Elongator and codon bias regulate protein levels in mammalian peripheral neurons

Joy Goffena1, Frances Lefcort2, Yongqing Zhang3, Elin Lehrmann3, Marta Chaverra2, Jehremy Felig1, Joseph Walters1, Richard Buksch1, Kevin G. Becker3 & Lynn George1

Familial dysautonomia (FD) results from mutation in IKBKAP/ELP1, a gene encoding the scaffolding protein for the Elongator complex. This highly conserved complex is required for the translation of codon-biased genes in lower organisms. Here we investigate whether Elongator serves a similar function in mammalian peripheral neurons, the population devastated in FD. Using codon-biased eGFP sensors, and multiplexing of codon usage with transcriptome and proteome analyses of over 6,000 genes, we identify two categories of genes, as well as specific gene identities that depend on Elongator for normal expression. Moreover, we show that multiple genes in the DNA damage repair pathway are codon-biased, and that with Elongator loss, their misregulation is correlated with elevated levels of DNA damage. These findings link Elongator’s function in the translation of codon-biased genes with both the developmental and neurodegenerative phenotypes of FD, and also clarify the increased risk of cancer associated with the disease.
The genetic code consists of trinucleotide units termed codons, with each of these units encoding a specific amino acid. In all species of life, the same amino acid can be encoded by up to six different trinucleotide combinations, or synonymous codons. The reason for this redundancy is not well understood, although, interestingly, synonymous substitutions can impair protein function. Codon usage has recently been spotlighted as a key determinant of translation elongation rates and co-translational protein folding, with preferred codons enhancing translational efficiency and folding fidelity. The unequal usage of synonymous codons is referred to as codon bias and the universal nature of this bias, from yeast to humans, suggests the existence of a secondary code within the more familiar genetic code. This secondary code is emerging as a major regulator of translational speed and co-translational protein folding and thereby a significant determinant of the cellular levels of specific proteins.

A key component of the cellular machinery used to interpret codon bias is the molecular complex Elongator, composed of 6 subunits (ELP1-6), which chemically modifies transfer RNAs (tRNAs). Within the ribosomal A site, the tRNA anticodon loop forms Watson–Crick base pairs with its cognate messenger RNA codon. Flexibility at the first anticodon position allows for non-standard (wobble) pairing and, in particular, uridines in the wobble position (U34) are capable of recognizing multiple nucleotides. Specifically, lysine, glutamine, and glutamic acid are each encoded by two trinucleotide combinations that end in the standard (wobble) pairing and, in particular, uridines in the wobble position (U34) can recognize both the AA- and AG-ending (wobble) versions of these codons (Fig. 1). Chemical modifications to U34, including the addition of both a thiol (mcm5) and a methoxy-carbonyl-methyl (mcm5s2) refine U34 recognition and specifically enhance the translational efficiency of problematic AA-ending codons that impose restrictive codon–anticodon interactions (Fig. 1)4–8. The impact of the mcm5s2 modification on the translation of the AG-ending (wobble) codons for these amino acids is not completely understood, with some studies suggesting it has no impact and others indicating that it decreases AG translational efficiency. Multiple studies in a variety of model organisms, as well as in brain tissue and fibroblasts from patients with familial dysautonomia (FD) that carry mutations in IKBKAP/ELP1, have shown that Elongator is essential for the synthesis of the mcm5 modification. In addition, thiolation of the wobble uridine depends on the addition of the same mcm5 moiety, such that neither modification is present in the absence of Elongator. Importantly, in Saccharomyces cerevisiae and Caenorhabditis elegans, loss of tRNA U34 modification leads to widespread protein aggregation and imbalanced protein homeostasis, and multiple studies in fission yeast have demonstrated that proteins encoded by AA-biased transcripts are depleted in Elongator (Elp) mutants.

To gain a better understanding of the role that codon bias may have in mammalian neurons, we conducted a comprehensive analysis of protein levels in the dorsal root ganglia (DRG) of mice in which Ikbkap/Elp1 is conditionally ablated in the peripheral nervous system (PNS). Although loss of any of the Elongator subunits is sufficient to disrupt function of the complex, we used an Ikbkap/Elp1 knockout, because mutations in the human gene cause FD. The primary Ikbkap mutation in FD causes tissue-specific exon skipping, mRNA nonsense-mediated decay, and reduced levels of the encoded IKAP/ELP1 protein, with the PNS being the most gravely impacted. In Wnt1-Cre; Ikbkap-fllox/LoxP conditional knockout (CKO) mice, Ikbkap expression is selectively ablated in the PNS. These CKO mice exhibit reduced levels of mcm5s2-U tRNA modification (Supplementary Fig. 1) and recapitulate many FD hallmarks including a depleted number of pain and temperature-sensing TrkA+ neurons in the DRG, a primary site of Ikbkap expression. By comparing protein levels within the DRG of CKO and control mice, we demonstrate that the differential usage of AA- and AG-ending codons, in coordination with the Elongator complex, function together to regulate the levels of particular proteins in the mammalian PNS that are ultimately essential for neuron survival.

**Results**

**Codon-biased eGFP is misexpressed in CKO neurons.** To directly investigate a role for the Elongator complex in the translation of codon-biased transcripts in mammals, we engineered a codon-biased version of enhanced green fluorescent protein (eGFP) that exclusively uses AA-ending codons for the amino acids Lys, Gln, and Glu, and compared the expression level of this construct with mammalian optimized eGFP in which 42 out of 44 codons for these same amino acids end in AG. In neuron cultures from both control and CKO embryos, the AA-biased eGFP was expressed at a significantly lower level than the AG-biased construct (Fig. 2), indicating that even in the presence of the mcm5s2 modification, restrictive AA codon–anticodon interactions impede protein production. Somewhat surprisingly, there was no difference in the amount of this reduced expression between CKO and control neurons. This may have been due to the small size of the eGFP gene (see below). Unexpectedly however, AG-biased eGFP was expressed at a significantly higher level in CKO neurons compared with controls (Fig. 2). This raises the possibility that the mcm5s2 modification present in the control may indeed reduce the ability of U34 to pair with G-ending (wobble) codons, such that the absence of the modification actually increases the efficiency of AG translation, although further kinetic data are required to show this.

**Codon bias and mRNA size influence Elongator dependence.** Across both the mouse and human genomes, AG-ending codons for Lys, Gln, and Glu are preferentially used over AA-ending codons (~64% of all such codons end in AG). This bias is often measured as the number of AA-ending codons over the number...
of AG-ending codons, or the AA:AG ratio, with a ratio of 1 indicating an unbiased usage. In mouse and humans, the average AA:AG ratio across all open reading frames is 0.67 and 0.76, respectively (ftp://ftp.ncbi.nlm.nih.gov/pub/CCDS/). To visualize this bias toward the use of AG-ending codons, we plotted the total number of coding sequences as a function of AA:AG ratio. A total of 84.1% of all coding sequences in mouse and 74.2% in humans exhibit an AA:AG ratio < 1 (Fig. 3a). To further investigate a mammalian role for Elongator in the translation of codon-biased genes, we conducted a comprehensive proteomics analysis. DRG from control and CKO embryos were pooled, respectively (n = 7 embryos per genotype), and the proteins subjected to ultra-high-pressure liquid chromatography tandem mass spectrometry (UHPLC–MS/MS). To identify genes that were specifically misregulated at a posttranscriptional level, we overlaid the resulting proteome data with a transcriptome analysis of genes expressed in these same tissues (17.5 DRG) (GSE80130) and filtered the combined data sets to identify genes that were detected in both the transcriptome and proteome analyses, and that were transcribed normally (average fold change ≤ 1.5). Genes that were detected in one sample type (either control or CKO), and not detected in the counterpart sample, were omitted. The remaining 6,448 genes (Supplementary Data 1) were then analyzed for codon usage. Total transcript length in number of codons was also considered.

Using this data set, we analyzed the behavior of protein expression levels as a function of codon bias and found that increasing AA bias correlates with an increasing likelihood that a protein will exhibit decreased expression (fold change ≤ −2) (Fig. 3b and Supplementary Data 2). Importantly, we also found that large AA-biased genes are more likely to be impacted by Elongator loss than are small AA-biased genes, presumably because large genes contain the highest numbers of AA-ending codons. In mouse and humans, the average mRNA coding sequence consists of 556 and 572 codons, respectively (https://www.ncbi.nlm.nih.gov/CCDS/CdcsBrowse.cgi). As shown in Fig. 3b, 55.0% of proteins encoded by transcripts ≥ 1,755 codons (5,265 bp) with an AA:AG ratio > 1.5 are depleted in the CKO, as compared with only 26.5% of proteins with the same AA:AG ratio, but a transcript length ≥ 505 codons. Considering the small size of the eGFP coding sequence (240 codons), these data may explain why the expression of AA-biased eGFP was comparable between control and CKO neurons in our in vitro experiments (Fig. 2).

Proteins expressed normally (fold change > −2 and < 2) were also analyzed as a function of increasing AA-bias. Figure 3c shows that the ability to maintain normal expression levels becomes increasingly dependent on Elongator as AA-bias and transcript length increase. To assess the impact of total AA number in the context of no bias, we also plotted the percentage of depleted proteins as a function of increasing AA number, irrespective of the AA:AG ratio. When transcripts of all AA:AG ratios are considered, the percentage of depleted proteins is consistently about 27%, regardless of AA number (Fig. 3d). These data indicate that AA-bias (a high number of AA-ending codons in the context of a low number of AG-ending codons), rather than AA number, determines Elongator dependence.

We next examined the impact of Elongator loss on the expression levels of AG-biased genes. Interestingly, we found that the percentage of proteins that are upregulated (fold change ≥ 2) increases with increasing AG-bias (Fig. 3e) (Supplementary Data 2). These data suggest that the presence of mcm5 in control neurons decreases the translational output of AG-ending codons, whereas in Elongator knockout cells that lack the modification, highly AG-biased transcripts are translated more efficiently. These data also show that small, AG-biased transcripts are more likely to be impacted by Elongator loss than are large, AG-biased transcripts. This conclusion supports our in vitro observation that the expression of AG-biased eGFP, at only 240 codons, increases in the absence of Elongator (Fig. 2). These data are also consistent with a study in yeast elp3 mutants showing small decreases in ribosome occupancy for AG-ending codons and support the hypothesis that an mcm5 modification may reduce the ability of U34 to pair with G-ending codons. In addition, these data also explain why AA abundance in the context of an equally high number of AG-ending codons did not have an impact on protein levels in our study (Fig. 3d); the decreased translation rate of AA-ending codons may have been offset by an increased translation rate of an equally large number of AG-ending codons.

In addition to the nervous system, Elp genes are expressed somewhat ubiquitously in numerous other tissue types. To gain insight into the frequency with which codon bias might be generally used in the regulation of protein levels within mice and humans, we also filtered both ORFeomes with parameters that are indicative of Elongator dependence (Fig. 3). Approximately 3.5% of mouse coding sequences, and 5% of human coding sequences fall within our most stringent parameters (a transcript size ≥ 1,755 total codons with an AA:AG ratio > 1.3, or a transcript size ≤ 150 total codons with an AA:AG ratio < 0.3) (Supplementary Data 3). Approximately 12% of mouse and 18% of human coding sequences fall within slightly less stringent size parameters (a transcript size ≥ 1,005 total codons with an AA:AG ratio > 1.3, or a transcript size ≤ 300 total codons with an AA:AG ratio < 0.3) (Supplementary Data 4). These data suggest that codon bias serves as a basic, somewhat pervasive mechanism for regulating protein levels within mammals.

**Large, AA-biased Elongator targets.** Having demonstrated that transcript length and AA:AG ratio regulate protein levels in an Elongator-dependent manner, we next investigated the identity of specific Elongator targets. For large, AA-biased genes, the proteome data set of normally transcribed genes was filtered for targets that consisted of ≥ 1,755 total codons with an AA:AG ratio...
To investigate whether Elongator loss and the diminished levels of both proteins in DRG neurons (Fig. 4 and Supplementary Fig. 2), we used quantitative IHC and RT-PCR to confirm decreased Elongator dependence. Quantitative reverse transcriptase-PCR (RT-PCR) confirmed normal transcription of both genes in the absence of Elongator loss and the reduced expression of less than half of these large, heavily AA-biased genes suggests that other posttranscriptional regulatory pathways (such as ubiquitination) may work to counteract the effects of Elongator loss for some genes. Alternatively, other as of yet unidentified gene characteristics may work in concert with codon bias and transcript size to mediate full Elongator dependence. Of the 7 under-expressed genes, Brca2 and Vcan consist of over 3,300 total codons with 425 (Brca2) and 374 (Vcan) of those codons ending in AA (average number of AA-ending codons = 36 in mouse and 40 in humans) and an AA:AG ratio > 1.5, making them particularly compelling as potential candidates for Elongator dependence. Quantitative reverse transcriptase-PCR (RT-PCR) confirmed normal transcription of both genes in the CKO and quantitative immunohistochemistry (IHC) confirmed diminished levels of both proteins in DRG neurons (Fig. 4 and Supplementary Fig. 2).

Brca2 has a key role in homology-directed repair (HDR) of DNA double-strand breaks. Interestingly, our proteome analysis indicated the possible misregulation of multiple other DNA damage response genes, with some of those genes, including Setx and Rbbp8/Ctip, being strong candidates for direct Elongator regulation (Supplementary Data 6). As mutations in the human SETX gene cause multiple neurodegenerative disorders, we used quantitative IHC and RT-PCR to confirm decreased SETX in the context of normal transcription (Supplementary Fig. 3).

HDR is specifically required by neural progenitor cells during proliferation. To investigate whether Elongator loss and the resulting misregulation of Brca2 and other DNA damage response genes leads to elevated levels of DNA damage during neurogenesis, we performed comet assays on control and CKO DRG at E12.5, during the peak of sensory progenitor proliferation. As shown in Figure 4, Wnt1-Cre; IkbkaploxP/loxP embryos indeed exhibit significantly elevated levels of fragmented DNA at E12.5. Elevated DNA damage is also associated with amyotrophic lateral sclerosis (ALS), a disease that strictly affects post-mitotic neurons that no longer have access to HDR. As mutations in the Elongator subunit ELP3 are associated with ALS, we also analyzed the codon usage of genes required for non-homologous end-joining (NHEJ), the primary DNA repair pathway used by post-mitotic cells (Supplementary Data 7). Three genes, Rif1, Pola1, and Prkdc, fall out as likely targets for downregulation in the absence of Elongator. In particular, Rif1 is a compelling candidate in that it is required for NHEJ and for suppression of the error-prone microhomology-mediated end joining pathway (MMEJ).

**Small, AG-biased Elongator targets.** We also investigated the identity of small, AG-biased Elongator targets. By filtering normally transcribed genes for an AA:AG ratio < 0.2 and transcript length of ≤ 150 codons, we narrowed our target pool to 49 genes, with 19 being upregulated, 27 expressed normally, and 3 being downregulated (Supplementary Data 8). Of the 19 upregulated genes, 9 encode different H2A histones (both type 1 and type 2), with each comprising fewer than 150 total codons including 2 or fewer AA-ending and 23–29 AG-ending (AA:AG ratios between 0.00 and 0.08). Hist1h2ab and Hist1h2aa1 were chosen as type 1 and type 2 representatives to verify normal transcription and
confirm upregulation at the protein level via quantitative IHC (Fig. 5). Although we did not detect significant misregulation of H2B histones, the extreme AG bias exhibited by the majority of both H2a and H2b genes suggests that histones may represent a family of functionally-related genes whose expression is coordinated by Elongator20.

**Upregulation of UPR genes in Ikbkap CKO embryos.** Loss of Elongator function in both yeast and mice has also been shown to cause ribosomal pausing and translational stress, which in turn triggered the unfolded protein response (UPR)27,38. Consistent with these findings, our transcriptome and proteome studies indicate similar changes induced by loss of Ikbkap/Elp1 in the developing PNS. Transcript levels of UPR genes Atf5, Ddit, Atf4, Egr1, Gadd45a, Herpud1, Trib3, and Eif2s2 were all significantly increased (GSE80130); out of these eight upregulated genes, three of them were also detected in our proteome analysis; DDIT3, EGR1, and EIF2S2 were increased 9.7, 10.8, and 2.1-fold, respectively (not shown).

**Gene Ontology analysis of misregulated proteins.** We determined Gene Ontology (GO) biological processes for genes that were strong candidates for regulation by Elongator including downregulated proteins encoded by transcripts with an AA:AG >
thermoceptors) are selectively depleted, whereas large diameter temperature subset of DRG neurons (TrkA) is a hallmark of the disease is a failure in development of the PNS; compromised. Importantly, our finding that Elongator loss leads to reduced expression of Brca2 and other genes needed to repair DNA damage sheds new light on the pathogenesis of FD. An important hallmark of the disease is a failure in development of the PNS; autopsy studies on infants show that the small diameter pain and temperature subset of DRG neurons (TrkA+ nociceptors and thermoceptors) are selectively depleted, whereas large diameter DRG neurons (TrkC+ proprioceptors) are less affected. An outstanding question in the FD field has been to understand why the small diameter TrkA+ population exhibits this increased susceptibility. A major difference between the two subsets is that TrkC+ neurons are born early, during a limited first wave of neurogenesis, whereas TrkA+ cells are born during a second, extended wave of neurogenesis from mitotically active progenitors that go through numerous rounds of cell division to generate more than 70% of the final DRG neuron population. We previously showed that these specific neural progenitors (Pax3+), as well as fully differentiated TrkA+ neurons, selectively exhibit elevated levels of apoptosis in the absence of Ikbpap/Elp1. The data presented here suggest that the decreased number of small diameter TrkA+ neurons in FD likely results from DNA damage that accumulates in mitotically active progenitor cells as they attempt to generate the massive nociceptor and thermoceptor populations. This conclusion is supported by studies showing a specific requirement for Brca2 and other homologous recombination repair genes during neural progenitor proliferation in the central nervous system. As Vcan encodes an extracellular matrix proteoglycan that also functions in neuronal proliferation and differentiation, its deletion may exacerbate defects in neurogenesis imposed by compromised DNA damage response pathways.

An accumulation of damaged DNA over time could also contribute to the neurodegenerative hallmarks of FD including the progressive gait ataxia and death of retinal ganglion cells as patients age, as well as to the progressive neuronal apoptosis observed in FD model mice. Genome maintenance poses a unique challenge for terminally differentiated neurons that go through their final cell division during embryogenesis and must therefore maintain genome integrity for many decades in spite of the thousands of DNA damage events they experience. As HDR is not accessible to post-mitotic cells, neurons depend on NHEJ and other genes needed to repair DNA damage. Therefore, maintain genome integrity for many decades in spite of the thousands of DNA damage events they experience. As HDR is not accessible to post-mitotic cells, neurons depend on NHEJ and other genes needed to repair DNA damage.

Discussion

In addition to FD, which results from a splice site mutation in Ikbkap/Elp1, allelic variation in other Elongator subunits is associated with ALS (Elp3), intellectual disability (Elp2 and 4), autism spectrum disorder (Elp4), and Rolandic epilepsy (Elp4). This study provides the first insight into how Elongator and codon bias may contribute to both the neurodevelopmental and neurodegenerative aspects of these diseases. We describe a novel mechanism for impacting gene expression directly at the protein level and demonstrate that the absence of this mechanism leads to misregulation of cellular pathways that are essential to neuron development, function, and survival. In total, our findings indicate that codon bias serves as a significant mechanism for regulating protein output, with grave cellular repercussions when the ability to interpret codon bias is compromised.

Importantly, our finding that Elongator loss leads to reduced expression of Brca2 and other genes needed to repair DNA damage sheds new light on the pathogenesis of FD. An important hallmark of the disease is a failure in development of the PNS; autopsy studies on infants show that the small diameter pain and temperature subset of DRG neurons (TrkA+ nociceptors and thermoceptors) are selectively depleted, whereas large diameter DRG neurons (TrkC+ proprioceptors) are less affected. An outstanding question in the FD field has been to understand why the small diameter TrkA+ population exhibits this increased susceptibility. A major difference between the two subsets is that TrkC+ neurons are born early, during a limited first wave of neurogenesis, whereas TrkA+ cells are born during a second, extended wave of neurogenesis from mitotically active progenitors that go through numerous rounds of cell division to generate more than 70% of the final DRG neuron population. We previously showed that these specific neural progenitors (Pax3+), as well as fully differentiated TrkA+ neurons, selectively exhibit elevated levels of apoptosis in the absence of Ikbkap/Elp1. The data presented here suggest that the decreased number of small diameter TrkA+ neurons in FD likely results from DNA damage that accumulates in mitotically active progenitor cells as they attempt to generate the massive nociceptor and thermoceptor populations. This conclusion is supported by studies showing a specific requirement for Brca2 and other homologous recombination repair genes during neural progenitor proliferation in the central nervous system. As Vcan encodes an extracellular matrix proteoglycan that also functions in neuronal proliferation and differentiation, its deletion may exacerbate defects in neurogenesis imposed by compromised DNA damage response pathways.

An accumulation of damaged DNA over time could also contribute to the neurodegenerative hallmarks of FD including the progressive gait ataxia and death of retinal ganglion cells as patients age, as well as to the progressive neuronal apoptosis observed in FD model mice. Genome maintenance poses a unique challenge for terminally differentiated neurons that go through their final cell division during embryogenesis and must therefore maintain genome integrity for many decades in spite of the thousands of DNA damage events they experience. As HDR is not accessible to post-mitotic cells, neurons depend on NHEJ and the more error-prone MMEJ pathway. If compromised expression of codon biased genes such as Rif1 were to shift neuronal...
DNA repair toward a heavier use of MMEJ, accumulated genetic damage would ensue. Indeed, an accumulation of DNA damage as a consequence of unchecked MMEJ is suspected to be the basis for neurodegeneration and tumor formation in ataxia telangiectasia52.

A compromised ability to deal with DNA damage could also clarify why individuals born with FD are more than twice as likely to develop cancer during their shortened life span (40 years) compared with the general population, as well as why cancer onset occurs at a young age; the average age of tumor diagnosis in FD is < 30 years, whereas cancers in non-FD children and young adults are very rare (~ 3.5%)53. The highly codon biased nature of multiple DNA damage response genes likely also explains why 20% of cancers occurring in FD patients are of neural crest origin, the cellular lineage in FD with the lowest levels of correctly spliced IKBKAP transcript24,33. Independent of FD, the Elongator subunit ELP3 has also been characterized as a tumor suppressor gene in studies correlating diminished ELP3 levels with a poor prognosis in cancer54–56. Thus, our findings provide a specific molecular pathway via which ELP3, the catalytic subunit of Elongator, may accomplish this tumor suppressor function.

Previous studies have demonstrated that the use of codon bias by families of functionally related genes helps coordinate and/or synchronize their expression19,20. It is also possible that by altering the levels of different tRNA modifications, a cell could quickly ramp up or slow down the production of an entire class of functionally related proteins, such as those needed to respond to DNA damage.

Encoded by mRNAs consisting of more than 3,300 total codons, both Brca2 and Vcan fall within the largest 1% of all genes detected in our proteome analysis out of more than 17,000 genes (not shown). Their transcripts also contain an extreme number of AA-ending codons, 425 and 374, respectively, and are highly AA-biased (AA:AG ratio = 1.5 for both). Depletion of proteins with this specific genetic profile in the absence of Elongator has not been previously described, and demonstrates a connection between codon usage, transcript length, and tRNA modification in determining the expression levels of specific neuronal proteins. Post-transcriptional changes in less biased as well as non-codon-biased genes were also present in our data set, suggesting that other, unidentified factors may function in concert with codon bias and transcript length to determine the overall impact of Elongator loss. Downstream effects from misregulation of other post-transcriptional pathways, such as the misregulation of numerous codon biased genes that function in ubiquitination (Supplementary Data 2), may also amplify the impact of elongator loss and at least in part explain the misregulation of non-codon-biased genes.

Recent studies have shown that acquisition of proper protein tertiary structure requires not only the correct amino acid sequence, but also a designated elongation rate as a growing peptide chain emerges from the ribosome1–3. Thus, for large AA-biased genes such as Brca2 and Vcan, loss of Elongator would presumably lead to depleted protein levels due to both a slower translation rate of AA-ending codons, as well as polyubiquitylation and proteosomal misfolding of misfolded proteins. Our data also support a model where altered translation rates and misfolding lead to ER stress and the UPR, which in turn could contribute to a loss of neuronal progenitors during development, as well as to the degeneration of mature neurons, both of which occur in mouse models of FD52,57.

In total, our study is the first to demonstrate the ramifications of codon bias in neurons. Our work indicates that a delicate interplay between tRNA modification and the genetic ratio of AA-ending to AG-ending codons may fine tune the translational efficiency of specific neuronal transcripts. Although multiple codons may indeed encode the same amino acid, our findings demonstrate that synonymous codons are not equivalent, with codon selection playing a key role in translational efficiency. In summary, our work highlights a new stratum of the genetic code wherein the differential use of synonymous codons serves as a regulator of specific protein levels, and underscores how breakdown of the machinery that reads this code can contribute to neurologic disease.

**Methods**

**tRNA modification analysis.** DRG were collected and pooled from four control and four CKO embryos (E17.5) per biological replicate. Total RNA was extracted using TRIzol (Life Technologies). The mRNA was isolated from total RNA by Urea-PAGE, followed by hydrolysis to single nucleosides and depurination. The nucleoside solutions were desalted using Satorius 10,000 Da MWCO (molecular weight cut-off) spin filters. LC–MS analysis was performed on an Agilent 6460 QQQ mass spectrometer with an Agilent 1290 HPLC system using Multi reaction monitoring detection mode. Peak information of modified nucleosides for each sample was extracted using Agilent Qualitative Analysis software. Peaks with a signal-to-noise ratio ≥ 10 were considered significant and reportable. Peak areas were then normalized to the quantity of purified tRNA for each sample.

**Production of codon-biased eGFP biosensors.** The mammalian optimized eGFP coding sequence, in which 42 out of 44 codons for Lys, Glu, and Gln are AG ending (AA:AG = 0.95) was used as the AG-biased biosensor (Vigene Biosciences, CV10002). For the AA-biased sensor, an eGFP construct in which all Lys, Glu, and Gln codons in the coding sequence, in which 42 out of 44 codons for Lys, Glu, and Gln are AG ending (AA:AG = 0.95) was used as the AG-biased biosensor (Vigene Biosciences, CV10002). Previous studies have demonstrated that the use of codon bias by families of functionally related genes helps coordinate and/or synchronize their expression19,20. It is also possible that by altering the levels of different tRNA modifications, a cell could quickly ramp up or slow down the production of an entire class of functionally related proteins, such as those needed to respond to DNA damage.

**In vitro neuronal cultures and lentiviral infection.** For culturing of DRG primary neurons, DRG from E17.5 embryos (Wnt1-Cre; Ikbkapfl/fl × Lcfl/fl; X IkbkapLoxp/Loxp) were collected and pooled into 1.5 ml microcentrifuge tubes containing neutral basal media on ice. Following genotyping, DRG from five control (Wnt1-Cre; Ikbkapfl/fl × Lcfl/fl; X IkbkapLoxp/Loxp) and five ELP3–/– (Wnt1-Cre; Ikbkapfl/fl × Lcfl/fl; X IkbkapLoxp/Loxp) embryos were pooled into 15 ml conical tubes containing neutral basal media. Cells were spun down at 1000 r.p.m. for 1 min at 4 °C. After decanting the neutral basal media, cells were incubated in 2 ml of 0.05% trypsin + EDTA (Thermo Fisher) at 37 °C for 15 min. Cells were then spun down as above and the trypsin quenched by addition of 2 ml 10% fetal bovine serum. After a final spin, cells were brought up in complete media containing 2% B-27, 2 mM L-Glutamine, 10 ng ml–1 NGF, 10 ng ml–1 NT3, and 10 ng ml–1 BDNF (brain-derived neurotrophic factor) in neutral basal media and triturated 20 times using a hemocytometer pipettes of varying bore sizes. Once dissociated, cells were counted using a hemocytometer and plated at ~ 50,000 cells per well in an 8-well glass bottom chamber slide. Lentivirus expressing either AG-biased eGFP or AA-biased eGFP was added to both control and CKO cultures at a concentration of 500 viral particles per neuron.

**Imaging and corrected total cell fluorescence measurements.** Before conducting the codon-biased eGFP experiments, IHC of cultured neurons and glia using the primary antibodies anti-Tuj1 (neurons) and anti-S-100 (Schwann cells) was used to verify that cell morphology could be definitively used to distinguish between neurons and glia. For antibody details, see below. In these preliminary experiments, 98.9% of cells were identified correctly.

After lentiviral infection and a 48 h incubation at 37 °C (5% CO2), a random number generator was used to randomly select 30 live eGFP positive neurons from 50 live neurons; neurons versus Schwann cells were identified based on morphology (see above). The 50 randomly selected neurons were imaged using a Leica TCS SP8 Confocal Laser Scanning microscope with identical laser intensity, gain, and offset.
settings used for all cells. Cell size was measured using Image J Image Processing and Analysis Software (https://imagej.nih.gov/ij/). The average cell areas for control neurons expressing AG- and AA-biased eGFP were 893.0 (range = 359–1,598) and 784.9 (range = 345–1,429) square pixels, respectively. Average cell areas for CKO neurons expressing AG- and AA-biased eGFP were 575 (range = 368–858) and 598 (range = 481–856) square pixels, respectively. Image J was then used to measure the corrected total cell fluorescent CLCF of all cells with an area between 400–800 square pixels (~30 cells per genotype and construct). Data are presented as the average CTCF ± SEM.

Proteomics. DRG from seven control (Ikkbkap+/LoxP) and seven CKO (Wnt1-Cre; Ikkbkapfl/+LoxP) embryos were collected on ice and snap frozen in liquid nitrogen. Global proteomic profiling was performed by Proximity (Proximity.com) as follows:

Samples were prepared for digestion using the filter-aided sample preparation method (https://doi.org/10.1038/nmeth.1322). Brieﬂy, the samples were suspended in RIPA buffer (pH 7.6, 3 mM DTT). Total RNA was used to generate biotin-labeled cRNA with the Illumina TotalPrep RNA Amplification Kit. In short, 0.5 μg of total RNA was first converted into single-stranded complementary DNA with reverse transcriptase using an oligo-dT primer containing the T7 RNA polymerase promoter site and then copied to produce double-stranded cDNA molecules. The double-stranded cDNA was cleaned and concentrated with the supplied columns and used in an overnight in-vitro transcription reaction where single-stranded RNA (cRNA) was generated incorporating biotin-16-UTP. A total of 0.75 μg of biotin-labeled cRNA from each of four control and three CKO biological replicates was denatured at 58 °C for 10 min and hybridized to an Illumina Sentrix mouse58 Expression BeadChip (Illumina, San Diego, CA). Each BeadChip has ~24,000 well-annotated RefSeq transcripts with ~30-fold redundancy. The arrays were washed, blocked, and labeled cRNA was detected by staining with streptavidin-Cy3. Hybridized arrays were scanned using an Illumina BeadStation 500 × Genetic Analysis Systems scanner and the image data extracted using Illumina’s GenomeStudio software, version 1.9.0. For statistical analysis, the expression data were filtered to include only probes with a consistent signal on each chip and an Illumina detection p-value < 0.02.

Gene Ontology analyses. The plots for Fig. 5 and Supplementary Fig. 2 were created using the GOplot R package (https://CRAN.R-project.org/package=GOplot). Significance was calculated using the DAVID 6.8 Functional Annotation Tool with the GO biological process direct database (https://david. ncifcrf.gov/).[38,39]

Real-time quantitative PCR. Real-time quantitative PCR (qPCR) was performed using the Viaa-7 qPCR System by Applied Biosystems and TaqMan Gene Expression Assays from Thermo Fisher Scientific (Mm01218747 [Bra2]), Mm01283063 (Vcam), Mm00185982, Mm01796543 (Hist2H2aa), and Mm00779772 (Hist2H2b). Relative expression levels were calculated and normalized against Gapdh (TaqMan GeneExpression Assay: Mm99999915). DRG from three E17.5 control (Ikkbkap+/LoxP) and three CKO (Wnt1-Cre; Ikkbkapfl/+LoxP) embryos. RNA was extracted using the Qiagen RNeasy Mini Kit and cDNA synthesized via the High Capacity RNA-to-cDNA Kit (Applied Biosystems). Each reaction was performed in triplicate. Bars represent SD calculated from three Delta Delta Ct Expression values for each genotype.

Comet assays. DRG were collected from three control and three CKO embryos at E12.5 and dissected as above. Approximately 5,000 cells were plated per each of three wells per embryo. Comet assays were performed according to the manufacturer’s instructions (Cell Biolabs, OxiSelect Comet Assay Kit, catalog number STA-350). For analysis, each well was imaged and ~200 randomly selected comets were analyzed per embryo using the ImageJ OpenPlots plugin.

Immunochemistry. All washing, blocking, secondary antibody, and post-fixation steps were performed at room temperature. All other steps were performed at 4 °C unless stated otherwise. Embryos were fixed in 4% paraformaldehyde/ phosphate-buffered saline (PBS) for 2 h, rinsed in PBS, cryoprotected through a series of sucrose solutions in PBS (15%, 30%), incubated for 2 h in a 1:1 mixture of 30% sucrose and optimal cutting temperature (OCT) compound (Tissue-Tek, Torrance, CA), followed by 2 h in OCT. Embryos were then frozen and cryosectioned at 16 μm. Slides were bathed in a 1:1 mixture of saline for 10 min, followed by NGS block (10% normal goat serum, 1% glycine, 0.4% Triton X-100 in 30 mM Tris, 150 mM NaCl) for 1 h, and overnight incubation in primary antibody (in NGS block). Slides were then rinsed in NGS block, incubated in Alexa Fluor secondary antibody (1:2000 in NGS block) for 1 h, rinsed in the NGS block, and mounted in ProLong Antifade (Invitrogen, La Jolla, CA). Control and experimental embryos were cryosectioned on the same day and sections were typically incubated in primary antibody on the same day that they were sectioned.

Microscopy images were captured using a Nikon TE200 inverted microscope and a QImaging QICAM 12 bit Mono Fast 1394 Cooled camera and QCapture software. Identical exposure times, gain, and offset settings were used to capture control and experimental images.

Primary antibodies included the following: BRCA2 (Bioirb, orb10203, 1:25), VCAN (Abcam, ab177480, 1:100), S-100 (Agilent Technologies, Z0311, 1:400), TuJ1 (Bielopharm, 801292, 1:1000), and Histone H2A (Bioirb, orb127582, 1:200).

Secondary antibodies used were Alexa Fluor goat anti-rabbit 488 and goat anti-mouse 568, (Invitrogen, 1:2,000).
Data analysis and statistics. For fetal crossectioning, two alternating series of slides with approximately twenty 10 μm sections per slide were generated from the mid-lumbar axial level (10 total slides). Different antibodies were used for each series of five slides, which typically contained sections from three to four DRG. For each series, the center-most section of each of 6 DRG (3 from the left half of the body and 3 from the right half of the body) was determined, and the 2 sections flanking that center section were imaged for a total of 12 sections per embryo. Identical gain and offset parameters were used for control and experimental slides and a minimum of three control and three CKO embryos were analyzed per antibody. CTCF was measured using Image J and the 10 brightest cells per section were averaged. Data are presented as the average CTCF ± SD. Statistical significance was determined by an unpaired Student’s t-test.

Mice. The generation of Wnt1-Cre; Ikbkapflox/flox (Ikbkap CKO) embryos has been previously described22. Wnt1-Cre mice were purchased from the Jackson Laboratory (stock number 030829) and all strains were maintained on a C57Bl/6j background. For all analyses, Wnt1-Cre; Ikbkapflox/flox E17.5 embryos were used as experimental and Ikbkapflox/flox littersmates were used as controls. All genotyping was performed via routine PCR22. Ikbkap CKO and wild-type alleles were distinguished using the following primers: forward (F), 5′-GCCAATCTATCTGTGACGGC-3′ and reverse (R), 5′-AGTGGGCGAGGAGACCA-3′. The presence of the Wnt1-Cre allele was detected using the following primers: F, 5′-GCAAATCTACCTCCTTACGAGC-3′ and R, 5′-CTCTATCGACCAAGCGTGG-3′. All performed according to the National Institutes of Health Guide for Care and Use of Laboratory Animals and protocols approved by the Montana State University Institutional Animal Care and Use Committee.

Data accessibility. The transcriptome data generated and analyzed in this study are available at Gene Expression Omnibus, accession number GSE80130 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE80130). Proteome data are available in the ProteomeXchange repository, accession number PXD007869 (http://massive.ucsd.edu/ProteoSAFe?queryPXDId=PXD007869).

Received: 20 March 2017 Accepted: 29 January 2018
Published online: 01 March 2018

References

1. Zhou, M. et al. Non-optimal codon usage affects expression, structure and function of clock protein FRO. Nature 495, 111–115, https://doi.org/10.1038/nature11833 (2013).

2. Yu, C. H. et al.Codon usage in yeast glutamic acid tRNA. Mol. Biol. Cell 29, 744–754 (2015).

3. Nikolskaya, D. D. & Leidel, S. A. Optimization of codon translation rates via tRNA modification maintains protein integrity. Cell 161, 1606–1618, https://doi.org/10.1016/j.cell.2015.05.022 (2015).

4. Murphy, F. Vt, Ramakrishnan, V., Agris, P. F. The role of codon discrimination by tRNA(Lys)UUU by elongator is essential for efficient translation of tRNA(Lys)UUU. PLoS Genet. 9, e1003647, https://doi.org/10.1371/journal.pgen.1003647 (2013).

5. George, L. et al. Familial dysautonomia model reveals Ikbkap deletion causes apoptosis of Pan3+ progenitors and peripheral neurons. Proc. Natl Acad. Sci. USA 110, 18698–18703, https://doi.org/10.1073/pnas.1308596110 (2013).

6. Slaughenhaupt, S. A. et al. Tissue-specific expression of a splicing mutation in the IKBKAP gene causes dysautonomia. Am. J. Hum. Genet. 88, 598–605 (2001). 30002-9297(07)63100-0 [pii].

7. Cuaudonc, M. P. et al. Tissue-specific reduction in splicing efficiency of IKBKAP due to the major mutation associated with familial dysautonomia. Am. J. Hum. Genet. 72, 749–758 (2003).

8. Pearson, J., Pytel, B. A., Grover-Johnson, N., Axelrod, F. & Dancis, J. Quantitative studies of dorsal root ganglia and neuropathologic observations on spinal cords in familial dysautonomia. J. Neurol. Sci. 35, 77–92 (1978).

9. Axelrod, F. B. Familial dysautonomia. Muscle Nerve. 29, 352–363 (2004).

10. Zinshteyn, B. & Gilbert, W. V. Loss of a conserved tRNA anticodon nucleoside mcmsU in tRNA. EMBO J. 24, 3637–3644 (2005).

11. Huang, B., Johansson, M. J. & Bystrom, A. S. A deep early step in wobble uridine tRNA modification requires the Elongator complex. RNA 11, 424–436 (2005). 11/4/424 [pii]10.1261/rna.7247705.0.
41. Strug, L. J. et al. Centrotemporal sharp wave EEG trait in rolandic epilepsy maps to Elongator Protein Complex 4 (ELP4). Eur. J. Hum. Genet.: ELHG 17, 1171–1181, https://doi.org/10.1038/ejhg.2008.267 (2009).
42. Simpson, C. L. et al. Variants of the elongator protein 3 (ELP3) gene are associated with motor neuron degeneration. Hum. Mol. Genet. 18, 472–481, doi:10.1093/hmg/ddn375 [pii] (2009).
43. Carr, V. M. & Simpson, S. B. Jr. Proliferative and degenerative events in the early development of chick dorsal root ganglia. II. Response to altered peripheralfields. J. Comp. Neurol. 182, 741–755, https://doi.org/10.1002/cne.901820411 (1978).
44. Rifkin, J. T., Todd, V. J., Anderson, L. W. & Lefcort, F. Dynamic expression of neurotrophin receptors during sensory neuron genesis and differentiation. Dev. Biol. 227, 465–480, https://doi.org/10.1016/j.dib.2000.08.1012-1606 (2000).
45. Ma, Q., Fode, C., Guillomet, F. & Anderson, D. J. Neurogenin1 and neurogenin2 control two distinct waves of neurogenesis in developing dorsal root ganglia. Genes Dev. 13, 1717–1728 (1999).
46. Orii, K. E., Lee, Y., Kondo, N. & McKinnon, P. J. Selective utilization of nonhomologous end-joining and homologous recombination DNA repair pathways during nervous system development. Proc. Natl Acad. Sci. USA 103, 10017–10022, https://doi.org/10.1073/pnas.0602436103 (2006).
47. Wu, Y. et al. Versican V1 isoform induces neuronal differentiation and promotes neurite outgrowth. Mol. Biol. Cell. 15, 2093–2104, https://doi.org/10.1091/mbc.E04-04-0295 (2005).
48. Sheng, W. et al. The roles of versican V1 and V2 isoforms in cell proliferation and apoptosis. Mol. Biol. Cell. 16, 1330–1340, https://doi.org/10.1091/mbc. E04-04-0295 (2005).
49. Mendoza-Santiesteban, C. E. et al. Pathological conformation of optic neuropathy in familial dysautonomia. J. Neuropathol. Exp. Neurol. 76, 238–244, https://doi.org/10.1093/jnen/ntv118 (2017).
50. Macefield, V. G. et al. Increasing cutaneous afferent feedback improves proprioceptive accuracy at the knee in patients with sensory ataxia. J. Neurophysiol. 116, 711–716, https://doi.org/10.1152/jn.00148.2015 (2016).
51. Chaverra, M. et al. The Familial Dysautonomia disease gene, Ikbkap/Elp1, is required in the developing and adult central nervous system. Dev. Biol. 370, 605–618, https://doi.org/10.1016/j.ydbio.2012.08.028 (2017).
52. Rahal, E. A. et al. ATM regulates Mre11-dependent DNA end-degradation and microhomology-mediated end joining. Cell Cycle 9, 2866–2877, https://doi.org/10.4161/cc.9.14.12363 (2010).
53. Gold-von Simson, G., Romanos-Sirakis, E., Maayan, C. & Axelrod, F. B. Neoplasia in familial dysautonomia: a 20-year review in a young patient population. J. Pediatr. 155, 934–936, https://doi.org/10.1016/j.jpeds.2009.04.055 (2009).
54. Wang, Y. et al. Prognostic significance of elongator protein 3 expression in endometrioid adenocarcinoma. Oncol. Lett. 3, 25–29, https://doi.org/10.3892/ol.2011.428 (2012).
55. Wang, D. et al. Promoter hypermethylation may be an important mechanism of the transcriptional inactivation of ARRD5C, GATA5, and ELP3 in invasive ductal breast carcinoma. Mol. Cell. Biochem. 396, 67–77, https://doi.org/10.1007/s11010-014-2143-y (2014).
56. Roessler, S. et al. Integrative genomic identification of genes on 8p associated with hepatocellular carcinoma progression and patient survival. Gastroenterology 142, 957–966, https://doi.org/10.1053/j.gastro.2011.12.039 (2012).
57. Walter, W., Sanchez-Cabo, F. & Ricote, M. GOplot: an R package for visually combining expression data with functional analysis. Bioinformatics 31, 2912–2914, https://doi.org/10.1093/bioinformatics/btv300 (2015).
58. Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4, 44–57, https://doi.org/10.1038/nprot.2008.211 (2009).
59. Huang da, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 37, 1–13, https://doi.org/10.1093/nar/gkr923 (2009).