutant H1.4. We focused on genes that had been related to the DNA methylation changes reported by Ciolfi et al (14), as well as complement factors as candidate genes implicated in aberrant synapse development in the brain. Although we saw differences in expression of some of the complement factors, we did not observe similar dysregulation of these genes by expression in control cells of frameshift mutant H1.4 alone which could suggest there are additional genetic determinants of this effect in the subjects (**Figure 5**). As to the methylation pathways, the Ciolfi et al. (14) paper used peripheral blood isolated directly from patients whereas our study used immortalized LCLs in culture, which could account for cell-type specific DNA methylation differences.

One of most significant challenges is unraveling the neurodevelopmental pathophysiology observed in individuals with frameshift mutant H1.4. However, this also presents an important opportunity for elucidating the functions of chromatin regulation in the developing brain. Like RMNS, many non-syndromic cases of ID and ASD have been shown to arise from *de novo* pathologic variants in chromatin regulators (3, 37). The other major class of ID/ASD-associated genes are synaptic proteins, however whether there is a privileged relationship between ID/ASD-associated chromatin regulators and synaptic development or function is poorly understood (38). There is some limited evidence of brain structural abnormalities in the small number of RMNS patients that have been studied by MRI, with thin corpus callosum as one abnormality identified (2) (**Supplemental File 1**). Identification of additional subjects with *H1-4* frameshift mutations will provide a larger subject pool for future evaluation of structural brain development.

Importantly, it is the fine tuning of the synaptic wiring diagram that has been implicated in many forms of ID and ASD (39, 40). Because there are not yet animal models of *H1-4* frameshift mutations, nothing is known of how frameshift mutant H1.4 impacts synapse development or function in the brain. We now show that exogenous expression of frameshift mutant H1.4 in primary embryonic rat hippocampal neurons in culture impairs the development of synaptic connectivity and reduces action potential firing (**Figure 7**). These changes in neuronal physiology occur despite normal dendritic architecture, suggesting that frameshift mutant H1.4 could selectively affect the expression of genes that play roles in synaptic connectivity and function or neuronal excitability. One limitation of our study is that we drove expression of H1.4 in neurons from an exogenous promoter that may not mimic expression of the endogenous protein. However, the histone H1s have been shown to be expressed in the brain, and levels of the rat homolog of H1.4 in particular have been shown to rise over the postnatal period, consistent with this protein playing a role in neuronal development (10). **Going forward, identifying the expression patterns**

and chromatin binding sites of WT and mutant H1.4 in rodent model systems will provide a key foundation for investigating possible disruptions of neuronal physiology driven by expression of frameshift mutant H1.4.

Limitations

While our data expand the current base of knowledge into how mutations in *H1-4* lead to RMNS, the total number of patients identified with clinically relevant *H1-4* mutations is still small. This limits our ability to know the full spectrum of phenotypes associated with *H1-4* mutations.

We selected only a small set of target genes for transcriptional analysis by qPCR and could have missed important gene targets of frameshift mutant H1.4 that might have been detected with a more unbiased method like RNA-seq. The genetic background of the subject and control cells interacted with the *H1-4* mutation to cause disease this could have led us to miss significant effects on gene expression. In the future, it may be valuable to CRISPR correct the *H1-4* mutation in the subject-derived LCLs. Finally, LCLs may not be the appropriate cell type to look for gene expression changes relevant to ID and future studies using iPSC-derived neurons would be highly useful.

Finally, as noted above, the studies on expression of WT and mutant H1.4 in neurons were performed using an exogenous promoter that may not mimic expression of the endogenous *H1-4* gene or maintain ratio of WT vs mutant H1.4. Future studies using knock in animal models will offer a more rigorous method to discover *H1-4* functions in neuronal physiology, especially *in vivo*.

Conclusions

In summary, this study combines human clinical data and molecular experiments to better understand the spectrum of *H1-4* mutations associated with variable clinical presentations in humans, with a specific focus on understanding how frameshift mutant H1.4 disrupts neuronal function. Our clinical reports of three *H1-4* loss-of-function cases in which the individuals lack key core phenotypes of RMNS suggests gain-of-function cause of RMNS. Our data show for the first time that frameshift mutant H1.4 expression can alter chromatin distribution in the nucleus and disrupt neuronal physiology, presumably via active roles in neuronal gene transcription.

List Of Abbreviations

RMNS: Rahman syndrome

EBV: Epstein-Barr Virus

LCL: Lymphoblastoid cell lines

ESC: Embryonic stem cells

ASD: Autism spectrum disorder

ID: Intellectual disability

WT: Wild type

qRT-PCR: Real-time quantitative reverse transcription PCR (polymerase chain reaction)

DIV: Days in vitro

Declarations

Ethics approval and consent to participate: This study was approved by the Duke University Institutional Review Board and the Yale University School of Medicine. Consent to participate was obtained from all patients.

Consent for publication: Consent for publication was obtained from all subjects in this study.

Availability of data and materials: Please contact authors to request materials

Competing interests: The Department of Molecular and Human Genetics at Baylor College of Medicine receives revenue from clinical genetic testing conducted at Baylor Genetics.

Funding: NIH grant 1R01NS098804 (A.E.W.), MH098114 (YH.J), MH104316 (YH.J), HD087795 (YH.J). HIST1H1E Genetic Syndrome Research Foundation (YH.J).

Authors' contributions: MWT and YHJ designed the human clinical data portion of study. MWT designed and performed the LCL portion of this study. MWT, MVG, and AEW designed the neuron experiment portion of the study. MWT, MVG, WDT performed all experiments. MWT, YHJ, JAR, HS, WC, NB, SAT collected patient genetic and phenotypic data. MWT and YHJ reviewed patient data. MWT, MVG, AEW and YJH wrote the manuscript. All authors revised the manuscript and approved the final version for publication.

Acknowledgements: The authors would like to thank the subjects described in this study and their families. This study used LCLs from the NINDS Repository. NINDS Repository sample numbers corresponding to the samples used are: ND23960, ND23648, AG09393. Lentiviral expression plasmids were gifts from David Baltimore and Bob Weinberg.

The authors would also like to thank: Fowzan S Alkuraya and Hanan E Shamseldin for confirming the *de novo* variant in subject #10. Janet Wooton for assistance in gathering clinical data from patients and their families. Joy Cogan and team at Vanderbilt University School of Medicine for generating patient derived LCLs. Scott Solderling for sharing MEA equipment.

References

1. Tatton-Brown K, Loveday C, Yost S, Clarke M, Ramsay E, Zachariou A, et al. Mutations in Epigenetic Regulation Genes Are a Major Cause of Overgrowth with Intellectual Disability. *Am J Hum Genet.* 2017;100(5):725-36.
2. Burkardt DD, Zachariou A, Loveday C, Allen CL, Amor DJ, Ardisson A, et al. HIST1H1E heterozygous protein-truncating variants cause a recognizable syndrome with intellectual disability and distinctive facial gestalt: A study to clarify the HIST1H1E syndrome phenotype in 30 individuals. *Am J Med Genet A.* 2019;179(10):2049-55.
3. Duffney LJ, Valdez P, Tremblay MW, Cao X, Montgomery S, McConkie-Rosell A, et al. Epigenetics and autism spectrum disorder: A report of an autism case with mutation in H1 linker histone HIST1H1E and literature review. *Am J Med Genet B Neuropsychiatr Genet.* 2018;177(4):426-33.
4. Takenouchi T, Uehara T, Kosaki K, and Mizuno S. Growth pattern of Rahman syndrome. *Am J Med Genet A.* 2018;176(3):712-4.
5. Flex E, Martinelli S, Van Dijck A, Ciolfi A, Cecchetti S, Coluzzi E, et al. Aberrant Function of the C-Terminal Tail of HIST1H1E Accelerates Cellular Senescence and Causes Premature Aging. *Am J Hum Genet.* 2019;105(3):493-508.
6. Helsmoortel C, Vandeweyer G, Ordoukhanian P, Van Nieuwerburgh F, Van der Aa N, and Kooy RF. Challenges and opportunities in the investigation of unexplained intellectual disability using family-based whole-exome sequencing. *Clin Genet.* 2015;88(2):140-8.
7. Pelle A, Pezzoli L, Apuril E, Iascone M, and Selicorni A. A novel HIST1HE pathogenic variant in a girl with macrocephaly and intellectual disability: a new case and review of literature. *Clinical dysmorphology.* 2020.
8. Bednar J, Horowitz RA, Grigoryev SA, Carruthers LM, Hansen JC, Koster AJ, et al. Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *Proc Natl Acad Sci U S A.* 1998;95(24):14173-8.
9. Harshman SW, Young NL, Parthun MR, and Freitas MA. H1 histones: current perspectives and challenges. *Nucleic Acids Res.* 2013;41(21):9593-609.
10. Pina B, Martinez P, and Suau P. Changes in H1 complement in differentiating rat-brain cortical neurons. *Eur J Biochem.* 1987;164(1):71-6.
11. Parseghian MH. What is the role of histone H1 heterogeneity? A functional model emerges from a 50 year mystery. *AIMS Biophys.* 2015;2(4):724-72.
12. Fan Y, Sirotkin A, Russell RG, Ayala J, and Skoultchi AI. Individual somatic H1 subtypes are dispensable for mouse development even in mice lacking the H1(0) replacement subtype. *Mol Cell Biol.* 2001;21(23):7933-43.
13. Parseghian MH, Newcomb RL, and Hamkalo BA. Distribution of somatic H1 subtypes is non-random on active vs. inactive chromatin II: distribution in human adult fibroblasts. *J Cell Biochem.* 2001;83(4):643-59.

14. Ciolfi A, Aref-Eshghi E, Pizzi S, Pedace L, Miele E, Kerkhof J, et al. Frameshift mutations at the C-terminus of HIST1H1E result in a specific DNA hypomethylation signature. *Clin Epigenetics*. 2020;12(1):7.
15. Alami R, Fan Y, Pack S, Sonbuchner TM, Besse A, Lin Q, et al. Mammalian linker-histone subtypes differentially affect gene expression in vivo. *Proc Natl Acad Sci U S A*. 2003;100(10):5920-5.
16. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. *Genome Res*. 2002;12(6):996-1006.
17. Lois C, Hong EJ, Pease S, Brown EJ, and Baltimore D. Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science*. 2002;295(5556):868-72.
18. Yang SM, Kim BJ, Norwood Toro L, and Skoultschi AI. H1 linker histone promotes epigenetic silencing by regulating both DNA methylation and histone H3 methylation. *Proc Natl Acad Sci U S A*. 2013;110(5):1708-13.
19. Ghosh RP, Horowitz-Scherer RA, Nikitina T, Shlyakhtenko LS, and Woodcock CL. MeCP2 binds cooperatively to its substrate and competes with histone H1 for chromatin binding sites. *Mol Cell Biol*. 2010;30(19):4656-70.
20. Ito-Ishida A, Yamalanchili HK, Shao Y, Baker SA, Heckman LD, Lavery LA, et al. Genome-wide distribution of linker histone H1.0 is independent of MeCP2. *Nat Neurosci*. 2018;21(6):794-8.
21. Stewart SA, Dykxhoorn DM, Palliser D, Mizuno H, Yu EY, An DS, et al. Lentivirus-delivered stable gene silencing by RNAi in primary cells. *RNA*. 2003;9(4):493-501.
22. Ko DC, Shukla KP, Fong C, Wasnick M, Brittnacher MJ, Wurfel MM, et al. A genome-wide in vitro bacterial-infection screen reveals human variation in the host response associated with inflammatory disease. *Am J Hum Genet*. 2009;85(2):214-27.
23. Tao X, West AE, Chen WG, Corfas G, and Greenberg ME. A calcium-responsive transcription factor, CaRF, that regulates neuronal activity-dependent expression of BDNF. *Neuron*. 2002;33(3):383-95.
24. McQuin C, Goodman A, Chernyshev V, Kametsky L, Cimini BA, Karhohs KW, et al. CellProfiler 3.0: Next-generation image processing for biology. *PLoS Biol*. 2018;16(7):e2005970.
25. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012;9(7):676-82.
26. Prevention CfDCa. Clinical Growth Charts. https://www.cdc.gov/growthcharts/clinical_charts.htm. Updated November 11, 2020.
27. Maddirevula S, Alzahrani F, Al-Owain M, Al Muhaizea MA, Kayyali HR, AlHashem A, et al. Autozygome and high throughput confirmation of disease genes candidacy. *Genet Med*. 2019;21(3):736-42.
28. Nishiyama M, Oshikawa K, Tsukada Y, Nakagawa T, Iemura S, Natsume T, et al. CHD8 suppresses p53-mediated apoptosis through histone H1 recruitment during early embryogenesis. *Nat Cell Biol*. 2009;11(2):172-82.
29. Weiss T, Hergeth S, Zeissler U, Izzo A, Tropberger P, Zee BM, et al. Histone H1 variant-specific lysine methylation by G9a/KMT1C and Glp1/KMT1D. *Epigenetics Chromatin*. 2010;3(1):7.

30. Kassner I, Barandun M, Fey M, Rosenthal F, and Hottiger MO. Crosstalk between SET7/9-dependent methylation and ARTD1-mediated ADP-ribosylation of histone H1.4. *Epigenetics Chromatin*. 2013;6(1):1.
31. Strom R, Santoro R, D'Erme M, Mastrantonio S, Reale A, Marenzi S, et al. Specific variants of H1 histone regulate CpG methylation in eukaryotic DNA. *Gene*. 1995;157(1-2):253-6.
32. Stephan AH, Barres BA, and Stevens B. The complement system: an unexpected role in synaptic pruning during development and disease. *Annu Rev Neurosci*. 2012;35:369-89.
33. Gupta S, Ellis SE, Ashar FN, Moes A, Bader JS, Zhan J, et al. Transcriptome analysis reveals dysregulation of innate immune response genes and neuronal activity-dependent genes in autism. *Nat Commun*. 2014;5:5748.
34. Fagan K, Crider A, Ahmed AO, and Pillai A. Complement C3 Expression Is Decreased in Autism Spectrum Disorder Subjects and Contributes to Behavioral Deficits in Rodents. *Mol Neuropsychiatry*. 2017;3(1):19-27.
35. Magdalon J, Mansur F, Teles ESAL, de Goes VA, Reiner O, and Sertie AL. Complement System in Brain Architecture and Neurodevelopmental Disorders. *Front Neurosci*. 2020;14:23.
36. Parseghian MH, and Hamkalo BA. A compendium of the histone H1 family of somatic subtypes: an elusive cast of characters and their characteristics. *Biochem Cell Biol*. 2001;79(3):289-304.
37. Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An JY, et al. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell*. 2020;180(3):568-84 e23.
38. Chen LF, Zhou AS, and West AE. Transcribing the connectome: roles for transcription factors and chromatin regulators in activity-dependent synapse development. *J Neurophysiol*. 2017;118(2):755-70.
39. Heavner WE, and Smith SEP. Resolving the Synaptic versus Developmental Dichotomy of Autism Risk Genes. *Trends Neurosci*. 2020;43(4):227-41.
40. Zoghbi HY, and Bear MF. Synaptic dysfunction in neurodevelopmental disorders associated with autism and intellectual disabilities. *Cold Spring Harb Perspect Biol*. 2012;4(3).
41. Deciphering Developmental Disorders S. Large-scale discovery of novel genetic causes of developmental disorders. *Nature*. 2015;519(7542):223-8.

Figures

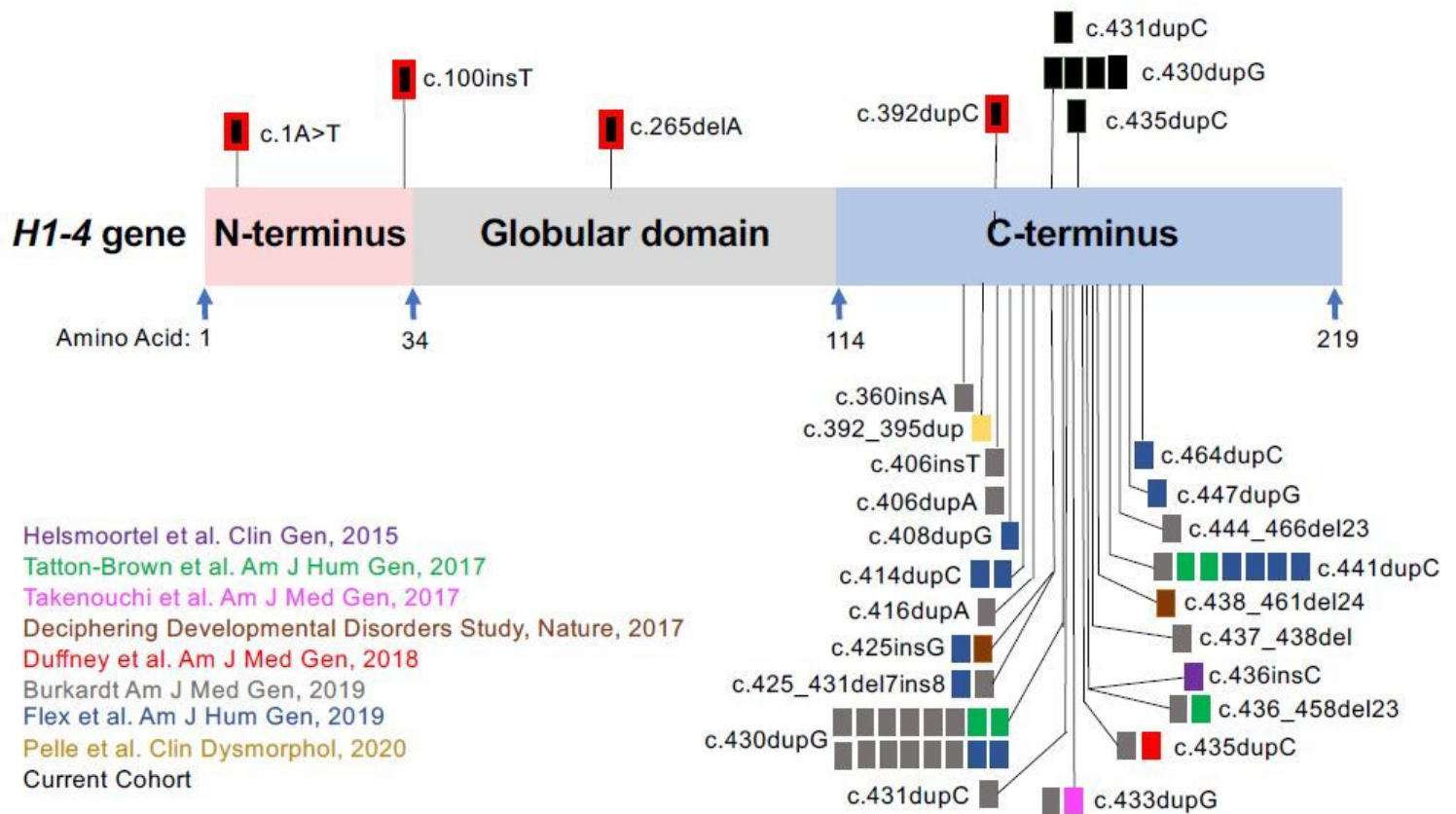


Figure 1

Summary of mutations in H1-4 The H1-4 gene encodes a 219 amino acid protein composed of 3 domains: a short N-terminal tail, a globular domain, and a long C-terminal tail. Above the gene schematic (black boxes) are mutations detected in the patients recruited for this study. Solid black boxes indicate that the individual was previously reported by another group. Black box with red outline indicates patients which have not been reported by other groups. Mutations which have previously been published are drawn under the peptide schematic, with color of the box corresponding with the publication reporting the mutation. (Purple (6); Green,(1); Pink,(4); Auburn,(41); Red,(3); Grey,(2); Blue,(5); Yellow(7)).

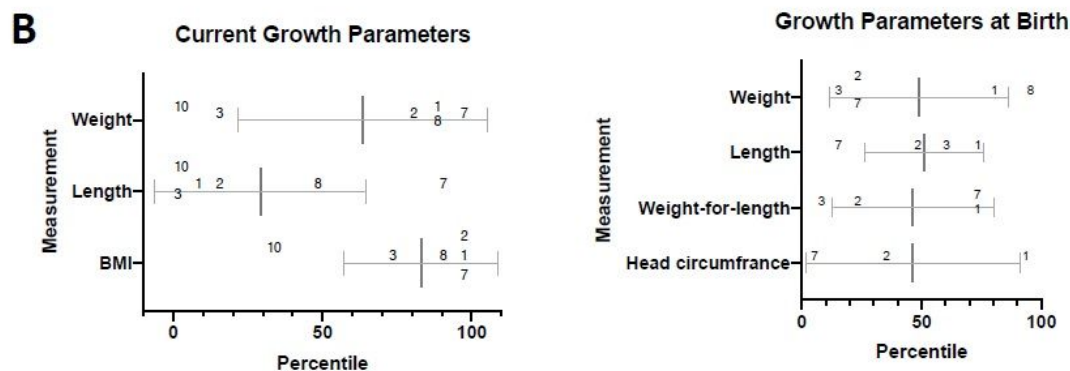
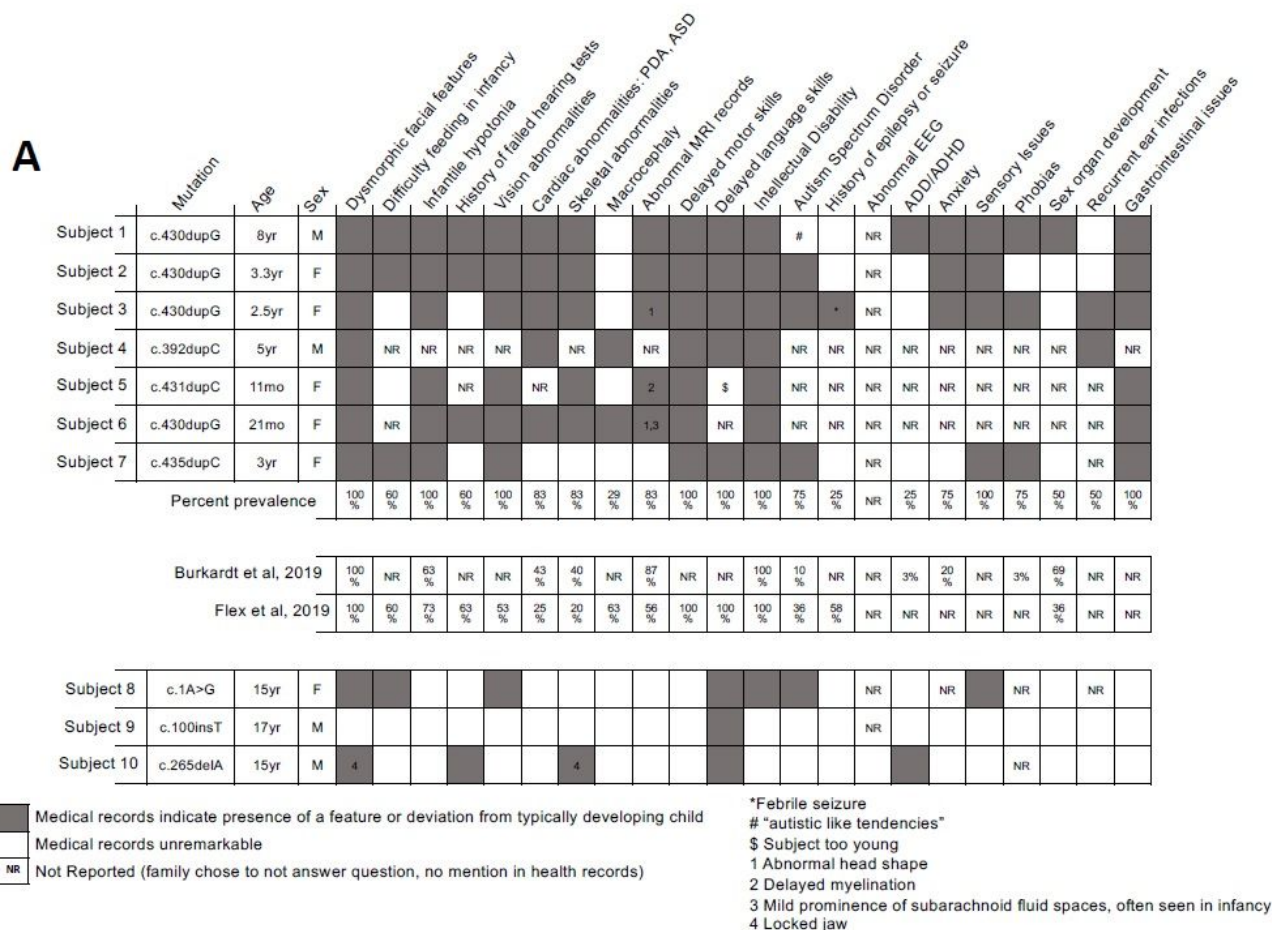


Figure 2

Clinical presentation of subjects in this study For subjects in this study (Subjects 1-9), medical and health history was collected, and guardians were asked to self-report in a survey covering additional health and development metrics. (A) In the table, grey boxes indicate that the patient's phenotype does differ from that of a typically developed child. White box indicates that respective questions in the self-reporting survey were answered and the patient's phenotype was unremarkable and does not differ from that of a

typical developed child. White box with NR indicates that self-reports or medical records were not available. Average prevalence of previously published phenotypes, as well as phenotype of haploinsufficiency Subject 8 and 9, are also included. (B). Comparison of growth parameters of our subjects against CDC clinical growth charts (percentile). Numbers correspond to growth parameters for each subject. Thick grey bar indicates average growth percentile, thin grey bar indicates standard deviation.

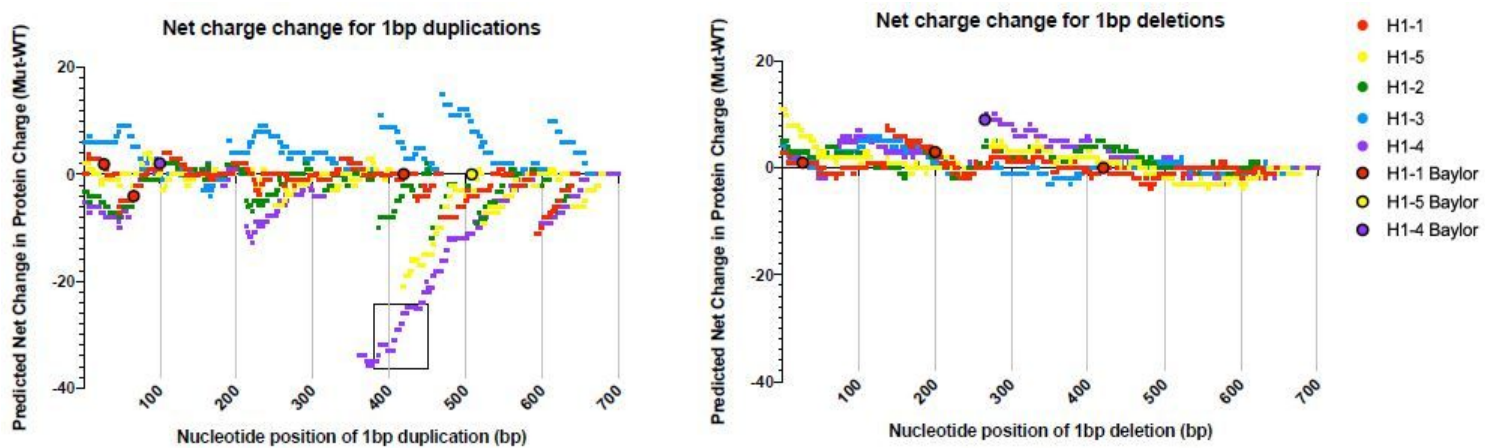


Figure 3

Frameshift modeling of the H1 family of genes R was utilized to systematically create one base pair (bp) duplications (left) or 1bp deletions (right) throughout the length of the H1-1 (red), H1-5 (yellow), H1-2 (green), H1-3 (blue), and H1-4 (purple) genes, translate the sequences until a stop codon was reached, and calculate net charge change from a WT protein of corresponding length. Frameshift mutations found in the clinical ES database are points outlined in black. Black box around H1-4 points shows previously published RMNS frameshift mutations (2, 5).

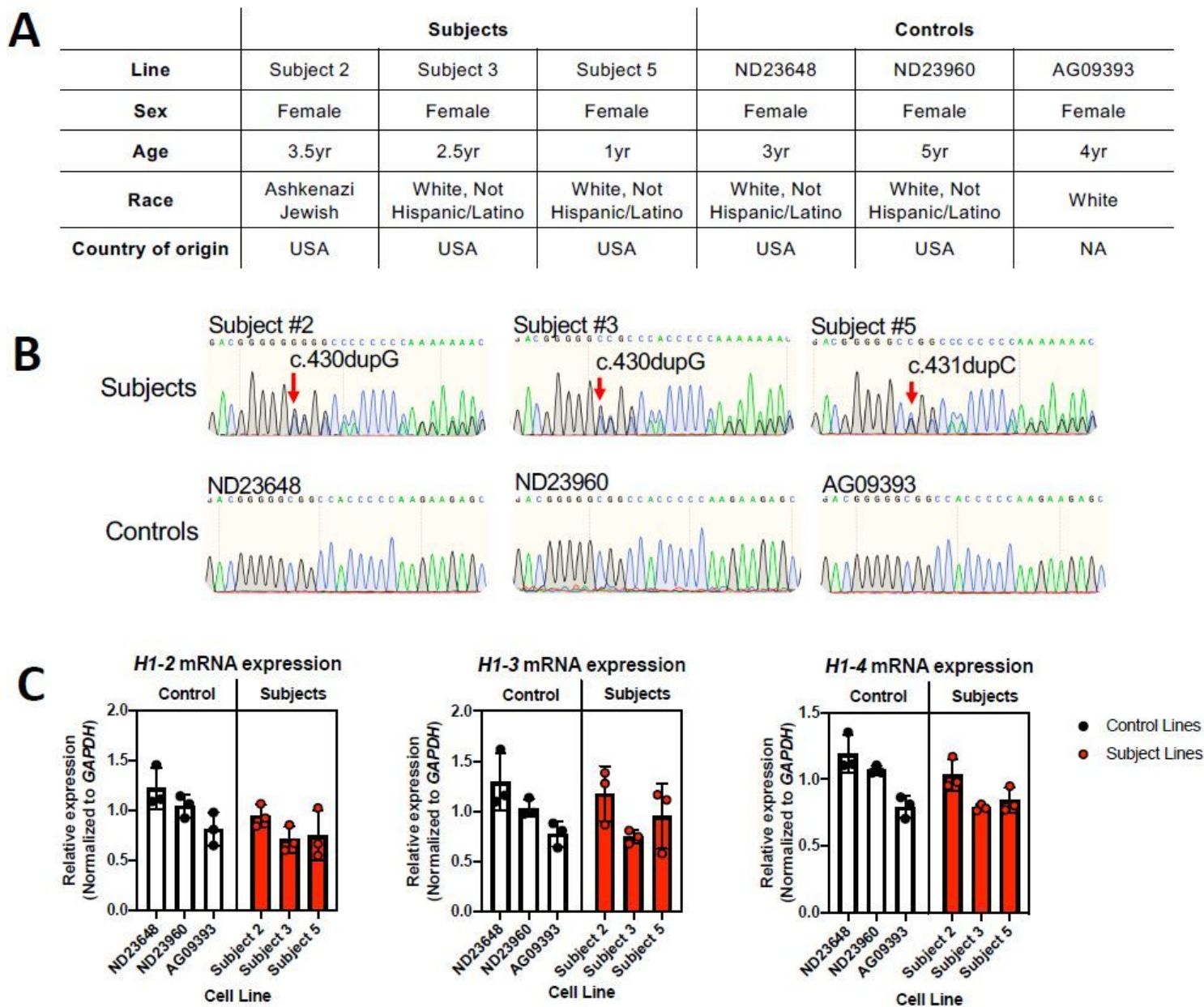


Figure 4

Transcript abundance of H1 family members in patient derived cells LCL cell lines (A) established from subject blood or ordered from Coriell cell repositories were (B) Sanger sequenced to confirm genotype of the H1-4 gene. Shown are representative chromatograms of the areas containing duplications in the patient cell lines. Red arrow indicates location of the mutations in the patient lines. (C) qPCR for H1-2, -3, and -4 in control and subject derived LCLs. Two-tailed Nested t-test, N=3 cell lines, n=3 for each cell line, error bars represent standard deviation. H1-2 p=0.179, H1-3 p=0.728, H1-4 p=0.402.

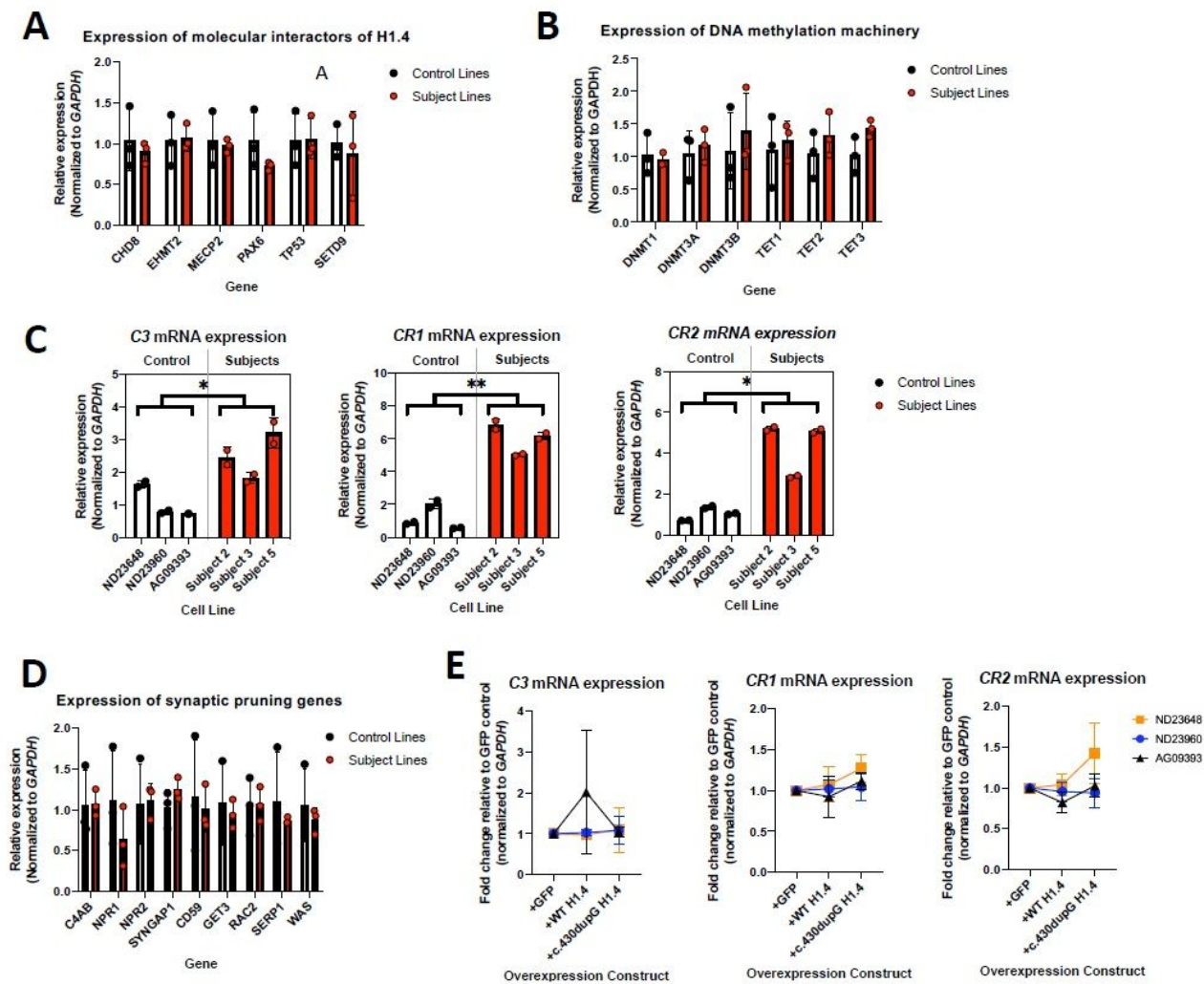


Figure 5

Transcriptional characterization of subject derived cells qPCR of (A) molecular interactors of H1.4 in mice and (B) DNA methylation machinery. N=3 cell lines/experiment. Multiple Two-tailed T-tests, CHD8 $p=0.545$, EHMT2 $p=0.862$, MECP2 $p=0.775$, PAX6 $p=0.203$, TP53 $p=0.920$, SETD9 $p=0.703$, DNMT1 $p=0.680$, DNMT3A $p=0.648$, DNMT3B $p=0.560$, TET1 $p=0.686$, TET2 $p=0.370$, TET3 $p=0.081$. (C) C3, CR1, and CR2 expression is increased in patient cells. Two-tailed Nested t-test, N=3 cell lines, n=3 experiments per cell line. C3 $p=0.0439$, CR1 $p=0.0021$, CR2 $p=0.0130$. *= <0.05 , **= <0.005 . Error bars represent standard deviation. (D) qPCR of genes implicated in complement-dependent synaptic pruning in mice. Two-tailed T-test, N=3 cell lines/experiment. C4AB $p=0.939$, NPR1 $p=0.319$, NPR2 $p=0.889$, SYNGAP1 $p=0.1996$, CD59 $p=0.749997$, GET3 $p=0.684$, RAC2 $p=0.935$, SERP1 $p=0.528$, WAS $p=0.527$. (E) Fold change of the expression of complement-associated genes, with FLAG-myc-WT or frameshift mutant FLAG-myc-c.430dupG H1.4 overexpression in LCLs, normalized to the cell line's GFP infection control. Repeated measures one-way ANOVA. C3 $p=0.449$, CR1 $p=0.265$, CR1 $p=0.286$. N=3 cell lines each with an exogenously expressed plasmid, n=3 experiments per cell line, error bars represent standard deviation.

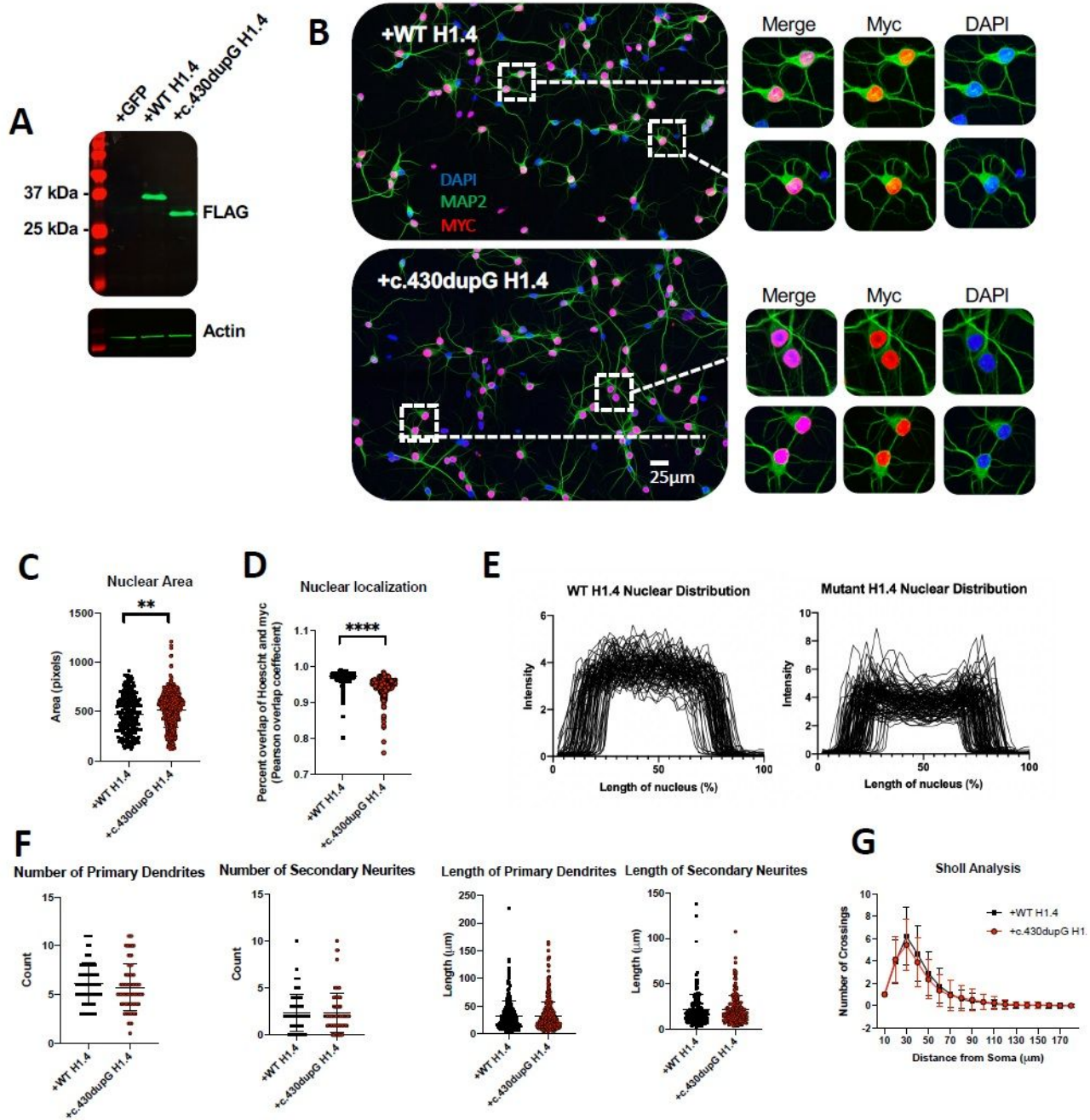


Figure 6

Overexpression of FLAG-myc-Tagged H1.4 in rat primary neurons (A) Western blotting of whole cell extract of primary neurons with anti-FLAG antibody. (B) Representative images of DIV7 rat hippocampal neurons exogenously expressing FLAG-myc-WT H1.4 or frameshift mutant FLAG-myc-c.430dupG H1.4 (stained with myc in red) is expressed in the nucleus of neurons (labeled with Map2 in green). (C) CellProfiler Quantification of percent overlap between Myc and the Hoescht signal of neurons ($p < 0.000001$) and (D) of nuclear size from Hoescht signal ($p = 0.0015$). N=3 experiments, n=234 WT H1.4

nuclei, n=511 mutant H1.4 nuclei. Unpaired two-tailed T-test, **=<0.005, ****=<0.0001. (E) Representative images of pixel intensity of myc signal (H1.4) across the diameter of neurons exogenously expressing FLAG-myc-WT H1.4 or frameshift mutant FLAG-myc-c.430dupG H1.4. (F) Neurite outgrowth tracing shows no difference between number of primary dendrites (p=0.248), length of primary dendrites (p=0.511), number of secondary neurites (p=0.926), or length of secondary neurites (0.835). N=3 experiments, n=74 WT H1.4 neurons, n=75 mutant H1.4 neurons. Unpaired two-tailed T-test. (G) Sholl analysis of these traces is not significant (p=0.2344), Repeated Measures Two-way ANOVA, Tukey correction. N=3 experiments, n=76 WT H1.4 neurons, n=75 mutant H1.4 neurons. Error bars represent standard deviation.

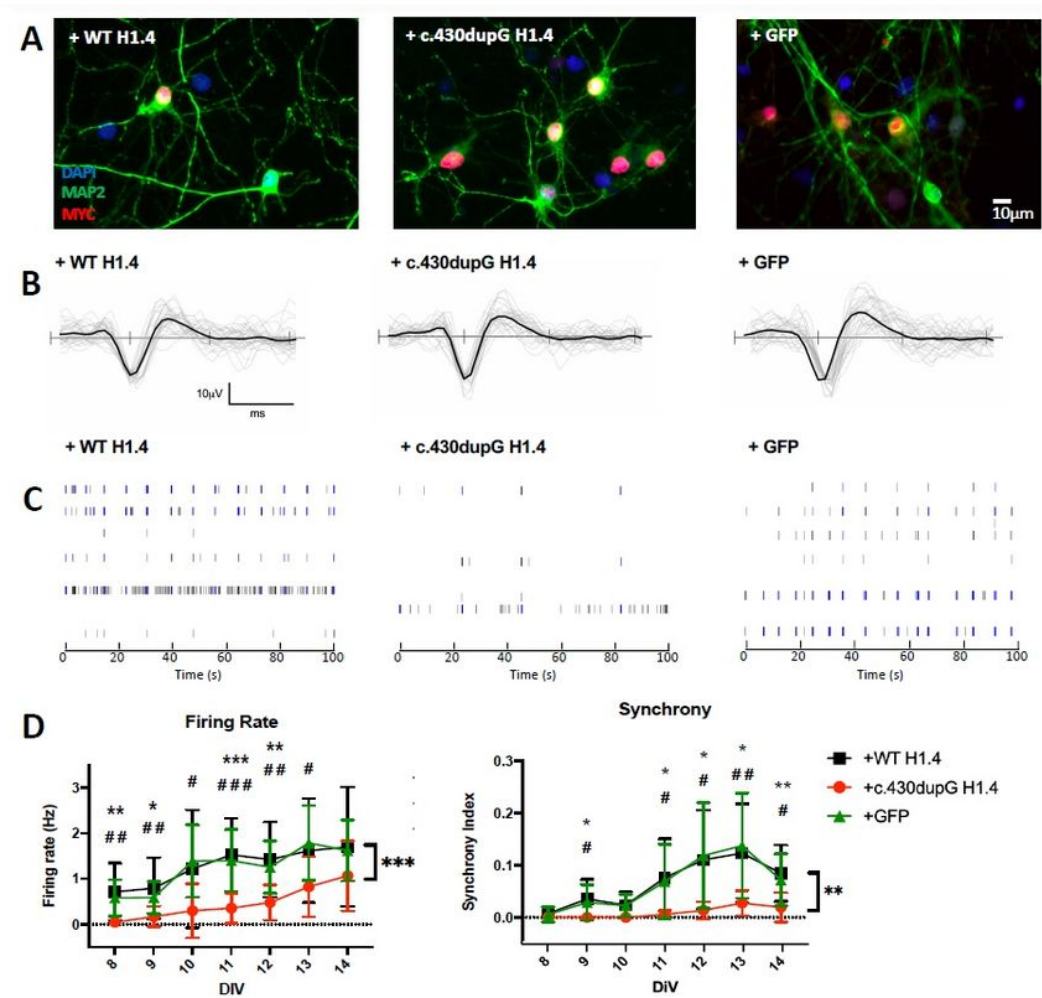


Figure 7

Effects of frameshift mutant H1.4 on firing rate and synchrony in rat hippocampal neurons (A) Representative images showing neurons at DIV14. Red= Myc (H1.4) or GFP, Green= MAP2, Blue=DAPI. (B) Representative traces of action potential (AP) waveforms from single electrodes of a 16 multi-electrode array recording. The average waveform is shown in black. (C) Raster plots showing spontaneous firing of APs in WT H1.4, frameshift mutant c.430dupG mut H1.4, and GFP expressing neurons over a 100-second time period. Each row represents an electrode, each dash on the plot indicates a spike. Synchronous spikes, in which multiple electrodes have activity simultaneously, are colored in blue. Non-synchronous spikes are colored in black. (D) Quantification of firing rate and synchrony throughout development in c.430dupG mut H1.4 neurons compared to WT H1.4 and GFP expressing neurons. N=3 experimental plates, n= 4-6 replicas of each condition per plate. Mixed effects analysis with multiple comparisons, For multiple comparisons: WT vs mut $\ast = <0.05$, $\ast\ast = <0.01$, $\ast\ast\ast = <0.001$, GFP vs mut $\# = <0.05$, $\#\# = <0.01$, $\#\#\# = <0.001$, brackets denote P value summary of overexpression (Firing rate $p = 0.0006$, Synchrony $p = 0.0017$).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementalFigure1Tremblay.pdf](#)
- [SupplementalFigure2Tremblay.pdf](#)
- [SupplementalFigure3Tremblay.pdf](#)
- [SupplementalFile1Tremblay.xlsx](#)
- [SupplementalFile2Tremblay.xlsx](#)
- [SupplementalTable1Tremblay.xlsx](#)
- [SupplementalTable2Tremblay.pdf](#)