Genetic compensation triggered by mutant mRNA degradation

Mohamed A. El-Brolosy, Zacharias Kontarakis, Andrea Rossi, Carsten Kuenne, Stefan Günther, Nana Fukuda, Khrievono Kikhi, Giulia L. M. Boezio, Carter M. Takacs, Shih–Lei Lai, Ryuichi Fukuda, Claudia Gerri, Antonio J. Giraldez & Didier Y. R. Stainier

Recent advances in reverse genetics tools have greatly enhanced our ability to analyse gene function in a wide range of organisms. These studies have reinforced previous observations that many engineered mutants do not exhibit an obvious phenotype, reviving interest in the concept of genetic robustness. Several mechanisms have been proposed to explain genetic robustness, including functional redundancy, rewiring of genetic networks, and, in the case of rapidly proliferating organisms such as yeast, the acquisition of adaptive mutations. In a previous report, we proposed genetic compensation—and specifically, transcriptional adaptation—as another underlying mechanism. According to this model, a mutation can lead to the increased expression of related genes that are themselves able to assume the function of the mutated gene. We provided evidence that this increased expression could be induced by a process upstream of the loss of protein function, indicating the existence of an unknown trigger. Here, to investigate the underlying molecular machinery, we developed and investigated several models of transcriptional adaptation in zebrafish and mouse.

Transcriptional adaptation in zebrafish and mouse

We began by analysing different zebrafish and mouse mutants that either have a premature termination codon (PTC) or have their last exon deleted (Extended Data Fig. 1). In zebrafish, hbegfa, vcl, hif1lab, vegfaa, egfl7 and alcama mutants exhibit increased mRNA expression levels of a gene paralogue or family member (hereafter referred to as ‘adapting gene’), namely, hbegfb, vclb, epas1a and epas1b, vegfab, emelin3a and alcamb, respectively (Fig. 1a). Injection of wild-type hif1lab, vegfaa, egfl7 and alcama mRNA into the respective mutants did not have any effect on this transcriptional adaptation response (Extended Data Fig. 2a), suggesting that it is triggered upstream of the loss of protein function. Moreover, we found that vcl, hif1lab and egfl7 heterozygous animals also displayed transcriptional adaptation, albeit less pronounced than that observed in the homozygous mutants (Extended Data Fig. 2b), indicating that transcriptional adaptation is a dominant phenomenon. Notably, we also observed an upregulation of the wild-type allele in hbegfa, hif1lab, vegfaa and alcama heterozygous embryos (Extended Data Fig. 2c). Similarly, we found that Fermt2 (also known as Kindlin-2) mutant mouse kidney fibroblasts (MKFs), Rela and Actg1 mutant mouse embryonic fibroblasts (MEFs) and Actb mutant mouse embryonic stem cells (mESCs) (hereafter referred to as the knockout alleles) displayed increased mRNA levels of Fermt1 (also known as Kindlin-1), Rel, Actg2 and Actg1, respectively (Fig. 1b). Transfection of wild-type Fermt2 and Rela in the respective knockout cells did not dampen the transcriptional adaptation response (Extended Data Fig. 2d–f), and we also found that Actb heterozygous mESCs upregulated Actg1 (Extended Data Fig. 2g). Altogether, these data strongly indicate that the loss of protein function is not the trigger for the transcriptional adaptation response observed in these models.

To determine whether the increased mRNA expression levels were caused by increased transcription of the adapting gene or increased mRNA stability, we measured the precursor mRNA (pre-mRNA) levels of hbegfb and emelin3a in hbegfa and egfl7 zebrafish mutants, and found that they were also upregulated (Extended Data Fig. 3a). Similar findings were obtained for Fermt1 and Rel pre-mRNA expression levels in Fermt2 and Rela knockout mouse cells (Extended Data Fig. 3b). Together, these data indicate an increase in transcription of the adapting genes. In addition, Fermt2 knockout MKFs displayed increased chromatin accessibility at the Fermt1 transcription start site (TSS), as observed by the assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq) (Extended Data Fig. 3c).

RNA decay triggers transcriptional adaptation

As the loss of protein function does not appear to be the trigger for the transcriptional adaptation response, we investigated two other possibilities: the DNA lesion and the mutant mRNA. In the context of the DNA lesion, we found that mutant mRNA degradation does not rise to more severe phenotypes than alleles displaying mutant mRNA decay. Transcriptome analysis in alleles displaying mutant mRNA decay reveals the upregulation of a substantial proportion of the genes that exhibit sequence similarity with the mutated gene’s mRNA, suggesting a sequence-dependent mechanism. These findings have implications for our understanding of disease-causing mutations, and will help in the design of mutant alleles with minimal transcriptional adaptation-derived compensation.

1Department of Developmental Genetics, Max Planck Institute for Heart and Lung Research, Bad Nauheim, Germany. 2ECCPS Bioinformatics and Deep Sequencing Platform, Max Planck Institute for Heart and Lung Research, Bad Nauheim, Germany. 3Department of Genetics, Yale University School of Medicine, New Haven, CT, USA. 4Present address: Leibniz Research Institute for Environmental Medicine, Düsseldorf, Germany. 5Present address: University of New Haven, New Haven, CT, USA. 6Present address: Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan. 7Present address: The Francis Crick Institute, London, UK. 8These authors contributed equally: Zacharias Kontarakis, Andrea Rossi. *e-mail: didier.stainier@mpi-hh.mpg.de
that pre-mRNA expression levels remained unchanged or were slightly upregulated compared to wild type, in contrast with the mRNA levels (Extended Data Fig. 4a). Similar findings were observed in Fermt2 and Rela knockout mouse cells (Extended Data Fig. 4b). Moreover, metabolic labelling of newly synthesized transcripts revealed similar or increased levels of Fermt2, Rela and Actg1 mutant transcripts compared to wild type (Extended Data Fig. 4c), while transcription inhibition assays showed that the mutant transcripts had shorter half-lives (Extended Data Fig. 4d–f). Together, these data indicate that mutants that exhibit transcriptional adaptation have reduced mutant mRNA levels because of mRNA decay.

To investigate the role of the mRNA surveillance machinery in transcriptional adaptation, we genetically inactivated the key nonsense-mediated decay (NMD) factor Upf1 in hbegfaΔ7, vegfaa and vclaΔ13 zebrafish mutants. Inactivating upf1 in these mutants led to a reduction in mutant mRNA decay (Extended Data Fig. 5a), and loss of transcriptional adaptation (Fig. 2a). We also observed a decrease in, or loss of, transcriptional adaptation in Rela and Actb knockout mouse cells when we knocked down proteins involved in the mRNA surveillance machinery (Fig. 2b, c, Extended Data Fig. 5b, c). Pharmacological inhibition of NMD in hbegfaΔ7 zebrafish mutants (Extended Data Fig. 5d, e), and blocking translation in Rela knockout mouse cells (an alternative approach to inhibit mRNA decay) (Extended Data Fig. 5f, g) led to similar observations. We next investigated whether inducing mRNA degradation in wild-type zebrafish or in mouse cells by using uncapped RNAs, which are known to be rapidly degraded by 5'- to 3'- exonucleases, could trigger transcriptional adaptation. Indeed, injection of uncapped hif1ab or vegfaa RNAs into wild-type embryos induced transcriptional adaptation (Fig. 2d), including an increase in endogenous hif1ab and vegfaa expression (Extended Data Fig. 5h). Similarly, transfection of uncapped Actb RNA into wild-type mESCs led to Actg1 upregulation (Extended Data Fig. 5i). Moreover, injecting wild-type embryos with uncapped hif1ab or vegfaa RNAs containing an upstream sequence that renders them resistant to 5'- to 3' exonuclease-mediated decay did not induce transcriptional adaptation (Extended Data Fig. 5j). Altogether, these data indicate that mRNA degradation is a major factor in triggering transcriptional adaptation. Notably, injection of uncapped RNA that was synthesized from the non-coding strand of hif1ab or vegfaa did not lead to an increase in epas1a or vegfab mRNA expression levels (Extended Data Fig. 5k), suggesting that the RNA sequence itself may have a role in transcriptional adaptation.

If mutant mRNA degradation is required for transcriptional adaptation, alleles that fail to transcribe the mutated gene should not display this response. To this end, we used CRISPR–Cas9 to generate such alleles (hereafter referred to as RNA-less alleles) by deleting promoter regions or the entire gene locus. RNA-less alleles of hbegfa, vegfaa and alcama failed to upregulate hbegfa, vegfaa and alcama, respectively (Fig. 3a). Similarly, in mouse cells, RNA-less alleles of Rela, Actg1 and Actb did not upregulate the adapting gene (Fig. 3b). We also attempted to generate a promoter-less allele of Fermt2 in MEFs; however, the obtained clones exhibited proliferation defects that prevented their expansion. As an alternative, we used CRISPR interference (CRISPRi) and found that reducing transcription of the mutant Fermt2 gene in Fermt2 knockout cells led to a decrease in Fermt1 mRNA expression levels (Extended Data Fig. 6a). Notably, we observed that the promoter-less Rela MEFs were more sensitive to TNF-induced apoptosis9 than the Rela knockout MEFs (Fig. 3c). Similarly, mESCs with a full locus deletion of Actb displayed less provocative activity than Actb knockout cells (Fig. 3d, e). We also generated an egfl7 RNA-less allele in zebrafish and observed that these mutants displayed pronounced vascular defects (Fig. 3f, g), and milder upregulation of the emilin genes (Extended Data Fig. 6b) compared to the phenotypically wild-type egfl7Δ4 allele4. Furthermore, vegfaa promoter-less mutants displayed a stronger central artery sprouting phenotype than vegfaaΔ10 mutants (Extended Data Fig. 6c); hbegfa RNA-less mutants displayed slow blood circulation—a phenotype not observed in hbegfaΔ7 mutants (Extended Data Fig. 6d).
data Fig. 6d); and alcama promoter-less mutants, but not alcama mutant, exhibited an elongated cardiac ventricle (Extended Data Fig. 6e). Therefore, use of RNA-less alleles can uncover phenotypes that are not observed in alleles exhibiting mutant mRNA degradation.

Sequence similarity and transcriptional adaptation

Next, we performed transcriptome analysis of Fermt2, Actg1 and Actb knockout cells and observed that hundreds of genes were upregulated in the knockout cells compared to wild-type cells (Extended Data Fig. 7a). Only 81 genes were upregulated in all three of the knockout models, with no signs of a stress-induced response (Extended Data Fig. 7b, c, Supplementary Table 1). A first-pass analysis of the upregulated genes in each model showed that a disproportionate number of them exhibited sequence similarity with the mutated gene. We thus explored the relation between upregulation and sequence similarity based on several similarity thresholds (Extended Data Fig. 7d; see ‘Sequence similarity and subsampling analyses’ section of the Methods) and found a significant correlation. Using the optimal ‘expect’ (E) values to identify similar genes, we observed that at least 50% of them were significantly upregulated in the different knockout cell models, compared to a maximum of 21% of non-similar genes (Fig. 4a, Extended Data Fig. 8a–c). Notably, seven of the twelve upregulated similar genes in Actg1 knockout cells were not upregulated in Actg1 RNA-less cells, and four of the six upregulated similar genes in Actb knockout cells were not upregulated in Actb RNA-less cells (Extended Data Fig. 8b, c). We also observed that four of the nine similar genes that were not upregulated in Actg1 knockout cells at the mRNA level were upregulated at the pre-mRNA level (Extended Data Fig. 8d). Additional studies showed that injection of uncapped mouse Actb RNA into zebrafish embryos led to an increase

in zebrafish actb1 mRNA expression levels (Extended Data Fig. 8e, Supplementary Data), in line with the sequence similarity analyses mentioned above.

To begin to investigate how sequence similarity could have a role in transcriptional adaptation, we first injected uncapped transcripts containing only similar or non-similar sequences into zebrafish embryos. Using the hif1ab model, we observed that only transcripts containing sequences similar to epas1a led to an increase in epas1a mRNA expression levels (Fig. 4b, Extended Data Fig. 8f, Supplementary Data).

Fig. 2 | Mutant mRNA decay is required for transcriptional adaptation. a, qPCR analysis of hbeqfb, vegfa and vcb mRNA expression levels in upf1;hbeqfb, upf1;vegfa and upf1;vcb double mutant zebrafish. b, qPCR analysis of Rel mRNA expression levels after siRNA-mediated knockdown of indicated proteins in Rela knockout mice. Scr, scrambled siRNA control. c, qPCR analysis of Actg1 mRNA expression levels after siRNA-mediated knockdown of indicated proteins in Actb knockout mice. d, qPCR analysis of injected hif1ab, epas1a, injected vegfaa and vegfab mRNA expression levels in 6 hpf wild-type zebrafish embryos injected with the indicated RNA. cap, capped; uncap, uncapped. Wild-type or control expression levels were set at 1. n = 3 biologically independent samples. Data are mean ± s.d., and a two-tailed Student’s t-test was used to calculate P values.

Fig. 3 | Alleles that fail to transcribe the mutated gene do not display transcriptional adaptation. a, qPCR analysis of hbeqfb, hbeqfb, vegfaa, vegfab, alcama and alcamb mRNA expression levels in zebrafish that lack the full hbeqfb locus (full locus deletion; full locus del.) or the vegfaa or alcama promoter (promoter-less), compared to wild type. b, qPCR analysis of Rela, Rela, Actg1, Actg2, Actb and Actg1 mRNA expression levels in MEFs and mESCs that lack the Rela promoter or the full Actg1 or Actb locus, compared to wild-type cells. c, Cytotoxicity assay following TNF treatment of wild-type, Rela knockout and Rela promoter-less MEFs. Percentages are normalized relative to dimethyl sulfoxide (DMSO)-treated cells. d, Confocal microscopy of wild-type, Actb knockout and Actb full locus deletion mESCs. Actin filaments are depicted in white and nuclei in red (4',6'-diamidino-2-phenylindole (DAPI)). e, Actin filament protrusion length in wild-type, Actb knockout and Actb full locus deletion mESCs. f, Confocal micrographs of 48 hpf Tg(fli1a:egfp) wild-type and egfl7 full locus deletion mutant zebrafish. Lateral views; anterior is to the left. Higher magnifications of the dashed boxes are shown in g. Scale bars, 20 μm (d), 500 μm (f). Wild-type expression levels were set at 1 (a, b), n = 3 (a, b); n = 5 (c); and n = 189 (wild type), 219 (Actb knockout) and 205 (Actb full locus deletion) (e) independent samples. Data are mean ± s.d., and a two-tailed Student’s t-test was used to calculate P values. The experiments in d and f were repeated twice independently with similar results.
Fig. 4 | Transcriptional adaptation favours genes that exhibit sequence similarity with the mutated gene’s mRNA, and is associated with permissive histone marks. a, Percentage of significantly upregulated (that is, where \( \log_{2}(\text{gene expression level in knockout cells/gene expression level in wild-type cells}) \) is greater than 0 and \( P < 0.05 \)) protein-coding genes that exhibit sequence similarity (red) or do not exhibit sequence similarity (blue) with Fermt2, Actg1 or Actb. b, qPCR analysis of eps1a mRNA expression levels in 6 haplotype-type zebrafish that were injected with uncleaved RNA composed solely of either the hif1ab mRNA sequences (that are similar (sim.) to eps1a) or the hif1ab mRNA sequences that are not similar (non-sim.) to eps1a. Control expression levels were set at 1. c, d, Chi–qPCR analysis of WDR5 (c) and H3K4me3 (d) occupancy near the TSSs of Fermt1, Rel and Actg2 in Fermt1, Rel and Actg1 knockout mouse cells, respectively, compared to wild-type cells. Quantification of enrichment is shown as fold enrichment over IgG control. e, Current putative simplified model of transcriptional adaptation to mutations. DFs, decay factors; RBPs, RNA-binding proteins; TC, termination codon. n = 3 (b) or \( n = 2 \) (c, d) biologically independent samples. Data are mean ± s.d., and a two-tailed Student’s \( t \)-test was used to calculate \( P \) values.

Data. We also generated synthetic transcripts that contained hif1ab sequences similar to the promoter, exons, introns or 3′ untranslated region (UTR) of eps1a. Injection of uncapped versions of these transcripts revealed that those exhibiting sequence similarity with exons or introns induced transcriptional adaption, whereas those exhibiting sequence similarity with the 3′ UTR did not; and the transcripts exhibiting sequence similarity with the promoter induced only a mild response (Extended Data Fig. 8g). These data are consistent with the transcriptome analyses of Fermt2, Actg1 and Actb knockout cells (Extended Data Fig. 8a–c)—genes that exhibited sequence similarity with the mutated gene’s mRNA in their 3′UTR were not upregulated, and genes that exhibited sequence similarity with promoters displayed a mild upregulation or were not upregulated. Altogether, these data suggest that, at least in some cases, sequence similarity has a role in transcriptional adaptation.

Epigenetic remodelling in adapting genes

In the past decade, it has become evident that the control of mRNA stability has an important role in gene expression10–13. Several studies have reported that mRNA decay factors can translocate to the nucleus and interact with histone modifiers and chromatin remodelers to modulate gene expression14–17. We thus performed a targeted small interfering RNA (siRNA) screen in Rela knockout cells to identify epigenetic modulators that are involved in transcriptional adaptation (Supplementary Table 2). Knockdown of the histone lysine demethylases KDM4 or KDM6, which remove the inhibitory H3K9me3 or H3K27me3 histone marks, respectively, dampened the transcriptional adaptation response; however, the strongest effect was observed after WDR5 knockdown (Extended Data Fig. 9a). WDR5 is part of the COMPASS complex, which generates the permissive histone mark H3K4me3. Chromatin immunoprecipitation (ChIP) revealed enrichment of WDR5 and H3K4me3 at the TSS of Fermt1, Rel and Actg2 in Fermt2, Rela and Actg1 knockout cells (Fig. 4c, d and Extended Data Fig. 9b). Moreover, knockdown of UFP1/EXOSC4 or XRN1 in Rela knockout cells led to the depletion of H3K4me3 at the Rel TSS (Extended Data Fig. 9c). Altogether, these data suggest a model in which concomitant with mutant mRNA degradation, decay factors translocate to the nucleus, where they bind to specific loci (possibly guided by decay intermediates) and recruit histone modifiers and/or chromatin remodelers to upregulate transcription (Fig. 4e and Extended Data Fig. 9d).

While investigating this model, we noted a study reporting that, in mESCs, transfection of short fragments of Cdk9 or Sox9 mRNA could lead to an increase in expression of these genes18. Mechanistically, these RNA fragments were found to downregulate native antisense transcripts that normally function as negative regulators of Cdk9 and Sox9 expression. Notably, we found that transfection of uncleaved Cdk9 or Sox9 RNA also led to upregulation of these genes (Extended Data Fig. 10a). Furthermore, in another study19, knockdown of a BDNF antisense transcript in HEK293T cells was reported to cause the upregulation of the sense transcript—a response that involved a decrease in the inhibitory H3K27me3 histone mark. We transfected HEK293T cells with uncleaved BDNF sense RNA and observed a downregulation of the antisense transcript and a concomitant upregulation of the sense one (Extended Data Fig. 10b). Notably, we also observed a downregulation of antisense transcripts at the hbegfb and vclb loci in hbegfb23 and vela21,25 mutants, respectively (Extended Data Fig. 10c, d). These data indicate that acting on antisense transcripts is another possible mechanism through which mRNA decay intermediates could induce transcriptional adaptation in a sequence-specific manner (Fig. 4e and Extended Data Fig. 9d).

Discussion

Despite its potential importance5, transcriptional adaptation to mutations and its underlying molecular mechanisms remain poorly understood. Here we show that the mRNA surveillance machinery is important not only to prevent the translation of defective transcripts, but also to buffer against mutations by triggering the transcriptional upregulation of related genes, including the mutated gene itself (see Supplementary Discussion).

For several human genetic diseases, missense mutations or in-frame insertions or deletions (indels)—which are less likely to lead to mutant mRNA degradation—are reportedly more common than potentially mRNA-desestabilizing nonsense mutations or out-of-frame indels20–25. Of note, a study on patients with Marfan syndrome reported that when compared to individuals with FBN1 missense mutations, the mildest form of the disease was observed in an individual displaying very low mutant FBN1 transcript levels owing to an out-of-frame indel that led to a PTC in the FBN1 coding sequence26. Similar observations have been described for mutations in the HBB gene27. The current dogma is that pathogenic missense mutations tend to be more common in affected individuals because they might lead to constitutively active or dominant negative proteins. However, we propose that nonsense mutations are less common as they might result in mRNA decay-triggered upregulation of related genes and therefore not cause noticeable symptoms. Detailed transcriptomic analyses of relevant individuals will help to test this hypothesis. Moreover, studies28,29 of healthy individuals have reported homozygous loss-of-function mutations in several genes (including EGFL7 and RELA, the zebrafish and mouse homologues of which we studied here), and it will be interesting to investigate whether degradation of the mutant transcripts is associated with a transcriptional adaptation response that protects them. Such analyses may help us to understand why some mutations cause disease whereas others do not. They may also help in the identification of modifier genes, the expression levels of which could be further modulated for therapeutic purposes.
Article

1. Tautz, D. Redundancies, development and the flow of information. BioEssays 14, 263–266 (1992).
2. Barabási, A. L. & Oltvai, Z. N. Network biology: understanding the cell’s functional organization. Nat. Rev. Genet. 5, 101–113 (2004).
3. Teng, X. et al. Genome-wide consequences of deleting any single gene. Mol. Cell 52, 485–494 (2013).
4. Rossi, A. et al. Genetic compensation induced by deleterious mutations but not gene knockdowns. Nature 524, 230–233 (2015).
5. El-Broly, M. A. & Stainer, D. V. R. Genetic compensation: a phenomenon in search of mechanisms. PLoS Genet. 13, e1006780 (2017).
6. Isken, O. & Maquat, L. E. Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. Nat. Rev. Genet. 5, 101–113 (2004).
7. Mukherjee, C. et al. Identification of cytoplasmic capping targets reveals a role for cap homeostasis in translation and mRNA stability. Cell Reports 2, 674–684 (2012).
8. Boehm, V., Gerbracht, J. V., Marx, M. C. & Gehrng, N. H. Investigating the degradation pathways of unstable mRNAs with XRN1-resistant sequences. Nat. Commun. 7, 13691 (2016).
9. Doh, T. S. et al. Absence of tumor necrosis factor rescues RelA-deficient mice from embryonic lethality. Proc. Natl Acad. Sci. USA 96, 2994–2999 (1999).
10. Hao, S. & Baltimore, D. The stability of mRNA influences the temporal order of the induction of genes encoding inflammatory molecules. Nat. Immunol. 10, 281–288 (2009).
11. Elkon, R., Zlotorynski, E., Zeller, K. I. & Agami, R. Major role for mRNA stability in shaping the kinetics of gene induction. BMC Genomics 11, 259 (2010).
12. Rabani, M. et al. Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. Nat. Biotechnol. 29, 436–442 (2011).
13. Sun, M. et al. Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. Genome Res. 22, 1350–1359 (2012).
14. Collins, S. R. et al. Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature 446, 806–810 (2007).
15. Berretta, J., Pinskaya, M. & Morillon, A. A cryptic unstable transcript mediates transcriptional trans-silencing of the Ty1 retrotransposon in S. cerevisiae. Genes Dev. 22, 615–626 (2008).
16. Pinskaya, M., Gourvenec, S. & Morillon, A. H3 lysine 4 di- and tri-methylation deposited by cryptic transcription attenuates promoter activation. EMBO J. 28, 1697–1707 (2009).
17. Haimovich, G. et al. Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. Cell 153, 1000–1011 (2013).
18. Ghanbarian, H. et al. Small RNA-directed epigenetic programming of embryonic stem cell cardiac differentiation. Sci. Rep. 7, 41799 (2017).
19. Modarresi, F. et al. Inhibition of natural antisense transcripts in vivo results in gene-specific transcriptional upregulation. Nat. Biotechnol. 30, 453–459 (2012).
20. Esensmith, R. C. & Woo, S. L. Molecular basis of phenylketonuria and related hyperphenylalaninemias: mutations and polymorphisms in the human phenylalanine hydroxylase gene. Hum. Mutat. 1, 13–23 (1992).
21. Myerowitz, R. Tay-Sachs disease-causing mutations and neutral polymorphisms in the Hex A gene. Hum. Mutat. 9, 195–208 (1997).
22. Genchel, J. & Schmidt, H. H. Mutations in the LMNA gene encoding lamin A/C. Hum. Mutat. 16, 451–459 (2000).
23. Chuzhanova, N. A., Anassis, E. J., Ball, E. V., Krawczak, M. & Cooper, D. N. Meta-analysis of indels causing human genetic disease: mechanisms of mutagenesis and the role of local DNA sequence complexity. Hum. Mutat. 21, 28–44 (2003).
24. Ferec, C. & Cutting, G. R. Assessing the disease-liability of mutations in CFTR. Cold Spring Harb. Perspect. Med. 2, a009480 (2012).
25. Zhou, Q. et al. Early-onset stroke and vasculopathy associated with mutations in ADAR2. N. Engl. J. Med. 370, 291–302 (2014).
26. Dietz, H. C. et al. Four novel FBN1 mutations: significance for mutant transcript level and EGF-like domain calcium binding in the pathogenesis of Marfan syndrome. Genomics 17, 468–475 (1993).
27. Hall, G. W. & Thein, S. Nonsense codon mutations in the terminal exon of the i-protocin gene are not associated with a reduction in i-mRNA accumulation: a mechanism for the phenotype of dominant i-thalassemia. Blood 83, 2031–2037 (1994).
28. Sulem, P. et al. Identification of a large set of rare complete human knockouts. Nat. Genet. 47, 448–452 (2015).
29. Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 536, 285–291 (2016).

Acknowledgements We thank V. Serebryan, F. Mueller, Z. Jiang, A. Beisaw and F. Gunawan for discussion and comments on the manuscript; J. Pestel for the alcaum mutant; A. Atzberger for support with cell sorting; and N. Gehring and V. Böhm for providing the XRN1-resistant sequence plasmid. M.A.E.-B. was supported by a Boehringer Ingelheim Fonds PhD fellowship. Research in the D.Y.R.S. laboratory is supported by the Max Planck Society, the EU, the DFG and the Leducq Foundation.
METHODS

Statistics and reproducibility. No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment. All experiments were performed at least twice unless otherwise noted. P < 0.05 was accepted as statistically significant.

Zebrafish husbandry. All zebrafish (Danio rerio, strain: TüB/AB) husbandry was performed under standard conditions in accordance with institutional (Max Planck Gesellschaft) and national ethical and animal welfare guidelines approved by the ethics committee for animal experiments at the Regierungsspräsidium Darmstadt, Germany (permit number B2/1017). All experiments were performed on zebrafish embryos or larvae between 6 hpf and 6 days post-fertilization (dpf). We used the following previously published mutant and transgenic lines: hfl (ref. 20), vegfas (ref. 38), vegfaas (ref. 39), epas1 (ref. 4), epas2 (ref. 4), vega (ref. 5), hfl (ref. 6), hfl (ref. 7), hfl (ref. 8), and hfl (ref. 9).

Rela expression levels of the mutated gene in the cell line qPCR experiments. rpl13 and gapdh, and Actb, B2d, d15a18s and Gapdh, were used to normalize zebrafish and mouse experiments, respectively. rpl13 was chosen as a reference gene for zebrafish experiments as its expression levels were not changed between wild-type and qfl27, hfl, vega, vegf, and alcama mutant embryos (unpublished microarray data). Primer sequences used for the qPCR experiments are listed in Supplementary Table 4. Fold changes were calculated using the 2−ΔΔCt method. All Ct and ΔCt values are listed in the Source Data files. C values for the reference genes ranged between 12 and 23 except for B2d, which ranged between 3 and 6.

In vitro transcription and RNA microinjections. cDNAs encoding hfl, vegfa, vegf, and alcama full-length mRNAs were amplified using whole-embryo cDNA as a template. PCR fragments were ligated into a pCS2+ vector between BamH1 and Xba1 (NEB) sites. All constructs were verified by sequencing. Plasmids were linearized using NotI (NEB) and in vitro transcribed using the mMESSAGE mMA- CHINE SP6 kit (Life Technologies). RNA was then purified using an RNA Clean and Concentrator kit (Zymo Research). Approximately 10–100 pg of each mRNA was injected into embryos from heterozygous incrosses at the one-cell stage. At 22–30 hpf, embryos were collected in TRIZol for qPCR analysis.

To transcribe uncap RNAs, cDNAs was used to amplify zebrafish hfl and vegf and mouse Actb, whereas genomic DNA was used to amplify mouse Cdk9 and Sox9 (single exon for Cdk9 and the entire genomic locus (exons and introns) for Sox9) as their expression levels were too low to amplify from cDNA. In all cases, a reverse primer containing the adaptor sequence 5′GGCAAGCTTTATATGTCATATGAG3′ was used for amplification. A reverse primer containing only the adaptor sequence was subsequently used for qPCR detection of the injected transcripts, as described above. To generate uncapped XRNA transcripts, the XRNA-resistant sequence 3′Δ was cloned upstream of the hfl and vegf a coding sequences in pCS2+, and then linearized using NotI for in vitro transcription. In vitro synthesis of uncap transcripts was performed using SP6 (or T7 in the case of the uncap RNAs corresponding to the non-coding strand) RNA polymerase (Promega) and the RNA was purified using an RNA Clean and Concentrator kit. Alternatively, a T7 sequence was added to the forward primer of the PCR reaction and the product was directly used for in vitro synthesis of uncap RNAs. The RNAs were injected into one-cell-stage zebrafish embryos (50 pg) or transfected into cells (1 μg) in 12-well plates.

In vitro generation of synthetic transcripts containing different sequences of hfl mRNA. Oligonucleotides containing hfl sequences similar to eps1a were ligated together along with a 5′T7 promoter sequence. Similar sequences were identified using a highly sensitive BLASTn analysis with a word size of 7 and an E value of up to 1,000,000. The alignment was visualized using Kalign at the less sensitive E value of 25 owing to display constraints (Extended Data Fig. 8f). The same approach was used to generate transcripts of hfl sequences that were not similar to eps1a, using the non-alignable sequences. In vitro transcription was performed using the T7 RNA polymerase (Promega).

Gene knockdown. Guided by RNA synthesis (Table 3) and expression analysis, this process was performed on embryos from at least three different crosses.

High-resolution melt analysis was used to genotype all mutant zebrafish and mouse cell lines, with the exception of the RNA-less alleles, which were genotyped by PCR. The primer sequences used for genotyping are listed in Supplementary Table 3.

qPCR expression analysis. qPCR was performed in a CFX Connect Real-Time System (Biorad). RNA was isolated using TRizol and at least 500 ng RNA was used for reverse transcription using the Maxima First Strand cDNA synthesis kit (Thermo). All reactions were performed in at least technical duplicates and the results represent biological triplicates. qPCR was performed at the embryonic or larval stage when the wild-type version of the mutated gene exhibits its highest level of expression. For hfl mutants treated with the NMD inhibitor, data were analysed at 6 dpf and not at 72 hpf, as the drug seemed to be effective only when used for three days between 72 hpf and 6 dpf; treatment of earlier-stage embryos was not possible owing to toxicity. Primers were designed using Primer 3 (http://bioinformatics.umassmed.edu/bioapps/primer3_web.cgi). Primers to detect pre-mRNA were designed to bind around intron–exon boundaries. Allele-specific primers were designed in such a way as to amplify the wild-type but not the mutant allele. To determine the injected capped and uncapped RNA levels, a universal reverse primer corresponding to the adaptor sequence at the 3′ end of the injected RNAs was used, along with forward primers designed in close proximity; in this way we were able to distinguish between endogenous and injected RNAs. Several primer pairs along the cDNA of the gene were used to assess transcription in the candidate promoter-less alleles. Only mutant alleles exhibiting less than 10% of wild-type mRNA levels with all the above mentioned primer pairs were used in the study. For the upf2 double mutant data, the figures show expression levels of the adapting gene in the double mutants relative to its expression levels in the upf2 single mutants. Expression levels of the mutated gene in the upf2 double mutants are relative to those in hfl, vegf, vega, or vela single mutants. Equal numbers of cells were used for the cell line qPCR experiments. rpl13 and gapdh, and Actb, B2d, d15a18s and Gapdh, were used to normalize zebrafish and mouse experiments, respectively. rpl13 was chosen as a reference gene for zebrafish experiments as its expression levels were not changed between wild-type and qfl27, hfl, vega, vegf, and alcama mutant embryos (unpublished microarray data). Primer sequences used for the qPCR experiments are listed in Supplementary Table 4. Fold changes were calculated using the 2−ΔΔCt method. All Ct and ΔCt values are listed in the Source Data files. C values for the reference genes ranged between 12 and 23 except for B2d, which ranged between 3 and 6.
5 μg nucleic plasmid DNA, according to the manufacturer's protocol. Two days after transfection, cells expressing eGFP were subjected to single-cell sorting into 96-well plates using a FACSARia III (BD Biosciences). Three to four weeks after single-cell sorting, clones were isolated and genotyped by PCR and sequencing. For generation of RNA-less alleles, cells were co-transfected with two nucleic plasmids.

Actg1ΔNSD (Actg1 knockout) and Actg1ΔNSD locus del mutant MEFs do not exhibit any gross morphological defects.

Overexpression plasmids. Fermi2 and Relax CDNAs were amplified using MEF cDNA as a template. PCR fragments were ligated into the pCDNA3.1 mammalian expression vector (Thermo) between the BamHI and XbaI sites. Plasmids were transfected into the respective knockout cells using FuGENE 6 (Promega) according to the manufacturer's protocol. Two days after transfection, cells were selected with 0.5 μg ml⁻¹ and 2 mg ml⁻¹ G418 (Sigma) for Fermi2 and Relax knockout cells, respectively. A week later, cells were lysed in modified radioimmunoprecipitation assay (RIPA) buffer for western blot analysis.

Uncapped RNA transfections. Uncapped eGFP, Acth, Cdk9 and S6x9 RNAs were transfected into mESCs or MEFS using Lipofectamine Messenger Max (Thermo) transfection reagent according to the manufacturer's protocol. Six hours after transfection, cells were trypsinized and collected in TRIZol for RNA extraction.

CrisPR interference. Three gRNAs targeting the Fermi2 promoter and TSS were each cloned into a plasmid encoding a fusion protein of catalytically dead Cas9 (in which cases cells were collected 24 h after transfection). In most cases, siRNAs were used at 10 nM and given a knockdown efficiency of 70–90% (as assessed by mRNA levels). However, to knock down XRN1 in mESCs, a lower concentration of siRNAs was used (2.5 nM, which led to a knockdown efficiency of 20%) as stronger knockdown affected housekeeping genes as well (data not shown). A scrambled siRNA (Sigma, SIC00), which does not bind to any of the mouse transcripts, was used as a negative control.

For each experiment, the figures show expression levels of the adapting gene in knockout cells relative to its expression levels in wild-type cells that were treated with the same siRNA. Expression levels of the mutant gene in the knockout cells treated with a given siRNA are relative to those in knockout cells treated with the same siRNA. Expression levels of the mutant gene in the knockout cells, or cells that were treated with the same drug. Expression levels of the mutated gene in mutant larvae or knockout cells, relative to its expression levels in wild-type larvae or cells that were treated with the same drug. Expression levels of the mutated gene in the mutant larvae, or knockout cells treated with a given drug, are relative to those in untreated mutant larvae, or knockout cells treated with DMSO. All C and ΔC values are shown in the Source Data files.

Pharmacological treatments. To inhibit NMD, 72 hpf wild-type and knockdown cells were treated with 10 μM NR1I2 (kindly gifted by Dr. Ralph W. Gray) or DMSO, and three days later, they were collected in TRIZol for qPCR analysis. No gross alterations were observed in the drug-treated larvae. To inhibit RNA degradation through translation blockade, wild-type and Relax knockdown cells were treated with 200 μg ml⁻¹-1 cycloheximide (Sigma) or DMSO for 5 h, then collected in TRIZol for RNA extraction.

For each experiment, the figures show expression levels of the adapting gene in mutant larvae or knockout cells, relative to its expression levels in wild-type larvae or cells that were treated with the same drug. Western blot analysis and antibodies. MKFs and MEFS were lysed in modified RIPA buffer (150 μM NaCl, 50 mM Tris-HCl pH 7.4, 1 mM IGEPA, 0.1 mM sodium deoxycholate, 1 mM EDTA) supplemented with protease inhibitors (complete ULTRA Mini, Roche) and phenylmethylsulfonyl fluoride (PMSF). Protein samples of 35 μg were separated using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) on precast TGX gradient gels (Biorad) (ACTB control, 9751, Cell Signaling Technology). The following antibodies were used: RELA (Cell Signaling Technology, 6956, 1:1,000), Cdk9 and XRN1 (in which cases cells were collected 24 h after transfection). In most cases, siRNAs were used (2.5 nM, which led to a knockdown efficiency of 20%) as stronger knockdown affected housekeeping genes as well (data not shown). A scrambled siRNA (Sigma, SIC00), which does not bind to any of the mouse transcripts, was used as a negative control.

For each experiment, the figures show expression levels of the adapting gene in knockout cells relative to its expression levels in wild-type cells that were treated with the same siRNA. Expression levels of the mutant gene in the knockout cells treated with a given siRNA are relative to those in knockout cells treated with the same siRNA. Expression levels of the mutant gene in the knockout cells, or cells that were treated with the same drug. Expression levels of the mutated gene in mutant larvae or knockout cells, relative to its expression levels in wild-type larvae or cells that were treated with the same drug. Expression levels of the mutated gene in the mutant larvae, or knockout cells treated with a given drug, are relative to those in untreated mutant larvae, or knockout cells treated with DMSO. All C and ΔC values are shown in the Source Data files.

RNA metabolic labelling. Metabolic labelling was performed as previously described 19,45. In brief, cells were treated with 200 μM 4-thiouridine (4sU; Sigma) for 1 h, followed by phenol–chloroform RNA extraction. RNA (80 μg) was incubated with biotin-HDPP (Thermo) to specifically biotinylate the newly transcribed proteins. Biotinylated RNAs were then pulled down using the Bioruptor (Diagenode) to generate fragments of 200–400 bp in size. Immunoprecipitation was then performed as previously described 48. The following antibodies were used: RELA (Cell Signaling Technology, 857, 1:1,000), anti-mouse horseradish peroxidase (HRP)-conjugated IgG (Thermo, 31,430, 1:10,000) and anti-rabbit HRP-conjugated IgG (Thermo, 31,460, 1:10,000). This experiment was performed once.

Chromatin immunoprecipitation. ChIP was performed using the truChIP Chromatin Shearing Reagent kit ( Covaris) using 30 million cells per immunoprecipitation according to the manufacturer's protocol. Chromatin was sheared using BioRuptor (Diagenode) to generate fragments of 200–400 bp in size. Immunoprecipitation was then performed as previously described 30. The following antibodies were used: rabbit IgG (4 μg per immunoprecipitation, 026102, Thermo Fischer Scientific) (P; IgG, 1:1,000), rabbit IgG (4 μg per immunoprecipitation, 31305, Cell Signaling Technology) (H3K14ac) and H3K27me3 (4 μg per immunoprecipitation, 9751, Cell Signaling Technology). Following immunoprecipitation and reverse cross-linking, samples were purified using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel), according to
to the manufacturer’s protocol for samples containing sodium dodecyl sulfate. ChiP–qPCR experiments on Rels knockout cells where UPE1/EXOSC8 or XRNI were knocked down were performed only once.

ATAC-seq material extraction and library preparation. Cells were trypsinized and washed with PBS. Counting of cells was performed by MOXI Z Mini Automated Cell Counter Kit (Ortho) and 50,000 cells were used for ATAC library preparation using the Tn5 transposase from the Nextera DNA Sample Preparation Kit (Illumina). The cell pellet was resuspended in 50 μl FMT and mixed with 25 μl fragmentation DNA (TD) buffer, 2.5 μl Tn5, 0.5 μl 10% NP-40 and 22 μl water. This mixture of Tn5 and cells was incubated at 37 °C for 30 min with occasional snap mixing. Transposase treatment was followed by 30 min incubation at 50 °C with 500 mM EDTA pH8.0 for optimal recovery of digested DNA fragments. For neutralization of EDTA, 100 μl of 50 mM MgCl2 was added, followed by purification of the DNA fragments using the MinElute PCR Purification Kit (Qiagen). Amplification of the library together with indexing was performed as previously described. Libraries were mixed in equimolar ratios and sequenced on a NextSeq500 platform using v2 chemistry.

ATAC-seq analysis. The samples were assessed for quality using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Trimmomatic v.0.33 was used to trim reads after a quality drop below a mean of Q20 in a window of five nucleotides. Only reads above 30 nucleotides were cleared for further analyses. To normalize all samples to the same sequencing depth, 27 million reads per sample were randomly selected for further analysis. These reads were mapped against the Ensembl mouse version mm10 (GRCh38) with STAR 2.4.2a(63) using only unique alignments to exclude reads with unclear placing. The reads were further deduplicated to eliminate PCR artefacts leading to multiple copies of the same original fragment. The MACS2 peak caller v2.1.0.1 was used to identify peaks. The minimum q-value was set to −1.5 and the false discovery rate was changed to 0.01. To determine thresholds for significant peaks, the data were manually inspected in the Integrated Genomics Viewer (IGV) 2.3.52(5). Peaks overlapping blacklisted regions (known mis-assemblies, satellite repeats) from ENCODE were excluded. To compare peaks between samples, the resulting lists of significant peaks were overlapped and unified to represent identical regions. After conversion of binary alignment map (BAM) files to bigWig format with deepTools bamCoverage(51), the counts per unified peak per sample were computed with bigWigAverageOverBed (UCSC Genome Browser Utilities, http://hgdownload.cse.ucsc.edu/downloads.html). Raw counts for unified peaks were submitted to DESeq2 for normalization(54). Spearman correlations were produced to identify the degree of reproducibility between samples using R. To allow a normal distribution of samples in IGV, the raw BAM files were normalized for sequencing depth (number of mapped deduplicated reads per sample) and noise level (number of reads inside peaks). Two factors were computed and applied for sequencing depth (number of mapped deduplicated reads per sample) and noise level (number of reads inside peaks). Only reads above 30 nucleotides were cleared for further analyses. To avoid contamination with genomic DNA, samples were treated by on-column DNAse digestion (DNase-Free DNase Set, Qiagen). Total RNA and library integrity were verified on LabChip Gx Touch 24 (Perkin Elmer). One microgram of the RNA was isolated using the miRNeasy micro Kit (Qiagen). 1 μl of 50 mM MgCl2 was added, followed by purification of the DNA fragments using the MinElute PCR Purification Kit (Qiagen). Amplification of the library together with indexing was performed as previously described. Libraries were mixed in equimolar ratios and sequenced on a NextSeq500 platform using v2 chemistry.

RNA sequencing. RNA was isolated using the miRNeasy micro Kit (Qiagen). To avoid contamination with genomic DNA, samples were treated by on-column DNAse digestion (DNase-Free DNase Set, Qiagen). Total RNA and library integrity were verified on LabChip Gx Touch 24 (Perkin Elmer). One microgram of the total RNA was used as input for the SMARTer Stranded Total RNA Sample Prep Kit –Hi Mammalian (Clontech). RNA sequencing (RNA-seq) was performed on a NextSeq500 instrument (Illumina) using v2 chemistry, resulting in an average of 25 million–30 million reads per library, with 1 × 75 bp single-end setup.

RNA-seq analysis. The resulting raw reads were assessed for quality, adapter content and duplication rates with FastQC. Reaper v.1.3-100 was used to trim reads after a quality drop below a mean of Q20 in a window of 10 nucleotides(64). Only reads of at least 15 nucleotides were cleared for subsequent analyses. Trimmed and filtered reads were aligned against the Ensembl mouse genome version mm10 (GRCh38) using STAR 2.5.3a with the parameters ‘–outFilterMismatchNoverLmax 0.1 –alignIntronMax 200000’(53). The numbers of reads that aligned to genes were counted with featureCounts 1.6.0 from the Subread package(57) to avoid PCR artefacts leading to multiple copies of the same original fragment. The MACS2 peak caller v2.1.0.1 was used to identify peaks. The minimum q-value was set to −1.5 and the false discovery rate was changed to 0.01. To determine thresholds for significant peaks, the data were manually inspected in the Integrated Genomics Viewer (IGV) 2.3.52(5). Peaks overlapping blacklisted regions (known mis-assemblies, satellite repeats) from ENCODE were excluded. To compare peaks between samples, the resulting lists of significant peaks were overlapped and unified to represent identical regions. After conversion of binary alignment map (BAM) files to bigWig format with deepTools bamCoverage(51), the counts per unified peak per sample were computed with bigWigAverageOverBed (UCSC Genome Browser Utilities, http://hgdownload.cse.ucsc.edu/downloads.html). Raw counts for unified peaks were submitted to DESeq2 for normalization(54). Spearman correlations were produced to identify the degree of reproducibility between samples using R. To allow a normal distribution of samples in IGV, the raw BAM files were normalized for sequencing depth (number of mapped deduplicated reads per sample) and noise level (number of reads inside peaks). Two factors were computed and applied to the original BAM files using bedtools genomecov(53), resulting in normalized bigWig files for IGV.

RNA sequencing analysis. The resulting raw reads were assessed for quality, adapter content and duplication rates with FastQC. Reaper v.1.3-100 was used to trim reads after a quality drop below a mean of Q20 in a window of 10 nucleotides(64). Only reads of at least 15 nucleotides were cleared for subsequent analyses. Trimmed and filtered reads were aligned against the Ensembl mouse genome version mm10 (GRCh38) using STAR 2.5.3a with the parameters ‘–outFilterMismatchNoverLmax 0.1 –alignIntronMax 200000’(53). The numbers of reads that aligned to genes were counted with featureCounts 1.6.0 from the Subread package(57). Only reads mapping at least partially inside exons were admitted, and these reads were aggregated per gene. Reads overlapping with multiple genes or aligning to multiple regions were excluded. Differentially expressed genes were identified using DESeq2 v.1.14.1(8). Genes were classified as significantly differentially expressed with P < 0.05 (Wald test), without assigning a specific minimum or maximum regulation fold change as transcriptional adaptation might not necessarily lead to strong upregulation levels. The Ensembl annotation was enriched with UniProt data (release 24 March 2017) based on Ensembl gene identifiers.

Sequence similarity and subsampling analyses. To identify similarity to one of the three query nucleotide sequences (Fernit2, Actg1 and Actb), the longest respective transcript was selected (ENSMUST00000071555, ENSMUST00000045905 and ENSMUST00000100497) and compared to the whole genome using BLASTN(52). Genes were defined to be similar to the mutated gene body when a partial match was identified inside the target gene body or its promoter region (that is, 2 kb upstream of the TSS). Several alignment parameters were surveyed to identify the optimal degree of similarity. Alignment length, bit score and E value were queried to identify the dynamic range and optimal values using subsampling analyses (see below). E value denotes the probability that the match resulted by chance, while considering the whole target database (the genome in this case).

A subsampling approach was used to calculate a ranked P value for the significance of the percentage of upregulated genes in subsamples of a specific size (equivalent to the number of similar protein-coding genes) for each cell line model. In brief, the following algorithm was repeated 10,000 times: 1) get X random protein-coding genes, 2) calculate the percentage of upregulated genes in this subsample. The resulting list was filtered for subsamples with an equal or higher than expected number of upregulated genes according to a previous comparison (for example, for Fernit2, 18 protein-coding genes exhibit sequence similarity to its mRNA (= subset size), 9 of which were also upregulated (≈ expectation)). The number of subsamples showing at least as many upregulated genes as the expectation represents the rank of the comparison. The ranked P value was computed by dividing that rank by the total number of iterations (= 10,000). Optimal thresholds varied for the different cell line models and ranged between 1) 20 and 180 nucleotides (alignment length); 2) 40 and 200 bit score (combination of alignment quality and length); and 3) 10 to 6.7 × 10−50 (maximum E value).

We selected the following maximum E values from the optimal range for the follow-up similarity analyses: 5.1 for Fernit2 and Actg1, and 2 × 10−48 for Actb. The much stricter E value for Actb was necessary because of its repetitive 3’UTR, which resulted in misreading ‘noisy’ matches. These E-value thresholds translate into local nucleotide sequence alignments that range from 24 to 1,901 nucleotides in length, with 75 to 96% identity.

Sequence alignments of hif1ab, Actb and epas1a. The BLASTn alignments of the longest transcript of hif1ab (ENSDART00000018500) with the epas1a gene body including its promoter (2 kb upstream of TSS) were visualized using Kablam(55) at word size 7 and an E value of 25 (Extended Data Fig. 8). Two additional alignments were performed to show homologous regions of the synthetic hif1ab transcript compared to 1) the original source transcript (ENSDART00000018500, using MUSCLE(53)); and 2) epas1a, including its promoter (TSS-2000, using BLASTn). The uncapped RNA composed solely of the hif1ab sequences that are similar to epas1a is necessary because of the repetitive 3’UTR, which led to misreading ‘noisy’ matches. These E-value thresholds translate into local nucleotide sequence alignments that are similar to epas1a in 1,929 nucleotides in length. The similarity of the coding sequences of zebrafish actb1 transcript ENSDART00000054987 (query) to the mouse Actb transcript ENSMUST00000100497 (subject) was assessed with a MUSCLE alignment. All of these alignments can be found in the Supplementary Data.

Gene-set enrichment analyses. Genes that were strongly upregulated in all three knockout samples versus their respective wild-type samples (P < 0.05 (Wald test), log2(gene expression level in knockout cells/gene expression level in wild-type cells) > 2.086) were used for gene-set enrichment analyses using KOBAS(62).

Data availability
ATAC-seq and RNA-seq data were deposited to the Gene Expression Omnibus (GEO) under accession codes GSE107057 and GSE114122.
40. Gapuzan, M. E., Yufit, P. V. & Gilmore, T. D. Immortalized embryonic mouse fibroblasts lacking the RelA subunit of transcription factor NF-κB have a malignantly transformed phenotype. Oncogene 21, 2484–2492 (2002).

41. Ran, F. A. et al. Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. 8, 2281–2308 (2013).

42. Thakore, P. I. et al. Highly specific genome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. Nat. Methods 12, 1143–1149 (2015).

43. Martin, L. et al. Identification and characterization of small molecules that inhibit nonsense-mediated RNA decay and suppress nonsense p53 mutations. Cancer Res. 74, 3104–3113 (2014).

44. Rädle, B. et al. Metabolic labeling of newly transcribed RNA for high resolution gene expression profiling of RNA synthesis, processing and decay in cell culture. J. Vis. Exp. 78, e50195 (2013).

45. Sun, W. & Chen, W. Metabolic labeling of newly synthesized RNA with 4sU to in parallel assess RNA transcription and decay. Methods Mol. Biol. 1720, 25–34 (2018).

46. Vanhollebeke, B. et al. Tip cell-specific requirement for an atypical Gpr124- and Reck-dependent Wnt/β-catenin pathway during brain angiogenesis. eLife 4, e06489 (2015).

47. Kwon, H. B. et al. In vivo modulation of endothelial polarization by Apelin receptor signalling. Nat. Commun. 7, 11805 (2016).

48. Biecher-Gonen, R. et al. High-throughput chromatin immunoprecipitation for genome-wide mapping of in vivo protein-DNA interactions and epigenomic states. Nat. Protoc. 8, 539–554 (2013).

49. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods 10, 1213–1218 (2013).

50. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120 (2014).

51. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013).

52. Robinson, J. T. et al. Integrative genomics viewer. Nat. Biotechnol. 29, 24–26 (2011).

53. Ramirez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. 42, W187–W191 (2014).

54. Anders, S. & Huber, W. Differential expression analysis for sequence count data. Genome Biol. 11, R106 (2010).

55. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842 (2010).

56. Davis, M. P., van Dongen, S., Abreu-Goodger, C., Bartonicek, N. & Enright, A. J. Kraken: a set of tools for quality control and analysis of high-throughput sequence data. Methods 63, 41–49 (2013).

57. Love, M. I., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930 (2014).

58. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. J. Mol. Biol. 215, 403–410 (1990).

59. Wintersinger, J. A. & Wasmuth, J. D. Kablammo: an interactive, web-based BLAST results visualizer. Bioinformatics 31, 1305–1306 (2015).

60. Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. SMC Bioinformatics 5, 113 (2004).

61. Xie, C. et al. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Res. 39, W316–W322 (2011).
Extended Data Fig. 1 | Schematic illustration of the mutant alleles generated for this study. Partial DNA sequences of the different mutant alleles generated for this study, and images of gels providing evidence for deletions in RNA-less alleles. Red indicates mutation (asterisks, deletion; upper-case letters, substitution; lower-case letters, insertion); green indicates stop codon in alleles with a PTC; arrows indicate genotyping primers.
Extended Data Fig. 2 | See next page for caption.
Extended Data Fig. 2 | Transcriptional adaptation is independent of the loss of protein function. a, qPCR analysis of epas1a, epas1b, vegfab, emilin3a and alcamb mRNA expression levels in wild-type and hif1ab, vegfaa, egfl7 and alcama mutant embryos injected (inj.) with eGFP mRNA (control; ctrl) or wild-type hif1ab, vegfaa, egfl7 or alcama mRNA. b, qPCR analysis of vclb, epas1a, epas1b and emilin3a mRNA expression levels in vela, hif1ab and egfl7 wild-type, heterozygous and mutant zebrafish. c, qPCR analysis of hbegfa, hif1ab, vegfaa and alcama mRNA expression levels in hbegfa, hif1ab, vegfaa and alcama wild-type and heterozygous zebrafish, using primers specific for the wild-type allele. d, qPCR analysis of Fermt1 and Rel mRNA expression levels in wild-type and Fermt2 and Rela knockout cells transfected with empty vectors (control) or plasmids encoding wild-type FERMT2 or RELA. e, Western blot analysis of FERMT2 and ACTB levels in Fermt2 knockout cells transfected with empty vectors (control) or plasmids encoding wild-type FERMT2. f, Western blot analysis of RELA and ACTB levels in Rela knockout cells transfected with empty vectors (control) or plasmids encoding wild-type RELA. g, qPCR analysis of Actg1 mRNA expression levels in wild-type and heterozygous Actb mESCs. n = 3 biologically independent samples. Wild-type or control expression levels were set at 1 for each assay. Data are mean ± s.d., and a two-tailed Student’s t-test was used to calculate P values (a–d, g). The experiments in e, f were performed only once. For the source data for western blots, see Supplementary Fig. 1.
Extended Data Fig. 3 | Transcriptional adaptation involves enhanced transcription and is independent of the DNA lesion itself. a, qPCR analysis of hbegfb and emilin3a mRNA and pre-mRNA expression levels in hbegfa and egfl7 wild-type and mutant zebrafish. b, qPCR analysis of Fermt1 and Rel mRNA and pre-mRNA expression levels in Fermt2 and Rela wild-type and knockout cells. c, IGV tracks of the Fermt1 locus showing ATAC-seq signals in wild-type and Fermt2 knockout cells. d, qPCR analysis of hbegfa, hbegfb, egfl7 and emilin3a mRNA expression levels in hbegfa and egfl7 wild-type and Δ3 mutant zebrafish. e, qPCR analysis of vegfaa, vegfab, egfl7 and emilin3a mRNA expression levels in vegfaa and egfl7 wild-type and 5’UTR mutant zebrafish. f, qPCR analysis of vcla and vclb mRNA expression levels in vcla wild-type and last exon (exon 22) mutant zebrafish. n = 3 biologically independent samples. Wild-type expression levels were set at 1 for each assay. Data are mean ± s.d., and a two-tailed Student’s t-test was used to calculate P values.
Extended Data Fig. 4 | Reduction in mutant transcript levels is caused by mRNA decay. a, qPCR analysis of *hbegfa*, *egfl7* and *alcama* mRNA and pre-mRNA expression levels in *hbegfa*, *egfl7* and *alcama* wild-type and mutant zebrafish. b, qPCR analysis of *Fermt2* and *Rela* mRNA and pre-mRNA expression levels in *Fermt2* and *Rela* wild-type and knockout cells. c, qPCR analysis of 4sU-labelled *Fermt2*, *Rela* and *Actg1* mRNA and pre-mRNA expression levels in *Fermt2*, *Rela* and *Actg1* wild-type and knockout cells. d, Fitted exponential decay curves of *Fermt2* mRNA expression levels in wild-type and *Fermt2* knockout cells. $t_{1/2}$, half-life. e, Fitted exponential decay curves of *Rela* mRNA expression levels in wild-type and *Rela* knockout cells. f, Fitted exponential decay curves of *Actg1* mRNA expression levels in wild-type and *Actg1* knockout cells. $n = 3$ (a, b, d–f) or $n = 2$ (c) biologically independent samples. Wild-type expression levels were set at 1 for each assay (a–c). Data are mean ± s.d., and a two-tailed Student’s t-test was used to calculate $P$ values.
Extended Data Fig. 5 | See next page for caption.
Extended Data Fig. 5 | RNA decay induces transcriptional adaptation. a, qPCR analysis of \( \text{hbegfa} \), \( \text{vegfaa} \) and \( \text{vcla} \) mRNA expression levels in \( \text{upf1;hbegfa} \), \( \text{upf1;vegfaa} \) and \( \text{upf1;vcla} \) double mutant zebrafish. b, qPCR analysis of \( \text{Rela} \) mRNA expression levels after siRNA-mediated knockdown of the indicated proteins in \( \text{Rela} \) knockout cells. c, qPCR analysis of \( \text{Actb} \) mRNA expression levels after siRNA-mediated knockdown of the indicated proteins in \( \text{Actb} \) knockout cells. d, qPCR analysis of \( \text{hbegfa} \) mRNA expression levels in 6 dpf \( \text{hbegfa} \) mutants treated with NMD inhibitor (NMDi). e, qPCR analysis of \( \text{hbegfb} \) mRNA expression levels in 6 dpf \( \text{hbegfa} \) mutants treated with NMDi. f, qPCR analysis of \( \text{Rela} \) mRNA expression levels in \( \text{Rela} \) knockout cells treated with cycloheximide (CHX). g, qPCR analysis of \( \text{Rel} \) mRNA expression levels in \( \text{Rela} \) knockout cells treated with CHX. h, qPCR analysis of endogenous \( \text{hif1ab} \) and \( \text{vegfaa} \) mRNA expression levels in 6 hpf wild-type embryos injected with uncapped \( \text{hif1ab} \) or \( \text{vegfaa} \) RNA. I, qPCR analysis of \( \text{Actg1} \) mRNA expression levels in mESCs transfected with uncapped \( \text{Actb} \) RNA at different times after transfection. J, qPCR analysis of injected \( \text{hif1ab} \), \( \text{epas1a} \), \( \text{vegfaa} \) and \( \text{vegfab} \) RNA expression levels in 6 hpf wild-type embryos injected with uncapped \( \text{hif1ab} \) or \( \text{vegfaa} \) transcripts with or without a 5' XRN1-resistant (xr) sequence. k, qPCR analysis of \( \text{epas1a} \) and \( \text{vegfab} \) mRNA expression levels in 6 hpf wild-type zebrafish embryos injected with uncapped sense or antisense (rev) \( \text{hif1ab} \) or \( \text{vegfaa} \) RNA; the same \( \text{eGFP} \) uncap. control samples were used for the \( \text{epas1a} \) experiments. Wild-type or control expression levels were set at 1 for each assay (a–d, f, h–k). \( n = 3 \) biologically independent samples. Data are mean ± s.d., and a two-tailed Student's \( t \)-test was used to calculate \( P \) values.
Extended Data Fig. 6 | Mutant mRNA decay helps confer genetic robustness. **a**, qPCR analysis of Fermt2 and Fermt1 mRNA expression levels following CRISPR interference-mediated knockdown of Fermt2 transcription in Fermt2 knockout cells. **b**, qPCR analysis of emilin3a, emilin3b and emilin2a mRNA expression levels in 20 hpf wild types, egfl7del4 mutants and egfl7full locus del. mutants. **c**, Number of central arteries (CtAs) connecting to the basilar artery (BA) in 58 hpf vegfaaΔ10 and vegfaapromoter-less mutants. **d**, Blood-flow velocity in 78 hpf wild types, hbegfaΔ7 mutants and hbegfalfull locus del. mutants. **e**, Quantification of the cardiac ventricle length in 100 hpf wild types, alcama Δ8 mutants and alcama promotor-less mutants. Wild-type or control expression levels were set at 1 for each assay (a, b). n = 3 (a, b); n = 13 (vegfaaΔ10−/−) and 19 (vegfaapromoter-less−/−) (c); n = 25 (d); and n = 18 (wild-type siblings of alcamaΔ8−/−), 7 (alcamaΔ8−/−), 22 (wild-type siblings of alcama promotor-less−/−) and 15 (alcama promotor-less−/−) (e) animals. Data are mean ± s.d., and a two-tailed Student’s t-test was used to calculate P values.
Extended Data Fig. 7 | Analysis of sequence similarity parameters in models of transcriptional adaptation. a, Numbers of differentially expressed genes in the different knockout cell line models; \( P \leq 0.05 \); these genes are distributed throughout the genome (data not shown). b, Venn diagram of genes upregulated in the three different cell line models with L2F knockout > wild-type and \( P \leq 0.05 \). c, KEGG pathway enrichment analysis for genes commonly upregulated in Fermt2, Actg1 and Actb knockout compared to wild-type cells. The top ten pathways based on \( P \) value are displayed. The dashed line marks a \( P \) value of 0.05. Circle sizes provide an estimation of scale; outer grey circles represent the total number of genes in the pathway; and centred coloured circles represent the number of genes in the pathway that are commonly upregulated. d, Impact of various values of three different BLASTn alignment-quality parameters (alignment length, bit score and \( E \) value) on the significance of the observed correlation between upregulation and sequence similarity, and therefore the identification or prediction of putative adapting genes.

The \( E \) value describes the probability of the match resulting from chance (a lower value corresponds to a lower probability), and the bit score evaluates the combination of alignment quality and length (a higher value corresponds to a better alignment). The \( y \) axis of each diagram shows the negative \( \log_{10} \) of the \( P \) value and the \( x \) axis shows the respective parameter value. A \( P \) value of 0.05 is marked with a black horizontal line. The \( E \) value thresholds used in our analyses are highlighted with a circle. Lines ending preliminarily indicate a lack of any remaining alignments after that point. The first row of diagrams explores large variations of thresholds, in an attempt to identify the total range, whereas the second row focuses on the most relevant window for the three genes investigated. The optimal thresholds differ considerably depending on the gene analysed. \( n = 2 \) biologically independent samples. \( P \) value was computed by bootstrapping random subsamples (see the ‘Sequence similarity and subsampling analyses’ section of the Methods). \( P \) values were not corrected for multiple testing.
Extended Data Fig. 8 | See next page for caption.
Extended Data Fig. 8 | Expression level of genes exhibiting sequence similarity in the different mouse cell line models. a–c, RNA-seq analysis of genes exhibiting sequence similarity with Fermt2 (a), Actg1 (b) or Actb (c) in knockout compared to wild-type cells. Bold, significantly upregulated in knockout relative to wild-type cells; red, L2F > 0, blue, L2F < 0; green, P value or adjusted P value ≤ 0.05; purple; genes exhibiting sequence similarity with the mutated gene’s mRNA in their promoter region; yellow, genes exhibiting sequence similarity with the mutated gene’s mRNA in their 3′UTR region. Other non-coloured genes exhibit sequence similarity with the mutated gene’s mRNA in their exons or introns. Boxed, upregulated in knockout but not RNA-less cells; no Fermt2 RNA-less allele was analysed. d, qPCR analysis of Ubapl, Fmnl2, Cdk12 and Actr1a pre-mRNA expression levels in Actg1 knockout relative to wild-type cells. e, qPCR analysis of actb1 mRNA expression levels in 6 hpf wild-type zebrafish injected with uncapped mouse Actb RNA. f, Schematic representation of regions of sequence similarity between hif1ab mRNA and the epas1a locus. Grey shaded triangles represent the alignments; intensity represents the alignment quality; and width at the base represents the length of the similarity region. g, qPCR analysis of epas1a mRNA expression levels in 6 hpf wild-type zebrafish embryos injected with uncapped RNA composed solely of the hif1ab sequences similar to the epas1a promoter, exons, introns or 3′UTR; the same eGFP uncap. control samples were used for all comparisons. n = 2 (a–c) or n = 3 (d, e, g) biologically independent samples. DESeq2 tests were used to test for significance of coefficients in a negative binomial generalized linear model with the Wald test (a–c). P values were not corrected for multiple testing. Wild-type or control expression levels were set at 1 for each assay (d, e, g). Data are mean ± s.d., and a two-tailed Student’s t-test was used to calculate P values.
Extended Data Fig. 9 | See next page for caption.
Extended Data Fig. 9 | Transcriptional adaptation involves chromatin remodelling that is dependent on the activity of decay factors. a, qPCR analysis of Rel mRNA expression levels after siRNA-mediated knockdown of the indicated proteins in Rela knockout cells. b, ChIP–qPCR analysis of H3K4me3 occupancy at non-promoter regions (as a control) of Fermt1, Rel and Actg2 in Fermt2, Rela and Actg1 knockout cells, respectively, compared to wild-type cells. c, ChIP–qPCR analysis of H3K4me3 occupancy near the Rel TSS and a non-promoter region (as a control) after siRNA-mediated knockdown of the indicated proteins in Rela knockout cells. d, Current expanded model of transcriptional adaptation to mutations. RNA decay fragments may act as intermediates to bring decay factors and chromatin remodellers to adapting gene loci, thereby triggering increased gene expression. Alternatively, RNA decay fragments may function by repressing antisense RNAs at the adapting gene loci, thus allowing for increased sense mRNA expression. It is, however, likely that additional mechanisms are involved in transcriptional adaptation, and possibly in a gene-dependent manner. n = 3 (a) or n = 2 (b, c) biologically independent samples. Data are mean ± s.d., and a two-tailed Student’s t-test was used to calculate P values.
Extended Data Fig. 10 | The potential role of antisense transcripts in the transcriptional adaptation response. a, qPCR analysis of Cdk9 and Sox9 mRNA expression levels in cells transfected with uncapped Cdk9 or Sox9 RNA. b, qPCR analysis of BDNF and BDNF-AS mRNA expression levels in HEK293T cells transfected with uncapped BDNF RNA. c, Integrated genome viewer tracks of vclb and hbegfb loci, showing the location of the annotated antisense transcripts. Two alignments of 105 and 147 bp were observed between vclb mRNA and vclb antisense RNAs, and an alignment of 39 bp was observed between hbegfa mRNA and hbegfb antisense RNA. Antisense transcripts shown were acquired from the datasets in GSE32898. d, qPCR analysis of vclb and hbegfb antisense (AS) RNA expression levels in vcla and hbegfa wild-type and mutant zebrafish at 24 and 72 hpf, respectively. Control expression levels were set at 1 for each assay. n = 3 biologically independent samples. Data are mean ± s.d., and a two-tailed Student’s t-test was used to calculate P values.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
- Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection  No Software was used.

Data analysis  No Software was used.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

ATACseq and RNAseq data were deposited to Gene Expression Omnibus under accession codes GSE107075 and GSE114212.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf
Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

No statistical methods were used to predetermine sample size. Sample size was determined according to the minimal number of independent biological replicates that significantly identified an effect. For most qPCR analyses we analyzed 3 sets of biological samples (ie, zebrafish or cultured cells). The variability between the three biological replicates was very minimal and therefore did not require increasing the sample size.

Data exclusions

No data were excluded.

Replication

Experiments in this study were independently replicated, with biological and technical replicates. All attempts at replication were successful, verifying the reproducibility of the findings.

Randomization

The experiments were not randomized due to the required experimental set up (ie, in most experiments, we first had to genotype the zebrafish in order to pool samples).

Blinding

The investigators were not blinded to allocation during experiments and outcome assessment for the reasons mentioned above.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

| n/a | Involved in the study |
|-----|-----------------------|
| ☐   | Antibodies            |
| ☑   | Eukaryotic cell lines |
| ☑   | Palaeontology         |
| ☑   | Animals and other organisms |
| ☑   | Human research participants |
|    | Clinical data         |

Methods

| n/a | Involved in the study |
|-----|-----------------------|
| ☑   | ChIP-seq              |
| ☑   | Flow cytometry        |
| ☑   | MRI-based neuroimaging |

Antibodies

Antibodies used

The following antibodies were used: KINDLIN-2 (Millipore, MA82617; clone 3A3, 1:1,000, Lot. 2739149), RELA (Cell Signalling Technology, #8457, 1:1,000), β-ACTIN (Cell Signalling Technology, #8457, 1:1,000), anti-mouse IgG-HRP (Thermo, 31430, 1:10,000, Lot. SG253594), anti-rabbit IgG-HRP (Thermo, 31460, 1:10,000, Lot. TG266717), rabbit IgG (4 μg/IP, # 026102, Thermo Fischer, Lot.: TD268143), WDR5 (4 μg/IP, #13105, Cell signaling) and H3K4me3 (4 μg/IP, #9751, Cell signaling, Lot. 10).

Validation

All antibodies were obtained from the indicated commercial vendors and have been validated for the application by the manufacturer to ensure quality. β-ACTIN, KINDLIN-2 and RELA antibodies were also validated by the authors in the sense that they do not recognize any protein in β-Actin, Kindlin-2 and Rela knockout mouse cells, respectively. Previous ChIP-seq data in mouse cells using the H3K4me3 antibody revealed enrichment at promoter regions (Sodersten et al. PLOS Genetics, 2014). In addition, all the antibodies detected intended proteins in control samples with the expected molecular weight. Specific references for each antibody can be found on the suppliers’ homepage.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

mESCs from the CS78L/6 mouse strain were a generous gift from Johnny Kim, MPI for heart and lung research, Bad Nauheim, Germany.

Wild-type and Rela knockout MEFs were a generous gift from Alexander Hoffmann, UCLA, USA.

Wild-type and Kindlin-2 knockout MKFs were a generous gift from Reinhard Fässler, MPI for biochemistry, Martinsried, Germany.

Authentication

None of the cell lines were authenticated by the authors.

Mycoplasma contamination

All cell lines tested negative for mycoplasma contamination.

Commonly misidentified lines (See ICLAC register)

No commonly misidentified cell lines were used.
## Animals and other organisms

Policy information about [studies involving animals](https:// ARRIVE guidelines](https://) recommended for reporting animal research

| Laboratory animals                  | Zebrafish (Danio rerio), strain: Tüb/AB. All experiments were performed on zebrafish embryos or larvae between 6 hpf and 6 dpf before sex is specified. |
|-------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|
| Wild animals                        | The study did not involve wild animals.                                                                                               |
| Field-collected samples             | The study did not involve samples collected from the field.                                                                           |
| Ethics oversight                    | All zebrafish husbandry was performed under standard conditions in accordance with institutional (Max Planck Gesellschaft) and national ethical and animal welfare guidelines approved by the ethics committee for animal experiments at the Regierungspräsidium Darmstadt, Germany (permit number: B2/1017). |

Note that full information on the approval of the study protocol must also be provided in the manuscript.