Histone H3.3 beyond cancer: Germline mutations in Histone 3 Family 3A and 3B cause a previously unidentified neurodegenerative disorder in 46 patients

Laura Bryant
Children's Hospital of Philadelphia, Philadelphia, PA
Marcia C Willing
Washington University School of Medicine in St. Louis
Linda Manwaring
Washington University School of Medicine in St. Louis
et al

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Recommended Citation
Bryant, Laura; Willing, Marcia C; Manwaring, Linda; and et al, "Histone H3.3 beyond cancer: Germline mutations in Histone 3 Family 3A and 3B cause a previously unidentified neurodegenerative disorder in 46 patients." Science Advances. 6, 49. (2020).
https://digitalcommons.wustl.edu/open_access_pubs/9978

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
Histone H3 beyond cancer: Germline mutations in Histone 3 Family 3A and 3B cause a previously unidentified neurodegenerative disorder in 46 patients

Although somatic mutations in Histone 3.3 (H3.3) are well-studied drivers of oncogenesis, the role of germline mutations remains unreported. We analyze 46 patients bearing de novo germline mutations in histone 3 family 3A (H3F3A) or H3F3B with progressive neurologic dysfunction and congenital anomalies without malignancies. Molecular modeling of all 37 variants demonstrated clear disruptions in interactions with DNA, other histones, and histone chaperone proteins. Patient histone posttranslational modifications (PTMs) analysis revealed notably aberrant local PTM patterns distinct from the somatic lysine mutations that cause global PTM dysregulation. RNA sequencing on patient cells demonstrated up-regulated gene expression related to mitosis and cell division, and cellular assays confirmed an increased proliferative capacity. A zebrafish model showed craniofacial anomalies.

INTRODUCTION

The study of histones and their role in epigenetics is a rapidly expanding field. Histones are nuclear proteins that associate with DNA to facilitate packaging into condensed chromatin. Histones are dynamically decorated with posttranslational modifications (PTMs), which regulate these processes as DNA repair, gene expression, mitosis, and meiosis. Dysregulation of PTMs leads to cancer, neurodevelopmental syndromes, psychiatric disorders, and even cardiovascular disease (1–4). The vast array of diseases that stem from histone dysregulation makes understanding histone biology vital to understanding the pathophysiology of many diseases and developing treatments. Histones are highly evolutionarily conserved, and only a few germline disease-causing variants in the histones themselves have been found (5). Rather, most disease-causing variants affecting
histone regulation are found in histone-modifying proteins or histone chaperones. Somatic mutations in histone H3.3 (H3.3), encoded by both H3F3A and H3F3B, have recently been identified in pediatric glia and other tumors. This discovery revolutionized the field of epigenetic changes in tumors and how they lead to cancer progression (6). There has been one case report suggesting a relationship between a neurodevelopmental syndrome and H3F3A, but no additional patients or functional data have been published previously (7). Studying how histone mutations cause disease can provide the key to understanding how mutations in the histone network lead to disease.

Histone 3 family 3 (H3F3) histones (H3.3) mark active genes, maintain epigenetic memory, and maintain heterochromatin and telomeric integrity. Every nucleosome contains a version of H3. H3.1 and H3.2 are canonical histones that are replication dependent and, therefore, added to chromatin only during DNA replication. H3.3, however, is a replication-independent variant, which differs from the canonical H3.1 and H3.2 by only five and four amino acids, respectively (8). H3F3A and H3F3B code for an identical protein product, H3.3, a 135–amino acid protein (after cleavage of the first methionine) with a histone tail, four α helices, and two loop domains. These genes, however, contain different regulatory and coding sequences, which lead to different expression patterns and levels for H3F3A and H3F3B. Both H3F3B and H3F3A are expressed ubiquitously, during development and throughout life, with relatively high levels of expression in the brain, testes, ovaries, and uterus (9).

Because of the biologic importance of H3F3A and H3F3B, their orthologs have been studied extensively in multiple model organisms. When both H3F3a and H3F3b are knocked out in mice, it causes embryonic lethality at embryonic day 6.5, and reduced expression leads to sterility, growth retardation, and increased neonatal lethality.
RESULTS

Patient mutations and phenotype

Our patient cohort consists of 46 unrelated patients bearing de novo germline missense mutations in H3F3A or H3F3B (Fig. 1). Both H3F3A and H3F3B code for H3.3. Therefore, a heterozygous mutation in either gene would affect only about 25% of the total H3.3 in the cell, depending on different expression levels from the two genes. None of the patients have malignancies, but they share a phenotype of developmental delay, usually severe and often progressive, with mostly minor congenital anomalies (Table 1). These patients were identified through exome or genome sequencing performed for the indication of neurodevelopmental delays and/or congenital anomalies (Table 1, fig. S1, and table S1). There is substantial phenotypic heterogeneity in the patients, and future individuals are likely to be identified using unbiased molecular testing. Notably, 9 of the 46 patients (21%) have demonstrated clinical neurologic degeneration, which suggests that this may be a progressive disorder. Multiple patients (26% of the cohort) have cortical atrophy on brain magnetic resonance imaging (MRI), even without intractable epilepsy. For example, the oldest patient in this cohort (32 years) developed seizures at the age of 14 years and a progressive spastic paraparesis starting at the age of 29 years.

The severity of the phenotype does not correlate with the location of the mutation or whether the mutation is in H3F3A or H3F3B. Four unrelated patients have the p.T45I mutation and exhibit the full range of the phenotype severity. The age at sitting for patients with the p.T45I mutation ranges from 8 months to not achieved within 4.5 years. The age at walking was 21 months, 2 years, and 3 years and not achieved at 4.5 years. One exhibited developmental regression, while the other three did not, and only one has seizures. Even the tone abnormalities are inconsistent with three exhibiting hypertonia, while one has hypotonia. This strongly suggests that the variation in phenotype is not due to the location of the mutation but rather due to either modifying alleles or environmental factors.

Five mutations [H3F3A (NM_002107.4) p.R17G, p.T45I, p.A114G, p.Q125R, and H3F3B (NM_005324.4) p.P121R] were detected in two or more unrelated patients. p.Q125R was found in H3F3A in three individuals and in a different individual in H3F3B, while p.P121R was found in H3F3A in one individual and in H3F3B in two other individuals. Notably, in two patients, the mutations (H3F3A p.Q125R and p.V117L) were identified only by trio genome sequencing after negative exome sequencing, as the last exon of H3F3A is currently not covered by some exome capture kits. We speculate that similar mutations in H3F3A may be underdiagnosed in current exome studies. It should be noted for nomenclature consistency purposes that, historically, many publications on H3.3 exclude the initiator methionine in the residue count. In this publication, we will use the traditional histone nomenclature and exclude the first methionine. However, we will include traditional human genetics nomenclature in the clinical table of mutations and include this first residue in the mutation notation.

One mutation in H3F3B (p.S146X) is only present in the alternate transcript. It is p.V117V in the canonical transcript. Although the mutation is only in the alternate transcript, the patient phenotype is consistent with the rest of the cohort, and the patient cells have the same phenotype as the other patient cells analyzed. The role of the alternate transcript is currently unknown and requires additional study.

These 37 unique missense mutations in 46 patients are all de novo and, with one exception (H3F3A c.362 T > A; p.M120K, 1 incidence), are not found in a large database of 138,632 controls, the Genome Aggregation Database (gnomAD) (13). Upon closer viewing of the raw data for this p.M120K variant, it may be a technical mapping error, as it is only present on one strand and did not meet the previous Exome Aggregation Consortium (ExAC) quality control criteria for reporting. In the general population, both genes have a very low rate of missense variants; the gnomAD missense Z score is 3.21 for H3F3A and 2.95 for H3F3B (>2 is significant) (13).

Computational modeling of patient mutations

It is likely that these variants are pathogenic through various different mechanisms, as they are found throughout the entire coding sequence.
in different domains. To explore these mechanisms further, we used molecular modeling of all 37 of the variants reported here. These 37 variants were mapped to a total of 25 loci distributed over several experimental structures. Eleven locations were in the structure of H3.3 in the nucleosome; at the sequence level, they spread from central positions to near the C-terminal end of H3.3. The remaining 14 locations (as well as H39, which was also present in the nucleosome) were mapped to different complexes between H3.3 and epigenetic regulators. At the sequence level, they concentrated in the N-terminal tail of H3.3.

In Fig. 2 (top), we show the location of the variants in the H3.3 structure (for clarity, they are shown only in one H3.3 monomer) and (Fig. 2, bottom) a summary of the amount and type of interatomic contacts at each mutation locus. Together, these data suggest three broad scenarios for the variants’ impact. In the first one, variants are more likely to disrupt the H3.3-DNA interaction because the native residue is involved in a large number of contacts with the DNA. For instance, this would be the case of p.R83C because the arginine residue penetrates the DNA minor groove. In the second scenario, variants are more likely to disrupt the histone octamer, either because they affect the intranomer contacts of H3.3 (e.g., p.Q125R) or because they may alter the interaction of H3.3 with other nucleosomal histones (e.g., p.L48R). In the third scenario, the variants disrupt histone-protein binding such as the p.G90R mutation disrupting chaperone binding. In summary, through different mechanisms, the variants in this section are likely to affect the formation, the deposition, or the stability of the nucleosomes containing H3.3 or of the nucleosome pairing. This may result in a generalized loosening of chromatin structure in those processes requiring the incorporation of H3.3 to the nucleosome.

The available complexes between H3.3 and epigenetic regulators (Fig. 3A) involve short stretches of the H3.3 tail bound to sites of different shapes (Fig. 3B). The number of interprotein contacts at the variant locus may vary substantially. For example, p.H39 has 33 and 4 contacts with SETD2 (SET Domain Containing 2, Histone Lysine Methyltransferase) and ZMYND11 (Zinc Finger MYND-Type Containing 11) domains, respectively. We find that residues p.R8, p.S31, p.K36, and p.H39 are involved in more than 10 contacts across epigenetic regulators, suggesting that their mutation may disrupt one or more biologically relevant interactions. For the remaining residues,

<table>
<thead>
<tr>
<th>Table 1. Clinical findings of affected patients. Developmental delay was seen in varying degrees in the majority of patients; hypotonia, poor growth, oculomotor abnormalities, seizures, abnormal skull shape, and microcephaly were also commonly seen in the patients.</th>
</tr>
</thead>
<tbody>
<tr>
<td>33 H3F3A/13 H3F3B</td>
</tr>
<tr>
<td>25 unique H3F3A/12 unique H3F3B variants</td>
</tr>
<tr>
<td>3 variants in both H3F3A and B</td>
</tr>
<tr>
<td>Ages: 2 months to 32 years; 3 deceased at 2, 10 months, and 2.5 years</td>
</tr>
<tr>
<td>Growth</td>
</tr>
<tr>
<td>Height undergrowth 15/46 (33%)</td>
</tr>
<tr>
<td>Height overgrowth 5/46 (11%)</td>
</tr>
<tr>
<td>Weight undergrowth 4/37 (11%)</td>
</tr>
<tr>
<td>Weight overgrowth 9/37 (27%)</td>
</tr>
<tr>
<td>Microcephaly 12/46 (26%)</td>
</tr>
<tr>
<td>Development/neurologic</td>
</tr>
<tr>
<td>Achieved independent sitting (range: 6 months to 7 years) 33/45 (73%)</td>
</tr>
<tr>
<td>Achieved independent walking (range: 15 months to 8 years) 25/41 (61%)</td>
</tr>
<tr>
<td>Spoke at least one word (range: 12 months to 6 years) 16/43 (37%)</td>
</tr>
<tr>
<td>Progression of neurologic disease 9/46 (20%)</td>
</tr>
<tr>
<td>Brain anomalies on imaging 30/41 (73%)</td>
</tr>
<tr>
<td>Cortical atrophy 10/41 (24%)</td>
</tr>
<tr>
<td>Seizures 23/46 (50%)</td>
</tr>
<tr>
<td>Hypotonia 31/46 (67%)</td>
</tr>
<tr>
<td>Spasticity 10/46 (22%)</td>
</tr>
<tr>
<td>Contractures 12/46 (22%)</td>
</tr>
<tr>
<td>Oculomotor abnormalities 25/46 (54%)</td>
</tr>
<tr>
<td>Craniofacial</td>
</tr>
<tr>
<td>High/prominent forehead 14/46 (30%)</td>
</tr>
<tr>
<td>Hypertelorism 11/46 (24%)</td>
</tr>
<tr>
<td>Craniosynostosis/abnormal head shape (excluding plagiocephaly) 14/46 (30%)</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Atrial septal defect 8/46 (18%)</td>
</tr>
<tr>
<td>Skeletal abnormalities 16/46 (35%)</td>
</tr>
<tr>
<td>Genital abnormalities Males 12/27 (44%), Females 0/19 (0%)</td>
</tr>
<tr>
<td>Chronic constipation 11/45 (24%)</td>
</tr>
</tbody>
</table>

Fig. 2. Variant locations in the nucleosome. At the top of the figure, we show the structure of the nucleosome with the H3.3 variants identified with spheres; the H3.3 monomer carrying them is colored in dark emerald green. The coloring of the variants reflects the predominating interactions at each location: DNA binding (magenta), intranomer contacts (light orange), and contacts with other histones (dark blue). The same color code is used in the histogram below the structure, where we show the amount of the three interaction types at each location. Note that we use different y axis for these interactions: The y axis to the left corresponds to the H3.3-DNA binding contacts (magenta bars), and the y axis to the right corresponds to the intranomer (light orange bars) and intermonomer contacts (dark blue bars).
Dysregulation of PTMs is distinct for each mutation

To quantify this histone PTM dysregulation, patients’ Epstein-Barr virus–transformed lymphoblasts (H3F3A p.A15G, H3F3A p.R17G, H3F3A p.T45I, H3F3B p.A29P, and H3F3B p.P121R) and primary fibroblasts (H3F3A p.R17G, H3F3A p.G90R, and H3F3A p.T45I) were obtained from patients and age- and ethnicity-matched controls. Histones were extracted and analyzed by nano–liquid chromatography–tandem mass spectrometry (LC-MS/MS) as previously described (15). This allowed for a rigorous and robust quantification of histone PTM levels, providing insight into the global epigenetic state in each cell type. For reference, the average of the control lymphoblasts is depicted in Fig. 4A. Notably, analysis of coefficients of variation both within and across patients and controls revealed that the distribution of histone PTM abundances was very similar (Fig. 4B). Although these data show modest changes compared to cells that express p.K27M- or p.K36M-mutant H3.3, which display a near complete loss of di- and trimethylation at K27 or K36, a dominant negative effect cannot be ruled out (16).

It was also observed that the overall histone PTM variation was slightly greater in healthy donors than in patients. Nonetheless, some histone PTMs were reproducibly altered when comparing patients to controls (Fig. 4D). Although all patients share a common phenotype of developmental delay, only some of them developed major congenital malformations (i.e., cardiac and cranial anomalies). Further study of the specific local dysregulation during development may lead to insights into the transcriptional control in these processes (i.e., cardiac and cranial development). In general, the magnitude of the changes was modest and may reflect the cell types studied and developmental timing.

The tails of H3.3 and the other H3 histones (1 to 89 amino acids) are identical except for residue 31, where they contain alanine and serine, respectively. Because peptides are generated by cleavage at arginine residues, it is technically impossible to assign PTMs to a specific H3 histone, except in the case of p.A29P. For this case, two PTM abundances (K27 and K36) were compared between (i) protein transcripted from the mutant p.A29P allele from patient cells, (ii) protein transcripted from the wild-type (WT) allele from the same patient cells, and (iii) protein transcripted from the two WT alleles from a control (Fig. 4E). These data show notable local deregulation of PTMs on the mutant protein, while the WT protein from the affected patient shows fewer differences from the control. This is a more minor effect than that of published somatic oncogenic mutations; however, the difference neither confirms nor refutes a dominant negative effect (17).

Together, these data suggest that the mutant histones can be incorporated into the nucleosome, cause marked local deregulation of chromatin state, and modestly alter the global control of histone modifications. Of greatest interest are the local chromatin changes induced by mutant histone deposition: H3.3 is known to have roles in diverse functions, including gene expression and repression, chromatin stability, DNA damage repair, and differentiation. These mutant proteins and their aberrant PTM states could disrupt any of these processes to lead to the observed phenotype.

Dysregulation of transcripts associated with proliferation and mitosis

To evaluate which biological pathways were differentially perturbed in the patients, we performed RNA sequencing (RNA-seq) on fibroblast cells derived from patients and age- and ethnic-matched controls. All reads were aligned with the STAR (Spliced Transcript Alignment to a Reference) aligner, and Cufflinks was used for performing differential expression analysis. In fibroblast cells (three mutations listed above versus three controls), the H3.3 transcripts (H3F3A and H3F3B) contribution to total histone H3 expression ranged from 45.1 to 72.6% in cases and 64.1 to 81.7% in controls (table S2). We found 323 genes to be differentially expressed with at least two fold changes in fibroblast cells (P < 0.05; table S3). Of these 323 genes, 166 were up-regulated, and 157 were down-regulated in cases. Differentially expressed genes were analyzed through David Functional Annotation Resource, and we found no significant biological signal in the genes with lower expression in cases but showed significant enrichment for up-regulated genes important in mitotic cell cycle process [Benjamini-Hochberg–corrected P (BH-corrected) = 7.8 × 10^{-11}], mitotic nuclear division (P_{BH-corrected} = 5.8 × 10^{-10}), cell division (P_{BH-corrected} = 7.5 × 10^{-10}), and many other mitosis-related processes (table S4).

To assess whether up-regulation of mitosis-related genes alters cell proliferation, we quantified the cellular proliferation capacity
Fig. 4. qMS analysis of patient samples. (A) Average profile of PTMs on canonical histones H3 and H4 across control lymphoblasts. Error bars represent SD (n = 9 donors; 3 biological replicates each). (B) Tukey boxplot depicting the coefficients of variation of 73 modified histone H3 and H4 peptides detected by nano–LC-MS/MS (biological variance: across all 14 donors; patient variance: across five patients; control variance: across nine controls). (C) Average histone H3.3 protein abundance (relative to total histone H3) in patient and control lymphoblasts. Error bars represent SD. (D) Volcano plot demonstrating significantly altered histone PTMs in patients versus controls. Dotted line represents P < 0.05 significance threshold. (E) K9 and K14 PTM abundances were compared between (i) protein transcribed from the mutant p.A15G allele from patient cells, (ii) protein transcript from the wild-type (WT) allele from the same patient cells, and (iii) protein transcribed from the WT alleles from a control. Note that the peptide from amino acids 9 to 17 is indistinguishable between canonical H3 and H3.3, so the WT peptide encompasses both. (F) A29P is the only mutation occurring on the same peptide that distinguishes H3.3 from H3. PTMs that fall on this peptide are compared across the mutant peptide, the WT peptide from the mutant sample, and the average profile of the peptide from control samples. *P < 0.05, **P < 0.01, and ***P < 0.001. This shows notable local deregulation of PTMs on the mutant peptide. qMS, quantitative mass spectrometry.

of five patient fibroblast lines compared to six age- and sex-matched control fibroblast lines. Patient lines had increased cell proliferation, notably at 72 and 96 hours (Fig. 5A). Furthermore, all five patient lines shared similar viability to the six control lines (Fig. 5B). Cell cycle analyses showed that H3F3A p.G90R and H3F3A p.T45I had a similar cell cycle profile to the control lines, while H3F3A p.R17G showed a decrease in cells in G1 phase and an increase in cells in S phase compared to all three control lines (Fig. 5C).

Deficit in neural crest–derived cells in a zebrafish model

A dominant zebrafish model (p.D123N) derived from a forward genetic screen has been previously reported with craniofacial abnormalities, which were replicated in our patient cohort (Fig. 6A) (18). This heterozygous missense variant replicates the dominant inheritance observed in humans and is only two amino acids away from a mutation identified in an affected patient (p.Q125R). Further investigation of this model also reveals a defect in foxd3-positive neural crest–derived glia, as well as melanocytes and xanthophores (Fig. 6, B to D.) The loss of glial cells may relate to the hypomyelination phenotype that is noted on the brain MRIs of over one-third of the H3.3 cohort as glial cells are the cell type responsible for myelination.

DISCUSSION

The discovery of missense mutations in H3.3 that cause a neurodevelopmental disorder, but not cancer, has profound implications for future research in histone biology. Until now, H3.3 mutations have only been directly linked to cancer (6). These specific mutations that cause a neurodevelopmental phenotype are distinct from the ones that cause cancer. The cancer mutations are mostly in lysines in the histone tail, while the only lysine mutation in our cohort is p.K36E. When this lysine-36 is mutated in cancer, it is mutated to methionine and not glutamate (19). The patient with the p.K36E mutation was 32 years old when he was last evaluated and still had no signs of cancer, suggesting that the p.K36E mutation does not lead to a strongly increased risk of malignancy. The only other variant in our cohort that is similar to a cancer-linked variant is the p.G34V variant in H3F3B. This p.G34V variant in H3F3A has been shown to cause cancer only when combined with mutations in ATRX/DAXX (ATRX Chromatin Remodeler/Death Domain Associated Protein) and TP53 (TAR-DNA binding protein-43) (19). Understanding why the somatic p.G34V variant causes cancer, but germline variants cause a neurodevelopmental phenotype, is currently unknown and requires further study.

Notably, although the mutations are spread throughout the coding sequence, they all converge on a similar phenotype. Each mutation affects different specific interactions but converge on the same phenotype. This suggests that the phenotype is due to dysregulation of H3.3 in general and not limited to a single mechanism. Molecular modeling revealed that mutations in the histone core likely affect nucleosome stability, while the mutations in the tail affect various protein-protein interactions. In addition to molecular modeling data, previous work in other species such as yeast can provide insights
Fig. 5. Cellular dynamics of patient fibroblasts. (A) Five $H3F3A/B$ patient fibroblast lines ($H3F3B$: p.G34V; $H3F3A$: p.R17G; $H3F3A$: p.G90R; $H3F3A$: p.T45I; and $H3F3B$: p.V117V or p.S146X in alternate transcript) demonstrated increased proliferation over six matched controls. **$P < 0.005$ and ***$P < 0.0005$. Data represent means ± SEM of three biological replicates using three technical replicates each. (B) The same five $H3F3A/B$ patient fibroblasts and six controls show no major differences in cell viability. The data represent the means ± SEM of four biological replicates using two technical replicates each. (C) Cell cycle analysis showed differences in the S ($P = 0.0127$) and G2 ($P = 0.0338$) phase in the same five patient cell line compared to the six control fibroblast lines. Data represent the means ± SEM of four biological replicates using two technical replicates each.
into the pathogenic mechanism of these variants. Of particular interest, two of the five amino acids that differentiate H3.3 from canonical H3 (p.S31 and p.G90) are mutated in our patient cohort. Both p.S31 and p.G90 are essential for proper recognition of H3.3 by other proteins. Mutagenic analysis in yeast shows that mutations at p.G90 prevent H3.3-specific chaperones DAXX and UBN1 (Ubinucin 1) from binding (20). The serine at position 31 is required for recognition of H3.3 by ZMYND11 (21). Mutations in ZMYND11 cause autosomal dominant neurodevelopmental phenotype similar to that seen in our patient cohort, including hypotonia, seizures, dysmorphic facial features, and developmental delay (22). A second mutation in our patient cohort, p.G34V, has also been shown to disrupt ZMYND11 binding to H3.3 (21). Other variants may disrupt histone octamer formation, nucleosome sliding, and chaperone binding based on mutagenic analysis of both H3 and H3.3 in model organisms (23–25). We hypothesized that additional missense mutations in our cohort induce epigenetic dysregulation of histone PTMs. These histones’ PTMs within the nucleosome affect chromatin state, mitotic initiation, protein-chromatin interactions, and gene expression (26–30). Specific PTMs of unincorporated histones also mediate chaperone recognition before incorporation into the nucleosome (31).

The fact that all of these mutations are heterozygous is particularly noteworthy. Since there are two genes that produce H3.3, a heterozygous mutation in one of them means that 75% of the alleles are WT. A missense mutation in only 25% of the alleles is sufficient to cause global developmental delay, hypomyelination, cortical atrophy, and craniofacial anomalies. H3F3A and H3F3B are not expressed at the same levels so the amount of mutant protein could be more or less than 25% in any individual patient or tissue at any particular time; however, we observe the same phenotype regardless of whether the mutation is in H3F3A or H3F3B so a higher expression level from one of the genes cannot explain the severity of the phenotype. The fact that a single missense mutation in one allele causes such a severe phenotype strongly supports the importance of tight regulation of histones and histone PTMs. Furthermore, since our studies showed that the changes in the PTMs are in cis on the mutant H3.3 and not in trans on the WT H3.3 in the same nucleosome as is observed in the cancer causing mutations, the severity of the phenotype cannot be explained by a global dysregulation of PTMs.

Although the exact mechanism of the cellular pathology in these patients is unclear, H3.3 is vital for normal neurologic functioning. A recent study showed that H3.3 begins to replace H3.1 and H3.2 in postnatal mouse and human brains in a time-dependent manner and displaces these canonical H3 variants almost completely in adulthood. The important role of H3.3 over time may explain the unique neuregenerative phenotype, as mice with decreased H3.3 expression in the hippocampus have impaired long-term memory (32). Humans with major depressive disorder have increased percentages of H3.3 in the nucleus accumbens, which is modulated by antidepressant therapy (33). This suggests that H3.3 modulators may represent potential targeted therapies for H3.3-related disorders and perhaps associated neuropsychiatric conditions in the general population. Future studies in patient-derived induced pluripotent stem cell neural cell types may further elucidate the underlying pathology.

Although these are the first germline variants associated with histone H3, germline variants in histones H1 and H4 with similar features have been reported. The overgrowth and neurodevelopmental delay associated with Rahman syndrome (MIM 617537) are caused by truncating variants in H1 Histone Family, Member E (HIST1H1E), (34) Recently, two specific germline variants in histone H4, which caused delayed growth and neurodevelopment, have been described in two families (35). In addition, there are many neurodevelopmental disorders associated with the histone lysine methylases and demethylases. These histone-related disorders provide a unique window into the role of histones in the control of development and growth.

In conclusion, we show here that heterozygous de novo missense variants in H3F3A and H3F3B, coding for H3.3, are associated with a previously undescribed phenotype of developmental delay,
neurodegeneration, epilepsy, facial dysmorphism, and congenital anomalies. The functional effects of these mutations appear to be different from those that are well-studied in cancer and may offer a target for therapy for these and other patients.

MATERIALS AND METHODS
The Institutional Review Boards of Columbia University, the University of Michigan, Children’s Hospital of Eastern Ontario, University Medical Center Hamburg-Eppendorf, and the Children’s Hospital of Philadelphia approved this study. Informed consent was obtained from all individual participants included in the study. Additional informed consent was obtained from all individual participants for whom identifying information is included in this article.

Exome and genome sequencing
Genomic DNA was extracted from whole blood from the affected children and their parents. Exome or genome sequencing was performed with a variety of standard capture kits and the general assertion criteria for variant classification following ACMG/AMP (American College of Medical Genetics and Genomics/Association for Molecular Pathology) guidelines (36). There were no other variants in these patients that survived filtration and analysis using either dominant or recessive models and could explain the phenotypes.

Patient and control fibroblast cell lines
Six healthy control cells lines that were matched to available patient cells for passage number, age, and sex were obtained from the Coriell Institute for Medical Research tissue bank. All these fibroblast samples were derived from skin biopsies performed either on the arm or on the leg of the patient. Only a portion of the patients who participated in our study agreed to donate cellular material, and the patient mutations that were analyzed reflect the sum of the viable donations.

Histone derivatization and PTM analysis by nano–LC-MS
Histones were purified in acid as previously described by Karch et al. (37). Acid-extracted histones (15 to 25 μg) were resuspended in 100 mM ammonium bicarbonate (pH 8), derivatized using pion-ionic hydride, and digested with trypsin as previously described (37). The resulting histone peptides were desalted using the C18 Stage Tips, dried using a centrifugal evaporator, and reconstituted using 0.1% formic acid in preparation for nano–LC-MS analysis.

Nano-LC was performed using the Thermo Fisher Scientific Easy nLC 1000 System equipped with a 75 μm by 20 cm in-house packed column using Reprosil-Pur C18-AQ (3 μm; Maisch GmbH, Germany). Buffer A was 0.1% formic acid, and buffer B was 0.1% formic acid in 100% acetonitrile. Peptides were resolved using a two-step gradient from 0 to 26% B over 45 min and then from 26 to 98% B over 10 min at a flow rate of 300 nl/min. The high-performance LC was coupled online to an Orbitrap Elite mass spectrometer operating in the positive mode using a Nanospray Flex ion source (Thermo Fisher Scientific) at 2.40 kV. MS was performed using data-independent acquisition (DIA) as previously described with slight modifications (37). Briefly, two full MS scans [mass/charge ratio (m/z), 300 to 1000] were acquired in the Orbitrap with a resolution of 120,000 (at 200 m/z) every eight DIA MS/MS events using isolation windows of 50 m/z each (e.g., 300 to 350, 350 to 400... 650 to 700). The full MS scan is performed twice within the same duty cycle to allow for a more resolved definition of the precursor peak profile. MS/MS events were acquired in the ion trap operating in normal mode. Fragmentation was performed using collision-induced dissociation in the ion trap mass analyzer with a normalized collision energy of 35. Automatic gain control (AGC) target and maximum injection time were 10 × 10^6 and 50 ms for the full MS scan and 10 × 10^6 and 150 ms for the MS/MS scan, respectively.

Data were searched using EpiProfile. The peptide relative ratio was calculated using the total area under the extracted ion chromatograms of all peptides with the same amino acid sequence (including all of its modified forms) as 100%. For isobaric peptides, the relative ratio of two isobaric forms was estimated by averaging the ratio for each fragment ion with different mass between the two species. Two to three biological replicates were analyzed per condition, and the relative abundance of each peptide modification was averaged across the runs.

Western blot validations
Isolated histones were separated on a 4 to 12% NuPAGE bis-tris gel and transferred to a 0.2-μm nitrocellulose blotting membrane. Membranes were blocked in TBS-T [50 mM tris-HCl (pH 7.5), 150 mM NaCl, and 0.1% Tween 20] containing 5% milk powder and probed for H3.3 (Thermo Fisher Scientific, MA5-24667; 1:5000), H3 (Abcam, ab1791; 1:5000), and H4 (Abcam, ab7311; 1:5000) primary antibodies diluted in TBS-T containing 5% milk powder. This was followed by anti-rabbit horseradish peroxidase (Abcam, ab99697; 1:10,000) and visualized using enhanced chemiluminescence prime Western blotting detection chemiluminescent reagent.

RNA-seq expression analysis
RNA was extracted using the RNeasy Mini Kit (QIAGEN) on the six samples at the same time. Extracted RNA samples underwent quality control assessment using Bioanalyzer (Agilent, Santa Clara, CA, USA) and were quantified using NanoDrop from NanoDrop Technologies (Wilmington, DE, USA). RNA libraries were prepared using the Illumina TruSeq RNA sample prep V2 with ribosomal RNA depletion and were sequenced using HiSeq 2500 sequencer (Illumina Inc., San Diego, CA, USA) at the Center for Applied Genomics at the Children’s Hospital of Philadelphia per standard protocols (paired-end 100 base pairs). The RNA-seq data were aligned on the hg19 reference genome using STAR (www.encodeproject.org/software/star/) and processed using Cufflinks (http://cole-trapnell-lab.github.io/cufflinks/). For each gene, we compared the expression levels between cases and controls. A gene was considered differentially expressed if the P value is less than 0.05 and fragments per kilobase of transcript per million mapped reads has at least twofold change. To identify overrepresented functional categories among genes that are differentially expressed, we performed annotation analysis using the David Functional Annotation Resource (https://david.ncifcrf.gov/). Multiple testing was adjusted using the Benjamini-Hochberg approach, and enrichment was declared if the adjusted P value is less than 0.05.

Cellular analysis
Cellular proliferation was assayed with five patients and six control fibroblast cell lines; they were plated at 3 × 10^4 cells per well and then were manually counted at baseline, 24, 48, 72, and 96 hours. Then, the means ± SEM of three biological replicates using three technical replicates each were analyzed. Cell viability was analyzed with three patients and three control fibroblast cell lines; the cells were plated and grown until ~80% confluent. Then, they were stained for annexin V/propidium iodide (PI), and fluorescence-activated cell
sorting (FACS) was performed on Accuri C6 (BD Biosciences). Then, the means ± SEM of four biological replicates using two technical replicates each were analyzed. Cell cycle analysis was performed with three patients and three control fibroblast cell lines; the cells were plated and grown until ~80% confluent. They were then stained for PI, and FACS was performed on Calibur and analyzed by FlowJo. Then, the means ± SEM of four biological replicates using two technical replicates each were analyzed.

Zebrashf analysis
All zebrashf experiments were approved by the University of Southern California Institutional Animal Care and Use Committee (protocol #20258). h3F3aR28H mutations were genotyped, and control mCherry or dominant-negative D123N H3F3a mRNAs were prepared and injected at 900 ng/μl into one-cell stage embryos as described (18). Acid-free cartilage staining with Alcian Blue was performed as described (38). In situ hybridization was performed with the probes and procedures described in (18). Images were captured on a Leica DM2500 compound microscope.

Computational analysis of structural location of the variants
For each H3.3 variant location, we analyzed its pattern of interatomic contacts (see below) using experimental structural information. The list of known structures was retrieved from the UniProt record of H3.3 (P84243). In all cases, the structure of H3.3 was complete, that is, only a fragment in complex with other proteins is described. For each variant, we identified the Protein Data Bank (PDB) entries that provided information about its location. We provide a table of the PDB files used in table S5.

Computing interactions at the native locus
For each variant, we computed the network of interatomic interactions of the native residue. This network was obtained following a simple protocol: (i) retrieve the corresponding PDB, (ii) compute the network of interactions for all the protein residues using Ring, (iii) extract the interaction data corresponding to the native residue, and (iv) organize these interaction data into three groups [intramonomer (within the same H3.3 monomer), intermonomer (between H3.3 monomer and other proteins), and H3.3-DNA] (39). Intra-

SUPPLEMENTARY MATERIALS
Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/eabc9207/DC1

REFERENCES AND NOTES
1. Z. Zhao, A. Shiltatifard, Epigenetic modifications of histones in cancer. Genome Biol. 20, 245 (2019).


Acknowledgments: We thank the families for contribution. We also thank J.F. Deleuze, A. Boland, and V. Meyer from the CNRS/GB for genome sequencing and data primary analysis. Funding: Funding was provided by the French Foundation for Rare Diseases (Fondation maladies rares). This work was also supported in part by a grant from the sm FAIR (to W.K.C.), JPB Foundation (to W.K.C.), and the Morton S. and Henrietta K. Sellner Professorship in Human Genetics (to J.W.J.), as well as an NIDDK T35 training grant (T35DK099340) (to E.F.J.). E.J.B. was supported by a K12 training grant (K12HD043245-14) and the Roberts Collaborative. D.M. was supported by a T32 training grant (T32GM008275). Funding from NIH grants GM110174 and CA196539 to B.G. is acknowledged. This research was funded in part by the Estonian Genome Foundation grant PUT0355 and PR0471. Analysis for one patient was provided by the Broad Institute of MIT and Harvard Center for Mendelian Genomics (Broad CMG) and was funded by the National Human Genome Research Institute, the National Eye Institute, and the National Institute of MIT and Harvard Center for Mendelian Genomics (Broad CMG) and was funded by the American College of Genomics and Genomics Institute, Ontario Research Fund, Genome Quebec, and Children’s Hospital of Eastern Ontario Foundation. This work was also supported by the Instituto de Biologia Humana y de la Salud (CSIC) and the Instituto de Biologia Molecular y Celular de la日凌晨, 2021 http://advances.sciencemag.org/ Downloaded from
Histone H3.3 beyond cancer: Germline mutations in *Histone 3 Family 3A and 3B* cause a previously unidentified neurodegenerative disorder in 46 patients


Sci Adv 6 (49), eabc9207.
DOI: 10.1126/sciadv.eabc9207

ARTICLE TOOLS
http://advances.sciencemag.org/content/6/49/eabc9207

SUPPLEMENTARY MATERIALS
http://advances.sciencemag.org/content/suppl/2020/11/30/6.49.eabc9207.DC1

REFERENCES
This article cites 39 articles, 6 of which you can access for free
http://advances.sciencemag.org/content/6/49/eabc9207#BIBL

PERMISSIONS
http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service