nature genetics

Germline mutations affecting the histone H4 core cause a developmental syndrome by altering DNA damage response and cell cycle control

Federico Tessadori[1], Jacques C Giltay[2,3], Jane A Hurst[4], Maarten P Massink[2,3], Karen Duran[2,3], Harmjan R Vos[3], Robert M van Es[3], Deciphering Developmental Disorders Study[5], Richard H Scott[4], Koen L I van Gassen[2,3], Jeroen Bakkers[1,6] & Gijs van Haaften[2,3]

Covalent modifications of histones have an established role as chromatin effectors, as they control processes such as DNA replication and transcription, and repair or regulate nucleosomal structure[1,2]. Loss of modifications on histone N tails, whether due to mutations in genes belonging to histone-modifying complexes or mutations directly affecting the histone tails, causes developmental disorders[3–6] or has a role in tumorigenesis[7,8]. More recently, modifications affecting the globular histone core have been uncovered as being crucial for DNA repair, pluripotency and oncogenesis[9,10]. Here we report monoallelic missense mutations affecting lysine 91 in the histone H4 core (H4K91) in three individuals with a syndrome of growth delay, microcephaly and intellectual disability. Expression of the histone H4 mutants in zebrafish embryos recapitulates the developmental anomalies seen in the patients. We show that the histone H4 alterations cause genomic instability, resulting in increased apoptosis and cell cycle progression anomalies during early development. Mechanistically, our findings indicate an important role for the ubiquitination of H4K91 in genomic stability during embryonic development.

We performed sequencing on two unrelated individuals with developmental disorders without a genetic diagnosis from the UK Deciphering Developmental Disorders (DDD) study[11–13] and UMC Utrecht diagnostic services. The subjects presented with a similar array of phenotypes (**Fig. 1a**, **Table 1** and case reports in the **Supplementary Note**), comprising short stature, microcephaly, intellectual disability, characteristic facial features and foot ray anomalies. Trio-based whole-exome sequencing identified heterozygous missense mutations in the gene encoding *HIST1H4C* (ENSG00000197061) in both individuals, located at adjacent nucleotides and affecting the same amino acid, namely K91. In patient 1, a *de novo* mutation, c.274A>C, results in p.Lys91Gln, while in patient 2 an apparently *de novo* c.275A>G mutation results in a p.Lys91Arg substitution (**Fig. 1c** and **Supplementary Fig. 1**). After obtaining these findings, we became aware of an older sibling of patient 2 (hereafter referred to as patient 3; **Fig. 1b** and **Table 1**) who died shortly (11 d) after birth. The limited information available for patient 3 suggested that the phenotype for this individual overlapped remarkably well with those of the other two affected individuals (**Table 1**). Sequencing of a DNA sample from patient 3 confirmed the presence of the same *HIST1H4C* c.275A>G mutation as in patient 2. Interestingly, a low level of mosaicism for this mutation was detected in the father of patients 2 and 3 (**Supplementary Fig. 2**), explaining the occurrence of the same mutation in two siblings.

The human genome contains 15 histone H4 genes, all differing at the nucleotide level but encoding an invariant histone H4 protein (**Supplementary Table 1**). Histone H4 is extremely well conserved across kingdoms (**Supplementary Fig. 3**). Specifically, K91 and the amino acids surrounding it at the C-terminal end of the sequence show no variation. In human adults, *HIST1H4C* is reported to be expressed in a large variety of tissues, yet at a relatively low rate (EMBL-EBI Expression Atlas (see URLs); **Supplementary Fig. 4**), whereas the corresponding genes are well expressed in very early developmental stages of mouse[14] and zebrafish embryos[15].

We tested the p.Lys91Gln and p.Lyr91Arg alterations for dominant effects on development, along with five other possible disease-causing mutations identified in the DDD study in three different human histone H4 genes (*HIST1H4C* p.Lys79Cys; *HIST1H4E* (ENSG00000276966) p.Arg40Cys, p.Arg45Cys and p.Asp67Glu; *HIST1H4I* (ENSG00000276180) p.Arg40Leu), by expressing synthetic mRNA in zebrafish embryos (**Fig. 2**). Injection with wild-type human histone H4 mRNA for *HIST1H4C*, *HIST1H4E* or *HIST1H4I* had no visible effect on structural embryo development, nor did injection of mRNA encoding the p.Arg40Cys, p.Arg45Cys, p.Asp67Glu and p.Lys79Cys substitutions in histone H4, while injection with mRNA

[1]Hubrecht Institute–KNAW and University Medical Center Utrecht, Utrecht, the Netherlands. [2]Department of Genetics, University Medical Center Utrecht, Utrecht, the Netherlands. [3]Center for Molecular Medicine, University Medical Center Utrecht, Utrecht, the Netherlands. [4]North East Thames Regional Genetics Service, Great Ormond Street Hospital, London, UK. [5]A list of members and affiliations appears in the **Supplementary Note**. [6]Department of Medical Physiology, Division of Heart and Lungs, University Medical Center Utrecht, Utrecht, the Netherlands. Correspondence should be addressed to G.v.H. (g.vanhaaften@umcutrecht.eu) or J.B. (j.bakkers@hubrecht.eu).

Received 20 February; accepted 23 August; published online 18 September 2017; doi:10.1038/ng.3956

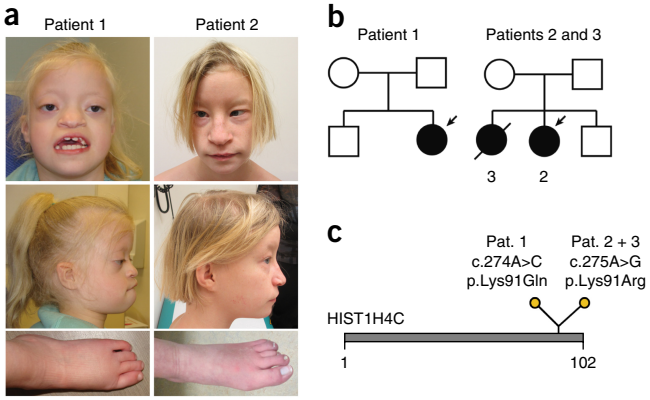

Figure 1 Clinical and molecular defects in patients with *HIST1H4C* mutations. (**a**) Frontal, lateral and foot views of patient 1 and patient 2. A detailed description of the clinical features is presented in **Table 1** and in the **Supplementary Note**. Informed consent to publish clinical photographs was obtained for subjects 1 and 2. (**b**) Succinct pedigrees of the patients belonging to this study. The patients reported are indicated with an arrow. Affected individuals are represented by a filled symbol. (**c**) Schematic of the linear histone H4 protein encoded by *HIST1H4C*. The positions of the missense variants carried by the patients are marked. The HGVS nomenclature for the substitutions is p.Lys92Gln and p.Lys92Arg. This manuscript bridges the fields of human genetics and molecular biology. For histones, the N-terminal methionine residue incorporated during translation initiation is excised, resulting in a −1 shift in amino acid number. Throughout this manuscript, the lysine is thus counted as residue 91, consistent with the molecular biology literature.

encoding p.Arg40Leu resulted in only mildly defective development in a low percentage of embryos. In contrast, expression of mRNA encoding either p.Lys91Gln or p.Lys91Arg severely affected the structural development of zebrafish embryos (**Fig. 2a–d**). An underdeveloped brain and eyes, defective body axis growth and a dysmorphic tail were observed in these mutants, as well as significant reduction of body axis length at 2 days post-fertilization (d.p.f.), matching the microcephaly and short stature seen in the patients in this study. Analysis of the embryos at the ten-somite stage (**Supplementary Fig. 5**) showed that these phenotypes were preceded by convergence-extension defects, which may be related to the facial features observed in the patients (**Table 1** and case reports in the **Supplementary Note**) as well as their short stature. The presence of two siblings (patients 2 and 3) from unrelated Dutch parents and another independent patient (patient 1) displaying similar clinical phenotypes, in combination with *in vivo* modeling evidence in zebrafish embryos, indicates that the substitutions affecting HIST1H4C K91 cause the abnormal phenotypes in patients.

To get more insight into the cellular effects of the mutations, we carried out RNA sequencing on primary fibroblasts from patients 1 and 2, as well as on a control fibroblast line. Expression levels of the mutated *HIST1H4C* alleles were comparable to those of the wild-type allele. Considering all sequenced reads of histone H4 genes, we estimate that about 8% of the histone H4 cDNA molecules contained the mutation (**Supplementary Fig. 6**). To address whether the mutant histone H4 protein is also incorporated into nucleosomes, we used high-resolution liquid chromatography coupled with tandem mass spectrometry (LC–MS/MS) to quantify in a label-free manner the ratio of each mutant histone to its wild-type counterpart (**Supplementary Fig. 7**). This analysis suggested that mutant protein is present in nucleosomes at a level of 1–2%. RNA sequencing identified 115 genes that were differentially expressed between the *HIST1H4C* mutation carriers and the control. Of these genes, 38 were expressed at higher levels and 77 were expressed at lower levels in the *HIST1H4* variant carriers as compared to the control, including 25 histone genes and numerous cell cycle–related genes (**Supplementary Table 2**). Furthermore, gene ontology terms relating, among other processes, to chromosome organization, histone binding, DNA packaging, nucleosomal organization and the cell cycle process were highly enriched among the differentially expressed genes (**Supplementary Table 3**), suggesting that the mutations have an effect on DNA replication or cell cycle progression.

Modifications affecting the histone core, such as acetylation, ubiquitination, methylation and phosphorylation, among others, are thought to have a variety of structural of functional roles, including in organization of the nucleosome, chromatin compaction and DNA replication[2]. The occurrence of post-translational modifications on H4K91 is highly conserved across species and has been associated with processes ranging from chromatin assembly and DNA damage sensitivity[16,17] to gametogenesis[18,19].

As a result of the amino acid substitutions, post-translational marks on K91 were absent on the fraction of mutated histone H4 molecules in patient-derived cells. To further comprehend the impact of the p.Lys91Gln and p.Lys91Arg alterations on genomic integrity, we fluorescently immunostained zebrafish embryos expressing the H4K91 mutants for γ-H2AX, which is deposited initially at the sites of DNA double-strand breaks (DSBs; **Fig. 2e**, representative pictures shown

Table 1 Genetic and clinical features of the patients in this study

Patient		1	2	3
HIST1H4C alteration	DNA	c.274A>C	c.275A>G	c.275A>G
	Protein	p.Lys91Gln[a]	p.Lys91Arg[a]	p.Lys91Arg[a]
Age at observation		7 years	13 years	11 d
Clinical features	Growth retardation	+	+	+
	Intellectual disability	+	+	NA
	Microcephaly	+	+	+
	Hypotonia	+	+	–
	Upslanting palpebral fissures	+	+	+
	Bifid flat nasal tip	+	+	+
	Median ridge on philtrum	+	+	?[b]
	Wide mouth	+	+	?[b]
	Foot ray anomaly	+[c]	+[d]	+[e]

NA, not applicable (patient died at 11 d from massive thrombosis).

[a]See the remarks on amino acid numbering in the legend of **Figure 1**. [b]Not documented as the patient was not seen by a clinical geneticist. [c]Absent postaxial digit (unilateral). [d]Proximal implantation fourth ray (bilateral). [e]Short fourth ray on the right foot.

for wild-type and K91Q injections). Quantification of γ–H2AX-positive cells identified a significant increase in the accumulation of DSBs with both mutations, with these breaks especially visible in the heads and tails of the larvae (**Fig. 2f**). Corroborating an accumulation of DSBs, embryos expressing the K91R and K91Q mutants also displayed increased cell apoptosis in the head and tail, as shown by extensive acridine orange staining and increased nuclear fragmentation (**Supplementary Fig. 8**).

Acetylation and ubiquitination of H4K91 regulate chromatin assembly[17,20] and DNA damage response[16], respectively. Perturbation of

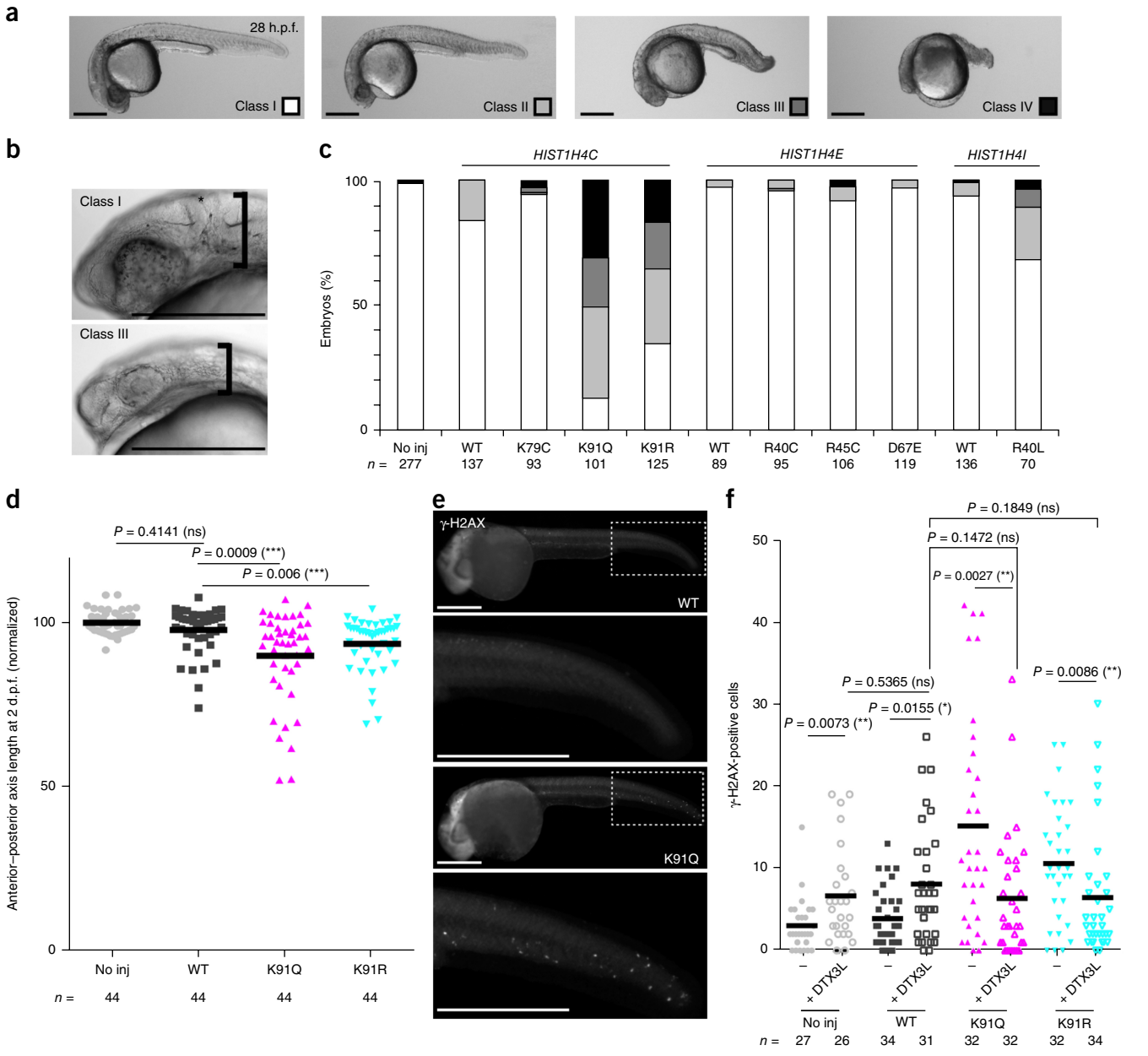

Figure 2 Histone H4 variants (encoded in *HIST1H4C*) induce defective growth, microcephaly and double-strand break accumulation in zebrafish embryos. (**a**) Representative pictures of observed phenotypes in zebrafish embryos 28 hours post-fertilization (h.p.f.) microinjected with *HIST1H4C* mRNA variants at the one-cell stage. Note the shortening of the anterior–posterior axis and defective head development in affected embryos. (**b**) Detailed view of the observed microcephaly phenotype. Note the absence of the midbrain–hindbrain boundary in class II embryos (boundary indicated by an asterisk in the class I embryo) and reduced brain size (bracket). (**c**) Quantification of the phenotypes reported in **a**. WT, wild type. (**d**) Quantification of the embryo anterior–posterior axis length at 2 d.p.f. (**e**) Fluorescent immunolabeling in 28 h.p.f. whole-mount embryos for γ-H2AX. The lower image in each pair corresponds to a magnified view of the boxed region in the upper image. (**f**) Quantification of the occurrence of γ-H2AX-positive cells in the tail region boxed in **e**. Expression of the K91Q and K91R histone H4 (*HIST1H4C*) variants results in significant accumulation of DSBs, as revealed by increased γ-H2AX signal. Coexpression of DTX3L significantly reduces the occurrence of γ-H2AX-positive cells in K91Q- or K91R-expressing embryos and rescues the phenotype to control levels. For all graphs, significance was determined by two-tailed Mann–Whitney *U* test: $*P < 0.05$; $**P < 0.01$; $***P < 0.001$; ns, not significant. The black horizontal bar indicates the mean value for each condition. Scale bars, 100 μm. Data were collected on one technical replicate for axis length measurement (**d**) and a minimum of two replicates for γ-H2AX scoring (**f**). All embryos analyzed originated from group matings of adult zebrafish.

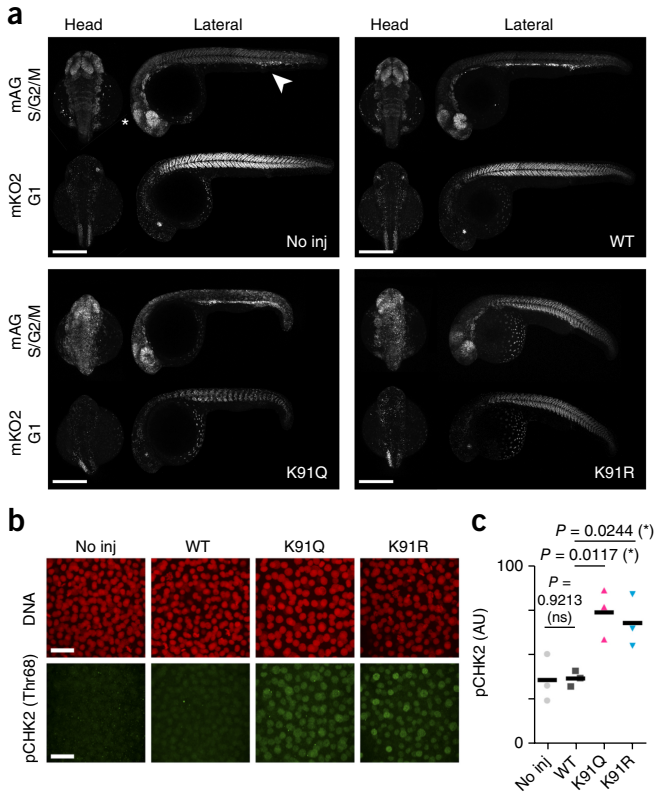

Figure 3 Histone H4 (*HIST1H4C*) alterations at K91 perturb cell cycle progression. (**a**) Representative confocal imaging (maximum projections of confocal stacks) of zFUCCI embryos (minimum of 30) expressing *HIST1H4C* mRNA encoding wild-type, K91Q or K91R histone H4 and no-injection control at 28 h.p.f. The asterisk and arrowhead indicate, respectively, the brain ventricles and localization of hematopoietic cells, which are highly proliferative at this developmental stage. Note the increased mAG fluorescence and defective organization of the tissues in these areas in the K91Q- and K91R-expressing embryos. mKO2, monomeric Kusabira orange 2; mAG, monomeric Azami green. (**b**,**c**) Embryos expressing *HIST1H4C* mRNA for K91Q or K91R histone H4 display significant upregulation of phosphorylated CHK2 (pCHK2 (Thr68); green) at late epiboly stages (5 h.p.f.). Nuclei were counterstained with TO-PRO-3 (red). For **b**, representative maximum projections of confocal stacks are shown. In **c**, for each of the three embryos imaged per condition, a minimum of 112 nuclei and a maximum of 284 nuclei were analyzed. Significance was determined by two-tailed unpaired Student's *t* test: $*P < 0.05$; ns, not significant. The black horizontal bar indicates the mean value for each condition. Scale bars, 100 µm (**a**) and 25 µm (**b**). All embryos analyzed originated from group matings of adult zebrafish. AU, arbitrary units.

H4K91 acetylation, whether resulting in constitutive or reduced acetylation levels, primarily leads to alterations in the nucleosomal structure[17,20]. Ubiquitination of H4K91 specifically protects cells against DNA-damaging agents[16]. These roles, or lack thereof in mutant-expressing cells, are consistent with the observed accumulation of γ-H2AX-positive cells in the K91R- and K91Q-injected larvae. To dissect the importance of H4K91 acetylation and monoubiquitination in the zebrafish phenotype induced by K91 alteration and better comprehend the mechanistic basis of the patients' condition, we coexpressed *HAT4*, which acetylates histone H4 at K91 (**Supplementary Fig. 9a**), or *BBAP* (also known as *DTX3L*), which is the E3 ligase responsible for H4K91 monoubiquitination[16] (**Supplementary Fig. 9b**), with wild-type, K91R and K91Q histone H4 in zebrafish embryos and quantified the presence of γ-H2AX. We hypothesized that increasing the acetylation or ubiquitination levels of endogenous wild-type histone H4 molecules would rescue the increase in DSBs resulting from K91R and K91Q expression. While coexpression of HAT4 did not relieve the occurrence of DSBs (**Supplementary Fig. 10**), in the presence of DTX3L, the number of γ-H2AX-positive cells significantly decreased (**Fig. 2f**) and the morphological anomalies were reduced (data not shown). This result indicates that perturbation of H4K91 acetylation is not the main cause of the defects described in this study. To further test the importance of H4K91 ubiquitination, we coexpressed an inactivated form of DTX3L, DTX3L-M2, which failed to rescue the γ-H2AX phenotype (**Supplementary Fig. 11**). These findings indicate that ubiquitination of H4K91 has a protective effect against DNA damage and plays a role in embryo development.

To test whether the increased accumulation of DSBs has consequences for cell cycle progression, we analyzed the effect of the p.Lys91Gln and p.Lys91Arg substitutions on cell cycle reporter zFUCCI[21] larvae. In control embryos, cells in the S, G2 or M (S/G2/M) phase were prominently present in brain ventricles and hematopoietic cells located in the tail region, whereas skeletal muscle cells were predominantly in G1 phase (**Fig. 3a**). Both K91Q- and K91R-expressing embryos displayed anomalous development and abnormal accumulation of cells in S/G2/M phase in the head and caudal regions (**Fig. 3a**), in agreement with the increases in DSBs and apoptosis observed in these regions. FACS analysis of single-cell suspensions of dissociated embryos indicated that expression of K91Q and K91R induced a relative increase in cells in S/G2/M phase (**Supplementary Fig. 12**). Corroborating these observations, K91Q- and K91R-expressing zebrafish embryos also displayed increased occurrence of activated checkpoint kinase 2 (CHK2; **Fig. 3b,c**), resulting in p53 stabilization (**Supplementary Fig. 13**). Together, these results demonstrate that expression of histone H4 mutated at K91 during embryonic development results in an increase in DSBs, leading to abnormal cell cycle progression and apoptosis, and suggest that these defects cause the syndromic phenotypes of the patients.

Numerous human disorders are associated with defective DNA damage response and are characterized by pathological *in vitro* responses to agents such as mitomycin C (Fanconi anemia) and gamma or UV radiation (for example, xeroderma pigmentosa). These agents were tested on fibroblasts from patients 1 and 2, but no abnormal responses were seen (data not shown). Yet, many of these disorders share essential features such as growth delay and microcephaly with the syndrome described here[22,23]. Interestingly, phenotypic features strikingly similar to those described in this study were caused by mutations in human *ATR*[24]. ATR has a crucial role in preventing DNA breaks at replication forks and is able to trigger cell cycle arrest through activation of the S–G2 DNA damage checkpoint[25]. This function is crucial in early embryonic stages, when the cell proliferation rate is high[26]. Mutations affecting zebrafish Stat3, which is a highly conserved cell cycle regulator, result in reduced cell proliferation, causing early convergence-extension defects in zebrafish embryos similar to those observed here (**Supplementary Fig. 5**)[27]. Intriguingly, it seems that, for ATR and Stat3, as well as histone H4 mutants at K91 encoded by *HIST1H4C*, the phenotype displayed by patients or zebrafish could ensue from delayed cell cycle progression in the very early stages of embryonic development.

Our results highlight the functional importance of the histone core and establish H4K91 and its modifications in the realm of human genetic disorders. Alteration of K91 on histone H4 acts in a genetically dominant manner. On a biological level, our data presented here point to a mechanism involving inherent DNA damage accumulation and early perturbation of the cell cycle, through which missense mutations in *HIST1H4C* affecting K91 of histone H4 are causative for an identifiable syndrome consisting of dysmorphic features and intellectual disability.

URLs. EMBL-EBI Expression Atlas, https://www.ebi.ac.uk/gxa/home.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank the subjects and their parents for being willing to be part of this study, S. van der Elst for technical assistance for the FACS analysis, P. Nguyen for the zebrafish embryonic cell suspension protocol, D. Guardavaccaro for advice on western blotting and A. de Graaff.

We acknowledge support from the Netherlands Cardiovascular Research Initiative, Dutch Heart Foundation grant CVON2014-18 CONCOR-GENES and the 'Proteins at Work' program of the Netherlands Organization for Scientific Research (NWO; project 184.032.201).

The DDD study presents independent research commissioned by the Health Innovation Challenge Fund (grant HICF-1009-003), a parallel funding partnership between the Wellcome Trust with the Department of Health and the Wellcome Trust Sanger Institute (grant WT098051). The views expressed in this publication are those of the author(s) and not necessarily those of the Wellcome Trust or the Department of Health. The study has UK Research Ethics Committee approval (10/H0305/83, granted by the Cambridge South REC, and GEN/284/12, granted by the Republic of Ireland REC). The research team acknowledges the support of the National Institute for Health Research, through the Comprehensive Clinical Research Network.

AUTHOR CONTRIBUTIONS

K.L.I.v.G., J.C.G., J.A.H. and R.H.S. identified and matched the study subjects. F.T., J.B. and G.v.H. designed the study. K.L.I.v.G. contributed to the study design. F.T. carried out the zebrafish experiments. K.D. performed molecular cloning and carried out RNA sequencing on patient-derived fibroblasts. H.R.V. and R.M.v.E. carried out the mass spectrometry analysis. M.P.M. performed the bioinformatics analysis on the RNA sequencing data. F.T., M.P.M., J.B. and G.v.H. wrote the manuscript, and F.T., J.C.G., R.H.S., J.B. and G.v.H. reviewed the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Bannister, A.J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res.* **21**, 381–395 (2011).
2. Tessarz, P. & Kouzarides, T. Histone core modifications regulating nucleosome structure and dynamics. *Nat. Rev. Mol. Cell Biol.* **15**, 703–708 (2014).
3. Koolen, D.A. *et al.* Mutations in the chromatin modifier gene *KANSL1* cause the 17q21.31 microdeletion syndrome. *Nat. Genet.* **44**, 639–641 (2012).
4. Ng, S.B. *et al.* Exome sequencing identifies *MLL2* mutations as a cause of Kabuki syndrome. *Nat. Genet.* **42**, 790–793 (2010).
5. Qi, H.H. *et al.* Histone H4K20/H3K9 demethylase PHF8 regulates zebrafish brain and craniofacial development. *Nature* **466**, 503–507 (2010).
6. Zollino, M. *et al.* Mutations in *KANSL1* cause the 17q21.31 microdeletion syndrome phenotype. *Nat. Genet.* **44**, 636–638 (2012).
7. Papillon-Cavanagh, S. *et al.* Impaired H3K36 methylation defines a subset of head and neck squamous cell carcinomas. *Nat. Genet.* **49**, 180–185 (2017).
8. Kallappagoudar, S., Yadav, R.K., Lowe, B.R. & Partridge, J.F. Histone H3 mutations—a special role for H3.3 in tumorigenesis? *Chromosoma* **124**, 177–189 (2015).
9. Lawrence, M., Daujat, S. & Schneider, R. Lateral thinking: how histone modifications regulate gene expression. *Trends Genet.* **32**, 42–56 (2016).
10. Tropberger, P. & Schneider, R. Scratching the (lateral) surface of chromatin regulation by histone modifications. *Nat. Struct. Mol. Biol.* **20**, 657–661 (2013).
11. Wright, C.F. *et al.* Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet* **385**, 1305–1314 (2015).
12. Firth, H.V. *et al.* DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am. J. Hum. Genet.* **84**, 524–533 (2009).
13. Deciphering Developmental Disorders Study. Prevalence and architecture of *de novo* mutations in developmental disorders. *Nature* **542**, 433–438 (2017).
14. Park, S.J. *et al.* Inferring the choreography of parental genomes during fertilization from ultralarge-scale whole-transcriptome analysis. *Genes Dev.* **27**, 2736–2748 (2013).
15. Yang, H. *et al.* Deep mRNA sequencing analysis to capture the transcriptome landscape of zebrafish embryos and larvae. *PLoS One* **8**, e64058 (2013).
16. Yan, Q. *et al.* BBAP monoubiquitylates histone H4 at lysine 91 and selectively modulates the DNA damage response. *Mol. Cell* **36**, 110–120 (2009).
17. Ye, J. *et al.* Histone H4 lysine 91 acetylation: a core domain modification associated with chromatin assembly. *Mol. Cell* **18**, 123–130 (2005).
18. Dai, L. *et al.* Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark. *Nat. Chem. Biol.* **10**, 365–370 (2014).
19. Govin, J. *et al.* Systematic screen reveals new functional dynamics of histones H3 and H4 during gametogenesis. *Genes Dev.* **24**, 1772–1786 (2010).
20. Yang, X. *et al.* HAT4, a Golgi apparatus–anchored B-type histone acetyltransferase, acetylates free histone H4 and facilitates chromatin assembly. *Mol. Cell* **44**, 39–50 (2011).
21. Sugiyama, M. *et al.* Illuminating cell-cycle progression in the developing zebrafish embryo. *Proc. Natl. Acad. Sci. USA* **106**, 20812–20817 (2009).
22. O'Driscoll, M. Diseases associated with defective responses to DNA damage. *Cold Spring Harb. Perspect. Biol.* **4**, a012773 (2012).
23. Ribezzo, F., Shiloh, Y. & Schumacher, B. Systemic DNA damage responses in aging and diseases. *Semin. Cancer Biol.* **37-38**, 26–35 (2016).
24. Mokrani-Benhelli, H. *et al.* Primary microcephaly, impaired DNA replication, and genomic instability caused by compound heterozygous *ATR* mutations. *Hum. Mutat.* **34**, 374–384 (2013).
25. Cimprich, K.A. & Cortez, D. ATR: an essential regulator of genome integrity. *Nat. Rev. Mol. Cell Biol.* **9**, 616–627 (2008).
26. Murga, M. *et al.* A mouse model of ATR–Seckel shows embryonic replicative stress and accelerated aging. *Nat. Genet.* **41**, 891–898 (2009).
27. Liu, Y., Sepich, D.S. & Solnica-Krezel, L. Stat3/Cdc25a-dependent cell proliferation promotes embryonic axis extension during zebrafish gastrulation. *PLoS Genet.* **13**, e1006564 (2017).

ONLINE METHODS

Clinical samples. Informed consent for whole-exome sequencing as a part of the diagnostic process (approved by the Medical Ethical Committee of the University Medical Center Utrecht) was obtained for subject 2 and her parents. Whole-exome sequencing for patient 1 has been performed within the framework of the DDD study[13]. Separate consent was obtained for patients 1 and 2 for publication of patient photographs.

Next-generation sequencing. After referral for routine diagnostic exome sequencing, the exomes of patient 2 and her parents were enriched using the SureSelect XT Human All Exon v5 kit (Agilent) and sequenced in rapid run mode on the HiSeq 2500 sequencing system (Illumina) at a mean target depth of 100×. Reads were aligned to hg19 using BWA (BWA-MEM v0.7.5a), and variants were called using the GATK haplotype caller (v2.7-2). For patient 1, please refer to the DDD study[13].

Fish lines and husbandry. Zebrafish (*Danio rerio*) were kept in standard laboratory conditions[28]. The zebrafish lines used in this study were Tübingen longfin (wild type) and Cecyil2 (Tg(EF1α:mKO2-zCdt1(1/190))^{rw0405d} and Tg(EF1α:mAG-hGem(1/60))$^{rw0412a;21}$), referred to as zFUCCI in the rest of the manuscript. Animal experiments were approved by the Animal Experimentation Committee of the Royal Netherlands Academy of Arts and Sciences.

mRNA expression assay in zebrafish. Human wild-type cDNAs encoding HIST1H4C, HIST1H4E, HIST1H4I, HIST1H4C K91Q, HIST1H4C K91R, HIST1H4C K79Q, HIST1H4E R40C, HIST1H4E R45C, HIST1H4I R40L, HAT4 and DTX3L (BBAP) were cloned into pCS2GW by Gateway cloning (Life Technologies). For the inactive form of DTX3L, DTX3L-M2, four single-nucleotide mutations resulting in p.Cys576Ser, p.His578Ser, p.Cys581Ser and p.Cys584Ser substitutions were introduced. The resulting constructs were linearized with NotI-HF (New England BioLabs) restriction endonuclease and used as template for *in vitro* synthesis of capped mRNA with the mMESSAGE mMACHINE SP6 Ultra kit (Life Technologies). One-cell-stage embryos were microinjected with approximately 50 pg of mRNA and kept at 28.5 °C in E3 medium. *In vivo* phenotypical analysis was carried out within 28 h of fertilization.

Whole-mount fluorescent immunolabeling. Zebrafish embryos at the selected developmental stage were fixed overnight in 2% paraformaldehyde (PFA) in PBS. After washing with 1× PBS–Triton X-100 (0.1%; PBS-T) and blocking in 10% goat serum in 1× PBS-T (blocking buffer), embryos were incubated overnight at 4 °C with rabbit anti-γ-H2AX (pSer139; 1:500 dilution in blocking buffer; Calbiochem, DR1017; used in ref. 29). For p53 (rabbit anti-p53; 1:250 dilution; Genetex, GTX128135) and phosphorylated Chk2 (Thr68; rabbit anti-pChk2 (Thr68); 1:250 dilution; Cell Signaling Technology, 2661) immunolabeling, before blocking, embryos were dehydrated in 100% methanol overnight, rehydrated in PBS-T, incubated for 7 min in 100% acetone at −20 °C and rinsed in PBS-T. After multiple washes in PBS-T, the embryos were incubated overnight at 4 °C in Cy2-conjugated goat anti-rabbit antibody (1:500 dilution in blocking buffer; Jackson Immunoresearch, 111-225-144). Embryos were then washed in PBS-T before imaging. Counterstaining for assessment of nuclear fragmentation was carried out with DAPI (Invitrogen; 1:2,500 dilution in PBS-T). Nuclear counterstaining for immunolabeling was carried out with TO-PRO-3 (1:1,000 dilution; Thermo Fisher Scientific).

Western blotting. Western blotting was carried out essentially as described[30]. Whole-protein samples were prepared with 24 h.p.f. embryos for p53 and H4K91ac analysis and with early-gastrulation embryos for H4 ubiquitination analysis. Samples (30–70 µg) were loaded and run on a gel (for each experiment, the sample amount was kept the same). Incubation in rabbit anti-p53 antibody (1:500 dilution; Genetex, GTX128135), rabbit anti-H4K91ac antibody (1:1,000 dilution; Abcam, ab4627) or mouse anti-Flag antibody (1:10,000 dilution; Sigma-Aldrich, F3165) was carried out overnight at 4 °C. If necessary, membranes were subsequently stripped and incubated with a 1:2,000 dilution of mouse anti-β-actin antibody (AC-15; sc-69879, Santa Cruz) or a 1:500 dilution of rabbit anti-H4 antibody (Millipore, 07-108) for loading control. Detection was carried out with a SuperSignal West Pico ECL system (Thermo Fisher Scientific). Quantification of band intensity was carried out with ImageJ (https://imagej.nih.gov/ij/).

Acridine orange staining. Embryos were incubated in 5 µg/ml acridine orange in E3 medium for 30 min in the dark, rinsed three times for 10 min in E3 medium, sedated in 16 mg/ml tricaine in E3 medium and imaged live.

Embryo dissociation and FACS analysis. 26 h.p.f. embryos (50 per condition) were dechorionated and washed briefly in sterile Ca^{2+}- and Mg^{2+}-free DPBS (Gibco). Embryos were subsequently incubated at 30 °C with gentle shaking in 2 mg/ml collagenase II (Gibco, 17101-015) in TrypLE Express (Gibco, 12605-010; 1 ml) until fully dissociated. The cell suspension was then pipetted through a 70-µm cell strainer (Greiner Bio-One, 542 070) into a 15-ml Falcon tube, diluted with 5 ml of DPBS and spun at 2,000 r.p.m. at 4 °C for 5 min. The pellet was then resuspended in 250 µl of DPBS supplemented with DNA dye (Hoechst 34580) and strained at 40 µm before FACS analysis with a FACSAria Fusion or FACSCanto II instrument (BD Biosciences) using FACSDiva software (BD Biosciences).

Fibroblast sample preparation for liquid chromatography and mass spectrometry. The nuclear fraction from 2 million primary fibroblasts was collected using the NE-PER kit (Thermo Scientific, 78833), following the manufacturer's directions. Samples were subsequently denatured with 8 M urea in 1 M ammonium bicarbonate (ABC) and digested with 250 ng of trypsin/LysC (Promega) after dilution to 2 M urea in 1 M ammonium bicarbonate buffer. After protein digestion, peptides were bound to an in-house-generated c18 stage tip washed with buffer A (0.1% formic acid) and stored at 4 °C until LC–MS/MS analysis.

LC–MS/MS analysis. Peptides were separated on a 30-cm pico-tip column (50-µm ID; New Objective) packed in house with 1.9-µm Aquapur Gold C18 material (Dr. Maisch) using a 140-min gradient (7–80% acetonitrile in 0.1% formic acid), delivered by an easy-nLC 1000 (Thermo Scientific), and electrosprayed directly into an Orbitrap Fusion Tribrid Mass Spectrometer (Thermo Scientific). The mass spectrometer was set in data-dependent mode with a cycle time of 1 s, during which a full scan over the mass range of 400–1,500 m/z was performed at a resolution of 240,000. The most intense ions (intensity threshold of 5,000 ions) were isolated by the quadrupole and fragmented with an HCD collision energy of 35%. The maximum injection time of the ion trap was set to 50 ms.

Data analysis. Raw files were analyzed with either Proteome Discoverer 2.1 (Thermo Scientific) in combination with Sequest HT or with Maxquant (software version 1.5.2.8)[31], both with oxidations of methionine set as variable modifications and cysteine alkylation with iodoacetamide set as a fixed modification. Both search engines were used with a dedicated human database including the mutant histone. In Proteome Discoverer 2.1, the node precursor area detector was used to calculate the ratio of the summed intensities of the different specific mutant or wild-type peptides in the separate replicates.

Imaging, confocal microscopy and image quantification. *In vivo* phenotypic assessment, acridine orange staining and immunolabeling for whole-embryo imaging were carried out on a Leica M165FC stereomicroscope and an AF7000 microscope (Leica Microsystems) with transmitted light or using the appropriate wavelength filters in combination with a fluorescent light source. Images were captured with a DFC420 (M165FC) or DFC360FX (AF7000) digital microscope camera (Leica Microsystems). Confocal imaging for zFUCCI embryos and nuclear fragmentation was carried out on a Leica SPE confocal microscope (Leica Microsystems).

3D quantification of phosphorylated Chk2 (Thr68) immunolabeling was carried out with Imaris software (Bitplane, Oxford Instruments). A Leica SPE (phosphorylated Chk2 (Thr68)) confocal system was used to generate confocal stacks of approximately 15 µm with a slice interval of 0.84 µm under identical imaging conditions for all samples in an experiment. The nuclear signal was used in Imaris to provide a mask for nucleus-localized signal from

immunolabeling. Quantification of mean nuclear immunolabeling fluorescence was then carried out with Imaris.

Total RNA extraction and RNA sequencing analysis. RNA-seq libraries were prepared according to the TruSeq Stranded Total RNA Sample Preparation, Low Sample (LS) protocol. The RNA-seq samples were run on a single HiSeq 2500 flow cell. Quality control in FastQ files was performed with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Alignment of cleaned sequencing reads to reference genome hg19 was performed with STAR[32], guided by gene annotation in the refGene UCSC table. The read count per gene was determined by HTseq[33]. Using edgeR[34], trimmed mean of *M* values (TMM) normalization and differential gene expression testing by general linear modeling (GLM) was performed and false discovery rate (FDR)-adjusted *P* values <0.05 were considered significant. Gene Ontology and Pathway analysis was performed for the 115 genes identified using the ToppGene Suite[35].

Statistical analysis. Sample size was not predetermined by statistical analysis. In all experiments involving zebrafish embryos, selection was random for scoring. Exact numbers of analyzed embryos are reported at relevant locations in the main text or the supplementary information. Statistical analysis was carried out with Prism (GraphPad). Distribution of the data sets was tested by D'Agostino–Pearson omnibus normality test. Depending on the outcome of this test, comparison of two conditions with each other was carried out using a Student's two-tailed *t* test or Mann–Whitney two-tailed *U* test throughout the manuscript.

Data availability. The RNA-seq data have been deposited at the European Genome-phenome Archive (EGA) under accession EGAS00001002330.

A **Life Sciences Reporting Summary** is available.

28. Westerfield, M. *The Zebrafish Book: A Guide for the Laboratory Use of Zebrafish (Brachydanio rerio)* (M. Westerfield, 1993).
29. Slaats, G.G. *et al.* Nephronophthisis-associated CEP164 regulates cell cycle progression, apoptosis and epithelial-to-mesenchymal transition. *PLoS Genet.* **10**, e1004594 (2014).
30. Tessadori, F. *et al.* Nodal signaling range is regulated by proprotein convertase–mediated maturation. *Dev. Cell* **32**, 631–639 (2015).
31. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
32. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
33. Anders, S., Pyl, P.T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
34. Robinson, M.D., McCarthy, D.J. & Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
35. Chen, J., Bardes, E.E., Aronow, B.J. & Jegga, A.G. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res.* **37**, W305–W311 (2009).

nature research

Corresponding Author: Gijs van Haaften

Date: Aug 16, 2017

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work we publish. This form is published with all life science papers and is intended to promote consistency and transparency in reporting. All life sciences submissions use this form; while some list items might not apply to an individual manuscript, all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

Sample size was based on previous work in our laboratory.

2. Data exclusions

Describe any data exclusions.

No exclusion criteria were used for analysis.

3. Replication

Describe whether the experimental findings were reliably reproduced.

The experimental findings presented were reliably reproduced.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

All embryos analyzed were chosen randomly

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

For the experiments reported in Figure 2f, Supplementary Fig. 10 and Supplementary Fig. 11, quantification of the immunolabeling was carried out blindly, ie samples were scored and only after completion of the scoring of all samples were the conditions of each sample revealed.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or the Methods section if additional space is needed).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒	A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly.
☐	☒	A statement indicating how many times each experiment was replicated
☐	☒	The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☒	☐	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒	The test results (e.g. *p* values) given as exact values whenever possible and with confidence intervals noted
☐	☒	A summary of the descriptive statistics, including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒	Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Nature Genetics: doi:10.1038/ng.3956

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	All software used to analyze data is reported in the SI of the manuscript

For all studies, we encourage code deposition in a community repository (e.g. GitHub). Authors must make computer code available to editors and reviewers upon request. The *Nature Methods* guidance for providing algorithms and software for publication may be useful for any submission.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	No unique materials were used

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	Rabbit-anti-γH2Ax, Calbiochem, Ref Nr: DR1017; Lot:D00157639 used in zebrafish WB by Slaats et al., PLoS Genet. 2014 Oct 23;10(10):e1004594. doi: 10.1371/journal.pgen.1004594. To the best of our knowledge the antibody we used was not previously used in IF in zebrafish, however the localization of the fluorescent signal is consistent with what is expected from γH2Ax nuclear localization . Moreover, we have tested the antibody on campthotecin-treated zebrafish embryos as positive control and could observe a strong increase in the IF signal. Imaging of this control experiment can be provided upon request. Rabbit-anti-pChk2(Thr68), Cell Signalling Technology #2661; Lot:11 used in zebrafish IF by Zhang et al., Cell Cycle. 2014;13(24):3828-38. doi: 10.4161/15384101.2014.967066. Rabbit-anti-zebrafish p53, Genetex GTX128135; Lot:42263 used in zebrafish IF and WB on the manufacturer's website http://www.genetex.com/p53-antibody-GTX128135.html and references therein. Rabbit-anti-H4K91Ac, Abcam ab4627, Lot: GR309535-1 used and published in Human and S.Cerevisiae as stated on manufacturer's website. We have tested the antibody on Trichostatin A-treated (HDAC inhibitor) zebrafish embryos as positive control and could observe increase in IF and WB signal. Imaging of this control experiment can be provided upon request. Rabbit-anti-H4, EMD Millipore 07-108, Lot: 2709119. used and published in numerous publications a reported on the manufacturer's website. Reactivity with Mouse, Human, Bovine, Avian, Chicken, Xenopus. Broader species cross-reactivity including Zebrafish is expected. Mouse-anti-FLAG, Sigma-Aldrich F3165; used in over 1500 publications as stated on the manufacturer's website and Hu et al., Sci.Rep. 2017; 7:2979 in zebrafish WB.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	Only patient-derived fibroblasts have been used for this study.
b. Describe the method of cell line authentication used.	Only patient-derived fibroblasts have been used for this study; cell lines have been authenticated by detection of the patient specific mutation by Sanger sequencing, RNAseq and Mass Spec
c. Report whether the cell lines were tested for mycoplasma contamination.	The patient-derived fibroblasts were tested negative for mycoplasma contamination.
d. If any of the cell lines used in the paper are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	No commonly misidentified cell lines were used.

Nature Genetics: doi:10.1038/ng.3956

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	Zebrafish strains used in the study are reported in The Experimental Procedures section of the SI. Age of analysis is reported in the relevant figures.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	Detailed patient reports can be found at the beginning of the SI section of the manuscript.

Nature Genetics: doi:10.1038/ng.3956