The predicted RNA-binding protein regulome of axonal mRNAs

Raphaëlle Luisier[1,2*], Catia Andreassi[3*], Lisa Fournier[1,2] and Antonella Riccio[3]

[1]Idiap Research Institute, Martigny 1920, Switzerland.

[2]SIB Swiss Institute of Bioinformatics, Lausanne 1015, Switzerland.

[3]UCL Laboratory for Molecular Cell Biology, University College London, London WC1E 6BT, UK.

Corresponding authors: a.riccio@ucl.ac.uk, raphaelle.luisier@idiap.ch

*These authors contributed equally to the work

Running title: RBPs binding to 3' UTR isoforms in axons

ABSTRACT

Neurons are morphologically complex cells that rely on the compartmentalization of protein expression to develop and maintain their cytoarchitecture. Targeting of RNA transcripts to axons is one of the mechanisms that allows rapid local translation of proteins in response to extracellular signals. 3’ untranslated regions (UTRs) of mRNA are non-coding sequences that play a critical role in determining transcript localisation and translation by interacting with specific RNA-binding proteins (RBPs). However, how 3’ UTRs contribute to mRNA metabolism and the nature of RBP complexes responsible for these functions remain elusive.

We performed 3’ end sequencing of RNA isolated from cell bodies and axons of sympathetic neurons exposed to either Nerve Growth factor (NGF) or Neurotrophin 3 (NT-3). NGF and NT-3 are growth factors essential for sympathetic neuron development through distinct signalling mechanisms. Whereas NT-3 acts mostly locally, NGF signal is retrogradely transported from axons to cell bodies. We discovered that both NGF and NT-3 affect transcription and alternative polyadenylation in the nucleus and induce the localisation of

specific 3' UTR isoforms to axons, including short 3' UTR isoforms found exclusively in axons. The integration of our data with CLIP sequencing data supports a model whereby long 3' UTR isoforms associate with RBP complexes in the nucleus, and upon reaching the axons, are remodelled locally into shorter isoforms. Our findings shed new light into the complex relationship between nuclear polyadenylation, mRNA localisation and local 3' UTR remodelling in developing neurons.

INTRODUCTION

Axonal RNA transport and local translation is a widespread phenomenon observed in most neuronal cell types and across species (Holt et al. 2019). In developing axons, local translation plays a crucial role in mediating cell survival and axon growth in response to extracellular cues. It enables axon extension, guides growth cone turning, and promotes nerve regeneration after injury (Dalla Costa et al. 2020).

During neuronal development, axons extend over long distances to reach their targets, driven by their growth cone, which responds rapidly to local cues along its migratory path (Dorskind and Kolodkin 2021). At early developmental stages, sympathetic neurons respond to neurotrophin 3 (NT-3) released from the vasculature surrounding the axons (Makita et al. 2008), eliciting a local signal that supports initial axon growth and cell survival (Elshamy and Ernfors 1996; Kuruvilla et al. 2004). At later stages, when axons have reached a considerable length and approach their final targets, Nerve Growth Factor (NGF) is released, internalised within signalling endosomes and transported back to the cell bodies where it activates transcription (Kuruvilla et al. 2004; Ascano et al. 2012; Scott-Solomon and Kuruvilla 2018; Scott-Solomon et al. 2021). In sympathetic neurons both neurotrophins bind to the same tyrosine kinase receptor TrkA, although their ability to signal to the nucleus depends on the internalization into signalling endosomes and transport to cell bodies. In contrast to NGF, NT-3/TrkA complexes are not found in signalling endosomes and are not retrogradely transported.

Thus, NT-3 in axons is thought to act mostly locally and to be a poor activator of gene expression (Kuruvilla et al. 2004).

In addition to conveying the genetic information from the chromatin to the translational machinery, mRNA transcripts also carry information stored in their untranslated regions (UTRs). The 3' UTRs regulate many aspects of RNA metabolism, including transcript localization, mRNA stability and translation by interacting with RNA-binding proteins (RBPs) (Andreassi et al. 2018; Andreassi and Riccio 2009; Mayr 2017; Andreassi et al. 2021). RBP complexes are initially assembled co-transcriptionally and are essential for regulating RNA splicing, polyadenylation site (PAS) choice, alternative polyadenylation (APA), and mRNA nuclear export (Van Nostrand et al. 2020; Hentze et al. 2018). Most RBPs are multifunctional, and the complexes undergo extensive remodelling in the cytoplasm (Hentze et al. 2018). RBP interaction with elements in the 3' UTR mediate mRNA targeting to dendrites and axons, and this event is necessary for the establishment of synapses, axon growth and nerve repair after injury (Cosker et al. 2016; Allen et al. 2013; Doyle and Kiebler 2011; Thelen and Kye 2019; Dalla Costa et al. 2020; Terenzio et al. 2017; Holt and Schuman 2013).

Here, we performed a comprehensive analysis of transcription, APA, mRNA transport and putative RBPs binding in cell bodies and axons of developing rat sympathetic neurons in response to neurotrophins. We aim to further characterise the shift in the transcriptional landscape that takes place in developing sympathetic neuron axons in response to extracellular cues.

RESULTS

NGF and NT-3 applied to distal axons induce distinct transcriptional programmes

To identify mRNAs localised in cell bodies and axons of sympathetic neurons in response to neurotrophins, we integrated previously published 3' UTR sequencing data obtained from compartmentalized cultures of sympathetic neurons exposed to NGF (Andreassi et al. 2021), with data simultaneously obtained from axons exposed to NT-3. Compartmentalized chambers allow the physical separation of cell bodies from distal axons and are especially suited for

sympathetic neurons because they grow in culture as a highly homogeneous population without glial cells. Neurons were seeded in the central compartment with NGF (100 ng/ml) and after 5 days, either NGF or NT-3 were added to the lateral compartments. A low concentration of NGF (10ng/mL) was maintained in the cell bodies to preserve viability. After 7 additional days necessary to achieve robust axon growth, RNA was isolated from cell bodies or axons, subject to two rounds of linear amplification and sequenced using 3'end-Seq. This technique allowed the sequencing of transcripts 3'ends independently of the length of the transcript (see methods for a detailed description of the protocol used; **Fig. 1A**) (Andreassi et al. 2021).

Analyses of the reads obtained from two independent biological replicates indicated the high reliability of the RNA sequencing in both compartments (**Supplemental Fig. S1A**). The lower correlation between the replicates from axonal compartments is likely due to the very low amount of RNA isolated in axons. Unsupervised analysis of the sequencing reads indicated the dominant effect of the compartments over the treatment condition. Indeed, despite axon only exposure to different neurotrophins, we primarily observed the effects of the treatments in the cell bodies (**Supplemental Fig. S1B,C**). Thus, we first aimed to examine the transcriptional changes in the cell bodies upon exposure of distal axons to NGF or NT-3. Because a control condition using neurons grown without neurotrophins is not possible, the experimental setting does not distinguish between changes due to the direct effect of each neurotrophin or the indirect effect due to the lack of thereof. Differential gene expression analysis of neurons whose axons were treated with NGF or NT-3 identified 232 genes up-regulated in NGF treated neurons compared to NT-3 (FC>1.5 and P value<0.01, **Fig. 1B**, **Supplemental Table S1**) enriched in terms related to cell adhesion and glucose metabolism (**Fig. 1C**, *upper*). Conversely, the expression of 119 genes was increased in the cell bodies of neurons whose axons were exposed to NT-3, and they were enriched for neuron projection and synaptic transmission terms (**Fig. 1C**, *lower*, **Supplemental Table S2**). Transcription factor binding site (TFBS) enrichment analysis indicated a neurotrophin-specific transcriptional regulation of the differentially expressed genes (**Fig. 1D**, **Supplemental Table S3**). The 1,000 nucleotide promoter regions of genes up-regulated in NGF compared to NT-3 were enriched in

14 TFBS motifs, including motifs bound by Early growth response 1 (EGR1), Early growth response 1 (EGR2), Early growth response 3 (EGR3; P-Value=3.46 x 10^{-04}, 43% of the promoter regions of NGF up-regulated genes), Fos proto-oncogene (FOS), FosB proto-oncogene (FOSB), FOS like 1 (FOSL1), JunB proto-oncogene (JUNB) and JunD proto-oncogene (JUND) (P-Value=3.91 x 10^{-03}, 16% the promoter regions of NGF up-regulated genes; **Fig. 1E**), all belonging to the Reactome biological pathway of NGF-stimulated transcription (HSA-9031628). Fewer motifs were enriched within the promoter regions of transcripts up-regulated in neurons whose axons were exposed to NT-3, and they all related to embryonic development. Such motifs were bound by Transcription factor SOX-2, POU domain, class 5, transcription factor 1 (POU5F1, 2.26 x 10^{-02}, 8% of the promoter regions of NT-3 up-regulated genes) and Transcription factor SOX-17 (P-Value=2.39 x 10^{-02}, 23% of the promoter regions of NT-3 up-regulated genes; **Fig. 1F**).

Alternative Polyadenylation (APA) takes place principally co-transcriptionally, generating transcripts expressing identical coding regions and 3' UTRs of different length (Tian and Manley 2013). To ask whether neurotrophins induce distinct APA, we investigated the shifts of the relative 3' UTR usage of transcripts expressed in cell bodies of neurons whose axons were exposed to NGF or NT-3. 27 transcripts showed distal-to-proximal promoter 3' UTR shift in NGF compared to NT-3 (**Fig. 1G, Supplemental Fig. S1D** and **Supplemental Table S4**), and 45 transcripts exhibited distal-to-proximal promoter 3' UTR shifts in response to NT-3 (**Fig. 1H**, **Supplemental Fig. S1D** and **Supplemental Table S5**). Differential relative 3' UTR isoform usage between the two conditions can be due to selective increase of promoter-proximal 3' UTR isoforms, promoter-distal isoforms or both. The 27 transcripts exhibiting significant promoter distal-to-proximal 3' UTR shifts in NGF included 20 promoter-proximal shifts in NGF and 17 promoter-distal shifts in NT-3 (**Fig. 1G**). The 45 transcripts exhibiting significant promoter distal-to-proximal 3' UTR shifts in NT-3 included 34 promoter-proximal shifts in NT-3 and 26 promoter-distal shifts in NGF (**Fig. 1H**). In both conditions some transcripts exhibited changes in both directions. Thus while neurons exposed to NGF exhibited a similar number of promoter-proximal and promoter-distal 3' UTR shifts (20/26), NT-3 induced twice as many

promoter-proximal shifts as promoter-distal shifts (34/17), indicating a preferential usage of isoforms with short 3' UTR in NT-3-treated condition. Depending on the neurotrophin, 3' UTR shifts targeted genes associated with distinct biological pathways: whereas target genes related to cellular stress and DNA damage were overrepresented in response to NGF (**Fig. 1I**, *upper panel*), target genes in neurons treated with NT-3 were enriched in cell migration and differentiation biological pathways (**Fig. 1I**, *lower panel*). When comparing changes in gene expression and APA between the two conditions, we found that the genes affected by APA were distinct from those undergoing transcriptional changes under NT-3 and NGF conditions (**Supplemental Fig. S1E**). We also observed a lack of correlation between changes in 3' UTR usage and overall gene abundance (**Supplemental Fig. 1F)**. These results are in line with previous studies (Tian and Manley 2017) showing the role of APA in global mRNA metabolism independent of mRNA expression. Thus, both NT-3 and NGF propagate signals from distal axons to the nucleus to regulate gene expression and co-transcriptional APA.

Identification of RBPs that regulate APA in response to NGF and NT-3

RBPs regulate mRNA processing and metabolism, including APA (Hentze et al. 2018; Erson-Bensan 2016). To identify putative regulators of neurotrophin-specific 3' UTR APA, we interrogated publicly available cross-linking and immunoprecipitation (CLIP) sequencing data for 126 RBPs assayed in human cell lines (**Supplemental Table S6**). Visual inspection of the distribution of RBP cross-link events along the 3' UTR isoforms revealed that the 500 nucleotide (nt) region preceding the 3' ends exhibited the highest percentage of bound RBPs (70%) (**Supplemental Fig. S2A**). This region is expected to serve a regulatory function given its high conservation score (**Supplemental Fig. S2B**). We first searched for associations between cross-linking events mapping to defined regions along the promoter-proximal and promoter-distal 3' UTRs, and the relative usage in cell bodies of the short 3' UTR isoform. 126 RBPs were tested individually both in NGF and NT-3 conditions (see Methods for details). A positive association between RBP binding and the expression of 3' UTR isoforms predominantly occurred at the 3' end of both the promoter-proximal (Ip) and the promoter-distal

3' UTR (Id), irrespective of the neurotrophin used (**Fig. 2A**). We identified 17 RBPs preferentially bound within the [-350:+150] region surrounding the 3' end, acting as positive regulators of polyadenylation (**Supplemental Fig. S2C,E** and **Supplemental Table S7**). These included the cleavage and polyadenylation (CPA) factors 1 (CPSF1, also known as CPSF160; **Fig. 2B**) and Cleavage stimulation factor (CSTF) subunit 2 (CSTF2; **Fig. 2C**), that are known to reside in the [-50:0] and [0:50] nt regions around the 3' end (Mitschka and Mayr 2022). Thus, the binding of RBPs to specific 3' end terminal regions was positively associated to both short and long 3' UTR isoforms, which is consistent with the fact that these factors may promote 3' end processing irrespective of 3' UTR length (Laishram and Anderson 2010).

Negative association between 3' UTR isoform usage and RBP binding revealed that negative regulators of polyadenylation are uniquely detected in the [0:150] nt region downstream of the promoter-proximal 3' end (**Fig. 2D**). Long 3' UTR isoforms were not associated with negative regulators. The 27 negative regulators of the short 3' UTR form a densely connected network of experimentally validated interacting proteins that are enriched in biological processes related to mRNA stability (**Fig. 2E,F, Supplemental Fig. S2D,F** and **Supplemental Table S8**). ELAV-like protein 1 (ELAVL1, also known as HuR) is a well characterised RBP that competes with the CSTF factors downstream of poly(A) sites to block polyadenylation (Zhu et al. 2007). ELAVL1 binding downstream of the 3' end of the promoter-proximal 3' UTRs was significantly associated with decreased promoter-proximal 3' UTR usage, promoting the transcription of long 3' UTR isoforms. In contrast, ELAVL1 binding upstream of the 3' end was not predicted to affect polyadenylation (**Fig. 2G**). Therefore, our analysis supports a model by which the selection between short and long 3' UTR isoform primarily depends on negative factors bound within the 150 nt region downstream of the promoter-proximal 3' end. Analysis of the potential to localise into cellular condensates revealed significant differences between the 17 positive and the 27 negative regulators of polyadenylation, with the former exhibiting higher probability to localise into P-bodies and cytoplasmic granules (**Fig. 2H**). Our analysis identifies new candidates and known regulators of polyadenylation along with their preferential location around the 3' end, with positive

regulators more likely to concentrate into condensates where they may operate to promote the cleavage of the 3' end (Andreassi et al. 2021).

Next, we tested whether the subtle but significant differences in APA observed in dozens of transcripts upon exposure of distal axons to NGF or NT-3 (**Fig. 1G**) were associated with specific RBPs. 3' UTR position-dependent Fisher enrichment in RBP cross-link events were studied for the four groups of APA isoforms previously identified (**Fig. 1G,H**). The distributions of the enrichment scores ($-\log_{10}$(P-value)) along the 3' UTR revealed that the regulatory regions with significant enrichment of RBPs were restricted to the region downstream the promoter-proximal 3' end, irrespective of the grouping (**Fig. 3A-D** and **Supplemental Fig. S3**), which aligns with the negative regulators of polyadenylation identified through the Welch's *t*-test (**Fig. 2D** and **Supplemental Fig. S2F**). Using this approach, four groups of RBPs that serve as candidate regulators of preferential promoter-proximal or promoter-distal usage between the two conditions were identified (**Fig. 3E-H**). While this analysis does not distinguish between positive and negative regulators of differential APA between each conditions, several of these factors, including heterogeneous nuclear ribonucleoprotein C (HNRNPC), Y-box-binding protein 3 (YBX3), and KH-Type Splicing Regulatory Protein (KHSRP), were previously identified as global negative regulators of APA (**Fig. 2E**). Strikingly, these RBPs bind within the 150 nt region downstream of the promoter-proximal 3' end (**Fig. 3A-D**), similar to the predicted negative regulators of APA (**Fig. 2D**). These findings collectively suggest that each neurotrophic treatment leads to a condition-specific change in the *negative* regulatory activity of these factors (**Fig. 3E-H**). Specifically, the *negative* regulatory activity on the promoter-proximal 3' UTR isoform is diminished in one condition compared to the other when there is a shift towards the usage of short 3' UTR isoforms in the former (**Fig. 3E,F,I,J**). Conversely, it is enhanced when there is a preference for long 3' UTR isoforms in the former (**Fig. 3G,H,K,L**). For example, HNRNPC is enriched in the regions downstream of the 3' end of the promoter-proximal 3' UTR isoforms whose related distal isoforms are up-regulated in the NGF condition compared to the NT-3 condition (**Fig. 3D,H**). Furthermore, HNRNPC is enriched in the regions downstream of the 3' end of the promoter-proximal 3' UTR isoforms that are up-regulated in

the NT-3 condition compared to NGF(**Fig. 3A,E**). These results support a stronger repressor activity of HNRNPC in response to NGF compared to NT-3. Tracks of representative transcripts expressing long or short 3' UTR in cell bodies of neurons treated with NGF or NT-3 along with CLIP-seq data are shown in **Fig. 3I-L.**

NGF and NT-3 induce the localisation of distinct mRNA transcripts in axons

We next investigated whether NGF and NT-3 prompted the transport of different mRNAs in axons. A similar number of transcripts and 3' UTR isoforms were found in cell bodies and axons independently of neurotrophin treatment (**Fig. 4A** and **Supplemental Fig. S4A**). However, a large number of axonal mRNAs with distinct cellular functions (**Supplemental Fig. S4B**) was uniquely detected in response to either NGF (n=1,962) or NT-3 (n=1,089; **Fig. 4B**). In line with previous studies (Tushev et al. 2018; Andreassi et al. 2021), a larger percentage of axonal transcripts expressed multiple (**Fig. 4C**) and longer (**Fig. 4D**) 3' UTR isoforms, when compared to cell bodies in both conditions.

The abundance of axonal mRNAs mostly correlated with both expression levels in the cell bodies and transcript length (**Supplemental Fig. S4C,D**), suggesting that diffusion and anchoring are important mechanisms regulating mRNA localisation (St Johnston 2005). To study the mechanisms responsible for active mRNA transport, we developed a statistical model and a novel metric that we named Localisation Score (LS). LS quantifies the efficiency of transcripts localisation in axons irrespective of mRNA length and abundance in the cell bodies (**Supplemental Table S9**; see Methods and Supplemental Material). A positive LS indicates higher axonal mRNA abundance than expected for transcripts with similar size and expression levels, and correlates with active transport and stabilisation. Conversely, negative LS values are indicative of either restricted transcript diffusion from the cell bodies or higher mRNA degradation. Because LS enables the identification of over- and under-transported mRNA irrespective of the expression levels (**Supplemental Fig. S4E-G**), it provides a better statistical tool than the ratio of the gene abundance between axons and cell bodies commonly used (**Supplemental Fig. S4H**), as the latter is more prone to identify extreme values for

highly expressed transcripts given the larger dynamic ranges in expression of this class of transcripts. Analysis of GO enrichment per range of abundance ratios *versus* LS shows that terms associated with over-transported transcripts using LS scores better reflect the biological system (axon development, cell adhesion) as compared to transcripts exhibiting excess in abundance ratios (**Supplemental Results** and **Supplemental Fig. S7**).

Using LS, thousands of over- and under-transported 3' UTR isoforms were identified in both NGF and NT-3 conditions (**Supplemental Fig. S4I,J**). To validate the statistical model, real-time quantitative PCR (RT-qPCR) was performed on transcripts predicted to be restricted to cell bodies. Analysis of *Eid2* and *Rab22a* indicated that these mRNAs were virtually absent in NGF and NT-3 treated axons, despite being highly expressed in the cell bodies (**Supplemental Fig. S4K**). LS analysis revealed that NGF and NT-3 promote axonal localisation of transcripts associated with largely similar GO biological pathways, such as vesicular localisation and axo-dendritic transport (**Supplemental Fig. S4L**). Subtle but significant differences in axonal transcripts associated with specific biological pathways were detected in both conditions. Transcripts related to the collagen catabolic pathway exhibited higher LS in NGF axons compared to NT-3 (**Fig. 4E,F**), while those related to the vascular endothelial-derived growth factors-related pathway had higher LS in NT-3 (**Fig. 4G,H**). The latter finding is especially interesting considering that when sympathetic neurons are exposed to NT-3, axons grow in close contact with blood vessels (Scott-Solomon et al. 2021), possibly inducing the transport of transcripts that mediate cross-signalling between neurons and endothelial cells. Analysis performed on individual transcripts identified 482 3' UTR isoforms with significantly higher LS in NGF-treated axons and 348 with higher LS in NT-3-treated axons (**Fig. 4I** and **Supplemental Table S10,11**). Single molecule RNA Fluorescence In Situ Hybridization (smFISH) of *Atf3*, a transcript predicted to localise in NT-3 but not NGF-treated axons, confirmed that neurotrophins induce mRNA axonal localisation in a highly specific manner (**Fig. 4J,K**) (Willis et al. 2007). We next investigated the interplay between mRNA regulation in the cell body (transcription and APA) and axonal localization. As differences in localization scores were not correlated with differences in gene expression (**Supplementary**

279 **Fig. S4M**), we found limited overlap between the significantly differentially expressed genes
280 (**Supplemental Tables S1, S2**) and the genes differentially localized between the two
281 conditions (**Supplemental Fig. S4N** and **Supplemental Tables S10, S11**). However,
282 comparison of the LSs of genes targeted by APA in response to NGF and NT-3
283 (**Supplemental Tables S5, S6**) revealed that genes exhibiting significant distal-to-proximal
284 shifts in the NGF condition displayed lower axonal localization compared to NT-3
285 (**Supplemental Fig. 4O**). Although the difference in LS between NGF and NT-3 is not
286 statistically significant (P-value=0.06), these findings suggest a specific correlation between
287 axonal localization and APA, especially in the NGF condition.
288
289 **Computational prediction of the RBP regulome for axonal mRNA localisation**
290 To study whether putative RBP binding accounted for the distinct axonal localisation
291 observed among 3' UTR isoforms in response to neurotrophins we conducted statistical tests
292 comparing the LS of transcripts exhibiting a cross-link event for individual RBPs in the 50-
293 nucleotide regions along the 3' UTR with those that did not exhibit such events (see Methods).
294 A [-200:-50] nt region preceding the 3' end exhibited the highest regulatory potential for axonal
295 localisation in both NGF and NT-3 conditions (median P-value=$10x10^{-25}$ across the 126 RBPs;
296 **Fig. 5A**). The preferential binding of *trans*-acting factors to this specific region is in line with
297 previous findings demonstrating the presence of NGF-dependent localisation elements within -
298 150 nt from the 3' end (Andreassi et al. 2010). Indeed the number of cross-link events (331
299 experiments for 126 RBPs) within the [-150:-100] nt region preceding the 3' end correlated with
300 the LSs **(Fig. 5B)**. The 32 RBPs exhibiting the highest positive association with localised
301 transcripts were enriched in mRNA transport biological pathway (**Fig. 5C** and **Table S12**; P-
302 value=$8.74x10^{-05}$). Regulators of mRNA transport, such as Fragile X messenger
303 ribonucleoprotein 1 (FMR1) (Antar et al. 2004; Dictenberg et al. 2008) (**Fig. 5D**), Insulin-like
304 growth factor 2 mRNA-binding protein 1 (IGF2BP1) (Kislauskis et al. 1994; Ross et al. 1997)
305 (**Supplemental Fig. S5A**) and TAR DNA-binding protein 43 (TDP43) (Nagano et al. 2020;
306 Štalekar et al. 2015) (**Supplemental Fig. S5B**), also exhibited significantly higher cross-linking

events in transcripts over-transported in response to NGF and NT-3 (Fisher count test). Although the 126 studied RBPs showed similar positive association with axonal localisation in both NGF and NT-3 (**Supplemental Fig. S5C**), Eukaryotic translation initiation factor 4 gamma 2 (EIF4G2) (**Supplemental Fig. S5D**) and SNRPB (**Supplemental Fig. S5E**) showed significant enrichment in over-transported 3' UTR isoforms in either NGF or NT-3 conditions, respectively. Interestingly, EIF4G2 is locally translated in rat sympathetic neuron axons where it supports axon growth (Kar et al. 2013). Finally 12 RBPs showed significantly higher binding occurrence (Fisher count test) in the [-250:-50] nucleotide region preceding the 3' end of the under-transported isoforms, when compared to either all transcripts or over-transported isoforms (**Supplemental Table S13**).

Logistic regression was used to investigate the *collective* contribution of 43 RBPs (consisting of 32 positive and 12 negative candidate regulators) that were *individually* associated with mRNA axonal localization. For discrimination between over- and under-transported 3' UTR isoforms (**Supplemental Fig. S5F**), we trained three types of classifiers: (1) model M1 utilized the complete set of 32 positive and 12 negative regulators of axonal localization (refer to **Supplemental Tables S12, S13**), (2) model M2 utilized exclusively the 32 positive regulators, and (3) model M3 utilized exclusively the 11 negative regulators. 42 interaction terms were also generated from a subset of positive and negative regulators to explore their potential synergistic regulatory effects (model M4; see Methods). Using these models, groups of RBPs were identified that enabled the classification of localised *versus* cell body restricted 3' UTR isoforms. RBPs were either positive regulators of axonal localisation (**Supplemental Fig. S5G**) or linked to decreased localisation of specific axonal mRNAs in response either to NGF or NT-3 (**Supplemental Fig. S5H**). While models utilising a higher number of predicted features (M1 and M4) demonstrated superior performance (**Fig. 5E**), the classifiers based on the 11 candidate negative regulators exhibited superior performance in discriminating between over- and under-transported transcripts than the classifiers based on 32 positive regulators in both NGF and NT-3 conditions (i.e. higher model performance of M2 compared to M3; **Fig. 5E**). These 11 negative RBPs form a network of interacting proteins

enriched in regulators of translation and mRNA stability (**Fig. 5F**) including ELAVL1, which reduces neuronal cytoplasmic mRNA, CUGBP Elav-like family member 4 (CELF4) and KHSRP (Engel et al. 2022; Olguin et al. 2022; Patel et al. 2022) (**Supplemental Fig. S5I)**. KHSRP and CELF4 regulate mRNA abundance in axons and dendrites (Snee et al. 2002) while Pumilio RNA binding family member 2 (PUM2) shapes the transcriptome in developing axons by retaining mRNAs in the cell body (Martínez et al. 2019). The visualization of positive and negative regulators shared by NGF and NT-3 along the 3' UTRs of over-transported (**Fig. 5G**) and cell body restricted transcripts (**Fig. 5H**) revealed a pronounced increase in the density of cross-linking events involving negative regulators of axonal localisation. These findings suggest that mRNA destabilization and compartment-specific restriction play crucial roles in regulating the axonal transcriptome (Patel et al. 2022; Wagnon et al. 2012; Shav-Tal and Singer 2005; Loedige et al. 2023).

Further comparison of model performances revealed that classifiers trained with data from the NT-3 condition exhibited superior performance compared to the model trained with data from the NGF condition (**Fig. 5E**), particularly in classifying a lower proportion of cell body-restricted 3' UTR isoforms as axonally localized (resulting in a reduced false positive rate). Thus, transcripts with restricted axonal localization in the NT-3 condition display a more coherent RBP code and a better defined regulation than in NGF.

The enhanced performance of the classifiers trained with 42 interaction terms derived from 14 pairs of RBPs, when compared to those trained with the 43 individual RBPs suggests a combined and dependent regulation among these RBPs (**Fig. 5E**). This finding also highlights the identification of RBP combinations that exhibit greater specificity for each cue (**Fig. 5I,J**). Furthermore, a systematic increase in the density of cross-linking events was observed for pairs of negative regulators in both NGF and NT-3 conditions (**Fig. 5K-N** and **Supplemental Fig. S5J**) highlighting the contribution of combined rather than individual RBPs.

The RBPome for 3' UTR remodelling in axons

We recently discovered that the 3' UTR of transcripts localised in axons may undergo local remodelling, resulting in the generation of polyadenylated isoforms expressing a shorter 3' UTR that are stable and efficiently translated (Andreassi et al. 2021). Analysis of isoforms localised in NGF and NT-3-treated axons revealed a similar number of 3' UTR isoforms with different usage of long and short 3' UTRs in both cell bodies and axons (665 and 458, respectively; **Supplemental Fig. S6A,B; Supplemental Tables S14-17**). Candidate transcripts of axonal remodelling were identified in both NGF and NT-3 conditions as isoforms expressing short 3' UTR in axons but virtually absent in cell bodies (**Fig. 6A**, *left*; **Supplemental Tables S18-19**). This unique expression pattern suggests that they are not the product of co-transcriptional APA and are cleaved in axons. At least some transcripts expressing short 3' UTR only in axons are not transported (Andreassi et al. 2021), however we cannot exclude that their localisation may be due to a very efficient and fast translocation from cell bodies to axons. Isoforms expressing short 3' UTRs only in axons were largely not overlapping between NGF and NT-3 (**Fig. 6A**, *right*). Tracks of representative transcripts expressing unique short 3' UTR in axons treated with either NGF or NT-3 are shown in **Fig. 6B,C**.

We identified Argonaute RISC catalytic component 2 protein (AGO2) as the endonuclease responsible for 3' UTR cleavage (Andreassi et al. 2021). Fisher enrichment in RBP cross-link events along the 3' UTR of the remodelled isoforms (n=80 in NGF and n=60 in NT-3) revealed that while NGF-related predicted regulators of axonal remodelling preferentially localised to the [0:+50] nt region downstream the 3' end, NT-3-related predicted regulators were found on both sides of the 3' end (**Fig. 6D**). Thirty-five enriched RBPs were observed in both NGF and NT-3 remodelled isoforms, whereas 10 and 18 were unique for NGF or NT-3, respectively (**Fig. 6E** and **Supplemental Tables S20-21**). Common RBPs included ELAVL1 which is known to inhibit polyadenylation when bound to the region downstream of the 3' end (**Supplemental Fig. S6C**). The RNA helicase and ATPase UPF1, which belongs to the protein complex that we previously showed to mediate the 3′ UTR cleavage (Andreassi et al. 2021) was enriched in

the 50 nt region downstream of the 3' end of remodelled isoforms in both NGF and NT-3 treated axons (**Fig. 6F** and **Supplemental Fig S6D**). Analysis of the positional preference of the top predicted regulators of axonal cleavage revealed that for NT-3-remodelled isoforms, binding to the [-200:-150] nt window upstream the 3' end is favoured, whereas regulatory regions of NGF-remodelled isoforms reside in the [50:100] nt region downstream the 3' end (**Fig. 6G**). Neurotrophin-specific predicted regulators of axonal remodelling include A-kinase anchor protein 8-like (AKAP8L), which is enriched in the [-200:-150] nt region of 3' UTR short isoforms and is uniquely detected in NT-3 condition, and polypyrimidine tract-binding protein 1 (PTBP1), which is enriched in the [0:50] nt region downstream of the pool of short 3' UTR isoforms detected only in NGF-treated axons (**Fig. 6H**). Analysis of the association of the RBPs predicted to regulate 3' cleavage in axons with APA in the nucleus revealed that these factors behaved as negative regulators of polyadenylation when bound downstream of the 3' end of short 3' UTR isoforms (**Fig. 6I**). Together these findings indicate that RBPs predicted to regulate 3' cleavage in axons may determine nuclear APA and 3' UTR isoform expression.

DISCUSSION

NGF and NT-3 are required for axon growth and neuron survival (Dorskind and Kolodkin 2021). In developing sympathetic neurons, NGF and NT-3 elicit distinct intracellular signalling pathways despite acting through the same receptor Trk-A. NGF-TrkA complexes are internalised into signalling endosomes that travel back long distances to the cell bodies where they activate transcription (Riccio et al. 1997, 1999; Bhattacharyya et al. 1997). In contrast, NT-3 cannot be endocytosed within signalling endosomes and is thought to act mostly locally (Kuruvilla et al. 2004; Scott-Solomon et al. 2021; Ascano et al. 2012; Scott-Solomon and Kuruvilla 2018; Harrington et al. 2011). It should be noted that sympathetic neurons do not express TrkC (Scott-Solomon et al. 2021), the higher affinity receptor for NT-3 (Barbacid 1994), and therefore any NT-3 dependent effect in these cells are likely due to NT-3 binding to TrkA. Our data demonstrate that both neurotrophins when applied to distal axons induce a robust transcriptional response of 3' UTR isoforms that are only partially overlapping (**Fig.1**).

Our findings challenge the current understanding that NT-3 does not signal retrogradely to the cell body compartment and open the possibility that it may initiate a signalling cascade that is faithfully propagated to the nucleus.

RBPs are key regulators of mRNA metabolism including mRNA transport and translation, and are essential in determining when and where specific proteins are expressed. By integrating our 3' end-seq data with CLIP data generated in human cell lines, we discovered novel candidates and known regulators of APA (**Fig. 2**), corroborating the suitability of human data to study RBP regulomes underlying mRNA metabolism in rodent neurons. RNA transport in axons is known to be bidirectional, and transcripts complexed with RBPs can move both anterogradely and retrogradely (Dalla Costa et al. 2020; Terenzio et al. 2017; Holt and Schuman 2013; Das et al. 2019). We identified a large repertoire of transcripts that are localised and stored in sympathetic neuron axons in response to either NGF or NT-3 (**Fig. 3**), and the putative RBPome responsible for mRNA localisation. We also showed that restricted axonal localisation in response to NT-3 is associated with a more defined RBP regulome compared to NGF (**Fig. 5**). Axonal localization is positively regulated by RBPs that facilitate mRNA transport, and negatively modulated by RBPs that regulate mRNA stability. Albeit universal axon-targeting motifs have not been identified so far, 32 RBPs were positively associated with axonal localisation. Employing a logistic regression method, groups of RBPs were identified that influenced mRNA localization to axons in response to either NGF or NT-3. Our analysis strongly suggests that mRNA stability and restriction to cell bodies play a pivotal regulatory role in axonal mRNA localization, aligning with recent research (Loedige et al. 2023). Future studies will clarify whether a similar combination of RBPs drives transcript localisation to other neuronal compartments, such as dendrites and dendritic spines, or in other cell types.

We recently discovered that some axonal transcripts including *Inositol Monophosphatase 1* (*IMPA1*) undergo 3' UTR remodelling in sympathetic neuron axons (Andreassi et al. 2021). Importantly 3' UTR cleavage is necessary for triggering the translation of *IMPA1* mRNA isoforms expressing a short, cleaved 3' UTR (Andreassi et al. 2021). Hundreds of transcripts

with a proximal-to-distal shift of the 3' UTR are detected in axons. Several short 3' UTR isoforms are expressed exclusively in axons and in some cases specifically in response to either NGF and NT-3 (**Fig.5** and **Supplemental Fig. S5**). Analysis of the RBP complexes interacting with the remodelled isoforms revealed an enrichment of UPF1 and PTBP1 previously identified as binding partners of the cleavage complex (Andreassi et al. 2021). It should be noted that predicted regulators of 3' UTR cleavage behaved as negative regulators of nuclear polyadenylation when bound downstream of the 3' end of short 3' UTR isoforms. Thus, we propose that RBPs responsible for 3' UTR cleavage are recruited co-transcriptionally to downstream regions of the promoter-proximal 3' end in the nucleus. They favour the expression of the long 3' UTRs required for axonal localisation (Cosker et al. 2016; Andreassi et al. 2021; Terenzio et al. 2018), possibly by competing with cleavage factors and suppressing the use of the proximal PAS. The factors may remain bound to the 3' UTR within RBP granules and hitchhike along the axons, eventually co-localising in the axonal compartment with the long 3' UTR isoform. Upon de-assembly of the transport granules in axons, some factors, including UPF1, promote the cleavage of the long 3' UTR (**Fig. 7**). Transport granules have been considered as "translation factories" that contain RBPs, mRNAs, ribosomes and translation factors, regulating local protein synthesis (Kanai et al. 2004; Krichevsky and Kosik 2001). Here, we propose a similar mechanism by which APA, RNA localization and neurotrophin-dependent translation are coupled and co-regulated. Release of the 3' UTR isoforms remodelling factors from granules may serve as the final step that allows the translational activation in axons of transcripts expressing shorter 3' UTRs.

While further mechanistic studies will be necessary to validate the integration of compartmentalised 3' end RNA-sequencing and CLIP sequencing data, our study sheds new light on the nature of axonal mRNA and the RBPs that regulate the transport and 3' UTR remodelling. Moreover, given that most neurological diseases are considered as disorders of the RNA (Nussbacher et al. 2019; Wang et al. 2007), our data provide new targets potentially amenable for the cure of degenerative disorders of the nervous system.

METHODS

Reagents

Cell culture reagents, molecular biology reagents and kits were purchased from Thermo Fisher Scientific and all other chemicals from Sigma, unless stated otherwise.

Compartmentalized cultures of rat sympathetic neurons

All animal studies were approved by the Institutional Animal Care and Use Committees at University College London. Superior cervical ganglia were dissected from postnatal day 1-2 Sprague Dawley rats, enzymatically dissociated and plated on glass coverslips or in compartmentalized chambers pre-coated with home-made collagen and laminin (5µg/mL), as previously described (Riccio et al. 1997). Undifferentiated cells seeded in the central compartment of the chambers were initially exposed to 100ng/mL NGF to support survival and cell differentiation. Five to six days after plating, NGF was reduced to 10ng/mL in the central compartment and neurons were maintained with either NGF (100 ng/ml) or NT-3 (1µg/mL) in the lateral compartment only where they promoted extensive axon growth. The concentration of NT-3 used was thoroughly tested to ensure a rate of axonal extension, similar and often even higher than NGF. Medium was changed every 48-72 hrs with a fresh amount of growth factors added. Cytosine arabinoside (ARA-C, 10µM) was added 24 hours after plating to block the proliferation of non-neuronal cells.

RNA isolation, reverse transcription, linear amplification and 3'end-RNA-seq

3' end sequencing based performed as previously (Andreassi et al. 2021) and as decribed in **Supplemental Methods**. Primer sequences and PCR conditions are provided in **Table S23**.

3' UTR isoform quantification and identification of transcripts localized to axons

As previously described (Andreassi et al. 2021), the last 500 nt portion of each transcript contains above 70% of the reads originating from that transcript irrespective of their length. We thus used the number of reads mapped to the -500 nt terminal region of each 3' UTR isoform as a proxy for the 3' UTR isoform expression levels. Because 3' end seq amplifies the 3' end of

the transcript, it is not in principle influenced by the transcript length as it is the case for classic RNA-sequencing and therefore no further normalization is performed to correct for transcript length as usually performed with RPKM. The density of mapped reads in -500 nt terminal region of 3' UTR isoforms is bimodal, with a low-density peak probably corresponding to background transcription, i.e. 3' UTR isoforms of low abundance or 3' UTR isoforms to which reads were spuriously mapped, and a high-density peak corresponding to expressed 3' UTR isoforms (**Supplemental Fig. S4A**). In order to identify 3' UTR isoforms expressed in axons and cell body, a two-component Gaussian mixture was fitted to the data using the R package mclust (Fraley and Raftery 2002). An isoform was called expressed if there were less than 5% chance of belonging to the background category in both replicates or if there was more than 10% chance of belonging to the expressed category in at least one replicate.

Analysis of Alternative cleavage and polyadenylation (APA)

The analysis of APA was performed as previously described (Andreassi et al. 2021) and as detailed in Extended Methods. In brief, we first extracted the $\log_2$ ratio of promoter-proximal and promoter-distal 3' UTR isoform expression levels, hereafter called RUD, in each sample. We also calculated the ratios between the read count in the promoter-proximal and the sum of the read counts in the promoter-proximal and the promoter-distal 3' UTR isoforms, hereafter called PUD. In the case of multiple promoter-distal 3' UTR isoforms per transcript ID, a value was computed for each individual promoter-distal 3' UTR isoform. In order to identify transcripts that show a marked change in the 3' UTR isoform between conditions, we scored the differences in proximal-to-distal poly(A) site usage using both the differences in RUD and PUD between conditions of interest. The statistical significance of the changes in proximal-to-distal poly(A) site ratio between conditions was assessed by Fisher's exact count test. We adjusted the P-Value controlling for False Discovery Rate (FDR) of 0.01.

Axonal localisation analysis

The ratio between genes' abundance in the axons and in the cell body is often used to quantify mRNA axonal localisation (Olguin et al. 2022; Martínez et al. 2019), however this metric correlates with the mRNA abundance in the cell body and the transcript length

(**Supplemental Fig 3C,D**). Consequently, such metrics fail to identify highly transported but lowly expressed transcripts, and similarly they are more likely to associate highly expressed transcript with high mRNA axonal localisation scores despite low transport efficiency, as compared to transcripts of similar expression level. Here we aimed to develop a novel metric that accurately infers the axonal transport efficiency and stability, irrespective of the transcript length and the expression in the cell body using a hierarchical Bayesian model procedure.

First, we looked at the global relationship between the normalised read count per 3' UTR isoforms in the axonal compartment and their corresponding normalised read counts in the cell body compartment. We found that the average 3' UTR isoform abundance level -average $\log_2$ expression- in the axonal compartments of either NGF or NT-3 condition is best approximated by a combination of polynomial regression model of degree four of the abundance in the cell body and a linear model of the transcript length as revealed by the Akaike's An Information Criterion from ANOVA analysis (**Supplemental Table S22** and **Supplemental Fig. S4E**). Given our goal to set out a metric of axonal localisation independent of the cell body read counts and the transcript length, we created 102 groups of 3' UTR isoforms of fixed expression ranges in the cell body (18 bins from 2^3 to 2^{20} nucleotides) and fixed ranges of transcript lengths (10 bins from 10^2 to $10^{4.5}$ nucleotides). For each of these 102 groups, we generated a matrix of 10^4 simulated draws of axonal read count values predicted using the fitted polynomial regression model of degree four, obtained from the total pool of transcripts, over a regularly interspaced grid of 100 possible transcript lengths -ranging from the minimal to the maximal transcript length of this specific group, and 100 possible cell body read count values -again ranging from the minimal to the maximal cell body read count of this specific group. We then extracted the average and variance over the predicted 10^4 axonal read count values for each of these 102 groups. Using these 102 pairs of averages and variances, we next fitted 102 negative binomial distributions to the 102 groups of axonal read counts by maximum likelihood (mle) using the fitdist function from the fitdistrplus R package (Dutang). These 102 posterior predictive distributions of axonal read counts could then serve to assess whether the observed

axonal abundance of each individual 3' UTR isoform y_i^{obs} are consistent with the fitted models given their associated transcript length and cell body abundance. Here we propose that the more extreme the y_i^{obs}, for a given 3' UTR isoform of specific transcript length and cell body abundance, is on the histogram of simulated values $y_i^{predict}$ from the corresponding predictive distribution, the more likely the axonal localization of its corresponding 3' UTR isoform has been actively regulated as opposed to the unspecific active transport which is expected to affect most transcripts detected in the axons. If y_i^{obs} is in the lower tail, we expect this 3' UTR isoform to be either restricted to the cell body or actively degraded in the axonal compartment; conversely if y_i^{obs} is in the higher tail of the histogram, we expect the 3' UTR isoform to be either actively transported or stabilised in the axonal compartment. Thus for each 3' UTR isoform i, we next computed the proportion of 10^4 values, randomly generated from the posterior negative distribution associated with its corresponding transcript length and abundance in the cell body, that were smaller or larger than the observed axonal read count. In order to get a single value per transcript, hereafter called *Localisation Score*, we selected the smallest of these two values (probabilities to observe smaller or larger axonal read count for a given isoform), transformed it using the $\log_{10}$, and multiplied it by -1 when the latter was selected. Hence positive values are associated with active transport or stabilisation, while negative values are associated with cell body restriction or degradation. Notably while the ratios between genes abundance in the axons and in the cell body depend on cell body abundance (**Supplemental Fig. S4F**, *upper*) and transcript length (**Supplemental Fig. S4G**, *upper*), this is not the case anymore with the novel localisation score metric (**Supplemental Figs. S4F,G**, *lower*). This analysis has been restricted on the 30,450 3' UTR isoforms detected in the axonal compartments of neurons either exposed to NGF or NT-3. The axonal localisation scores for these 3' UTR isoforms have been computed for NGF and NT-3 conditions independently and are reported in **Supplemental Table S9**.

Mapping and analysis of CLIP data

582 To identify RBPs that bind to 3' UTR regions, we examined iCLIP data for 18 RBPs (Attig et
583 al. 2018), and eCLIP data from K562 and HepG2 cells for 89 and 70 RBPs, respectively,
584 available from ENCODE (Van Nostrand et al. 2020; Sloan et al. 2016). In total we analysed
585 CLIP-seq data for 126 RBPs (see **Supplemental Table S6** for complete list of CLIP
586 sequencing data). Before mapping the reads, adapter sequences were removed using
587 cutadapt v1.9.dev1 (Martin 2011) and reads shorter than 18 nucleotides were dropped from
588 the analysis. Reads were mapped with STAR v2.4.0i (Dobin et al. 2013) to UCSC
589 hg19/GRCh37 genome assembly. To quantify binding to individual loci, only uniquely mapping
590 reads were used. The results were lifted to rn5 using liftOver (Hinrichs et al. 2006). The
591 mapping of the CLIP data has been done in 2018. While updated versions of the human
592 genome assembly have been released since then (GRCh38, T2T-CHM13), mapping the CLIP
593 to a newer version of the genome assembly is not expected to significantly change the results
594 given that novel annotated regions are not biased towards specific regions (3' UTR long *versus*
595 short isoform) or to specific groups of genes related to neuronal functions. Indeed the statistics
596 based on these data capitalise on the large variety of transcripts and regions where these are
597 mapped.

598 *RBP regulome underlying alternative polyadenylation*

599 In order to identify positive and negative RBPs regulators of alternative polyadenylation in
600 developing rat sympathetic neurons, we tested the association between RBP binding in
601 defined regions along the 3' UTR and the relative usage of the promoter-proximal 3' UTR
602 isoforms (PUD) in the cell body compartments of either NGF or NT-3 conditions. In particular
603 we used the Welch's *t*-test to compare the distributions of the PUD between groups of isoforms
604 exhibiting or not cross-link events in the following 30 defined regions surrounding the 3' ends
605 of the *promoter-proximal* 3' UTR isoforms : [-3000:-2950],[-2750:-2700],[-2500:-2450],[-2250:-
606 2200],[-2000:-1950],[-1750:-1700],[-1500:-1450],[-1400:-1350],[-1300:-1250],[-1200:-1150],[-
607 1100:-1050],[-1000:950],[-900:-850],[-800:-750],[-700:-650],[-600:-550],[-500:-450],[-450:-
608 400],[-400:-350],[-350:-300],[-300:-250],[-250:-200],[-200:-150],[-150:-100],[-100:-50],[-
609 50,0],[0,+50],[+50:+100],[+100:+150] for 126 RBPs with available CLIP-seq data in human cell

clines, thereby obtaining 30 *P*-values per RBPs. These were then minus $\log_{10}$-transformed and multiplied by the sign of the difference in PUD between the group of isoforms exhibiting or not a cross-link event in the defined region along the 3' UTR. Thus positive value indicates that RBP binding to regions surrounding the 3' end of the *promoter-proximal* 3' UTR promotes the usage of the promoter-proximal 3' UTR isoforms, hence acting as positive regulators of the promoter-proximal and negative regulators of promoter-distal 3' UTR isoform, while negative value indicates that RBP binding to regions surrounding the 3' end of the promoter-proximal 3' UTR promotes the usage of the promoter-distal, hence acting as negative regulators of the promoter-proximal and positive regulators of promoter-distal 3' UTR isoforms. We repeated this analysis to compare the distributions of the PUD between groups of isoforms exhibiting or not cross-link events in the same 30 defined regions surrounding the 3' ends of the *promoter-distal* 3' UTR isoforms, therefore inspecting the regulatory potential of RBP through the binding of the distal 3' UTR isoforms. Thus, we obtained a map of regulatory potential of the relative 3' UTR usage along the short and long 3' UTR for each of the 126 RBPs.

RBP regulome underlying APA between NGF and NT-3 and axonal remodelling

To identify RBPs regulators of APA between NGF and NT-3, we performed Fisher count enrichment analysis to test for significant enrichment in RBP cross-link events in defined regions (see previous paragraph) around the 3' end of either the promoter-proximal or the promoter-distal 3' UTR isoform between the groups of isoforms exhibiting significant shifts in either groups and the total pool of 3' UTR isoforms. This analysis is indeed more efficient in recovering candidate RBP regulators as compared to the Welch's *t*-test comparing the distributions in $\overline{\Delta_{PUD}}$ between groups of isoforms exhibiting or not a cross-link event given the relatively low number of pairs of isoforms exhibiting promoter-proximal or distal shifts in the cell bodies of NGF and NT-3 treated neurons compared to the full set of 3' UTR isoforms. We used a similar approach to identify candidate regulators of axonal remodelling. Specifically we used the Fisher count test to assess for significant enrichment in RBP cross-link events in defined regions along the 3' UTRs of predicted remodelled promoter-proximal 3' UTR isoforms (n=80 in NGF and n=60 in NT-3) as compared with the total pool of promoter-proximal 3' UTR.

RBP regulome underlying axonal transport

To identify positive and negative RBPs regulators of axonal 3' UTR isoform localisation and stability in developing rat sympathetic neurons, we tested the association between RBP binding in defined regions along the 3' UTR and the localisation score in NGF or NT-3 conditions. We used the Welch's *t*-test to compare the distributions of the localisation scores between groups of isoforms exhibiting or not cross-link events in 30 defined regions surrounding their 3' ends (see above for detailed regions), for 126 RBPs with available CLIP-seq data in human cell lines, thereby obtaining 30 *P*-values per RBPs. These were then minus $\log_{10}$-transformed and multiplied by the sign of the difference in localisation scores between the group of isoforms exhibiting or not a cross-link event in the defined region along the 3' UTR. Thus positive values indicates that RBP binding to regions surrounding the 3' end of 3' UTR isoform promotes axonal transport and/or stability, while negative values indicates indicate lower localisation scores for the group of isoforms exhibiting cross-link events as compared with the group that do not exhibit cross-link event hence indicating that that RBP binding to regions surrounding the 3' end of 3' UTR isoforms prevents axonal transport and/or promotes mRNA degradation.

Combinatorial regulatory potential of RBPs in axonal localisation

To investigate potential combinatorial effects of RBPs in regulating axonal mRNA localisation, we used a binary logistic regression classifier trained to distinguish between over- and under-transported 3' UTR isoforms. For this analysis, we focused on the 5% isoforms exhibiting the highest (n=1356) and lowest (n=1356) localisation scores in each treatment conditions to have a sufficiently high number of training data (**Supplemental Fig. S5F**), ensuring a balanced representation of both classes. We split the total pool of 3' UTR isoforms into a training (n=2168) and a testing set (n=542). The training set was utilized to perform randomized search on hyperparameters, where a cross-validation strategy with 20 folds was employed to assess the performance of the different hyperparameter combinations.

A combination of four classifiers were trained and tested where the predictor variables were respectively 1) evidence of cross-link events for the 32 positive and 11 negative regulators of

axonal transport identified using Welch's *t*-test (M1); 2) evidence of cross-link events for the 32 positive regulators of axonal transport identified using statistical testing (M2; **Supplemental Table S12)**, 3) evidence of cross-link events for the 11 negative regulators of axonal transport identified using statistical testing (M3; **Supplemental Table S13)**, and 4) combinations of cross-link events involving 42 selected pairs of the 43 RBPs identified as potential positive and negative regulators of axonal transport in the region preceding the 3' end, specifically the [-250, -50] nucleotide region (M4). An isoform was considered bound by an RBP if there were evidence of cross-link event in the region [-250:-50] nt preceding the 3 'end in at least one of the multiple CLIP sequencing data generated in different cell lines (HepG2, K652) and with different techniques (eClip and iClip).

The development of the M4 model aimed to study the synergistic potential of RBP was limited by the size of the training set (n=2000 3' UTR isoforms) making it challenging to learn the coefficients for the total number of possible pairwise interaction terms (ntot=583; n=528 interactions terms for 32 positive candidate regulators of axonal localisation and n=55 interactions terms for the negative regulators). We therefore selected the four RBPs exhibiting most positive weights and the four RBPs exhibiting most negative weights in the M1 regression models trained with the total set of 43 positive and negative regulators in either NGF (top positive regulators: SF3B4, SRSF1, SNRPB, EFTUD2; top negative regulators: CELF4, KHDRBS1, PUM2 and FUBP3) or NT-3 (top positive regulators: SF3B4, DDX24, SLTM and SRSF9; top negative regulators: CELF4, KHSRP, CELF2, HNRNPC) (M1; **Supplemental Figs. S6G,H**) conditions. We next created 42 pairwise interaction predictor variables (21 positive and 21 negative interaction terms) by identifying 3' UTR isoforms exhibiting evidence of cross-link events for pairs of RBPs.

All logistic regressions have been fitted using the *sklearn* library in Python (Pedregosa et al. 2011). For each set-up, we performed a randomized search to optimize the hyperparameters, namely the solver among ['lbfgs', 'newton-cg', 'liblinear', 'sag', 'saga'] and the regularization technique among ['l2', 'l1', 'elasticnet']. We fixed the inverse regularization term to C=0.01 for all models, to make it easier to compare and interpret the performance between the different

models, and the maximum number of iterations taken for the solvers to converge has been set to max_iter=10000. The selected regularization term for both NGF and NT-3 sets is the L2 penalty term for all four models, while the solver depends on the model (lbfgs for M1, M3 and M4 and liblinear for M2 in the NGF condition; sag for M1, liblinear for M2 and lbfgs for M3 and M4 in the NT-3 condition). The predictive power of this set of models was determined using the recall and precision, and F1-score.

To identify the statistically significant predictors in each model (RBPs or interaction of RBPs) in each culture condition, we compared the observed predictor coefficients to their respective null distributions. To generate a null distribution for each predictor, we shuffled the response variable label randomly, re-trained the model and re-calculated the predictor coefficients. We repeated this process 2000 times. To compare the observed predictor coefficients with the null distributions generated, we computed their *Z*-scores as:

$$z_{predictor} = \frac{x - \mu}{\sigma}$$

where x is the observed predictor coefficient, μ and σ are the mean and standard deviation of the null distribution for this predictor. If $|z_{predictor}| > 1.645$ (95% confidence; one-tailed), then the predictor is a significant regulator of localisation. In particular, if $z_{predictor} < 0$, the predictor is a negative regulator of localisation, while if $z_{predictor} > 0$, the predictor is a positive regulator of localisation. To compare regulators between the NGF and NT-3 conditions, we applied the following rule: if a predictor is significant in both the NGF and the NT-3 model, then this predictor is a common regulator of localisation. However, if a predictor is significant in only one of the two models, this predictor is a specific regulator of this model.

Statistical analyses

Data are expressed as averages ± SEM. *t*-test was used as indicated to test for statistical significance, which was placed at least $P < 0.05$ unless otherwise noted. Statistical analysis was performed with the R (Foundation for Statistical Computing) statistical package version 4.2.2 (2022) and Bioconductor libraries version 3.16 (R Core Team. R: A Language and

721 Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical
722 Computing; 2013).

723

724 **DATA ACCESS**

725 The processed data required to implement the code and generate figures are available
726 at Zenodo under the accession number 8047412 (https://zenodo.org/record/8047412).
727 Supplemental Tables have also been deposited in Zenodo (same accession number). Code
728 and scripts used to make the results are available at GitHub
729 (https://github.com/RLuisier/AxonLoc) and as Supplemental Code.

730

731 **COMPETING INTEREST STATEMENT**

732 The authors declare that they have no competing interests.

733

734 **ACKNOWLEDGEMENTS**

735 We thank all members of the Riccio and Luisier laboratories, as well as Pierre Klein for
736 stimulating discussions and critical review of the data. The work was supported by the
737 Wellcome Trust Investigator Awards 103717/Z/14/Z and 217213/Z/19/Z (to A.R.), the MRC
738 LMCB Core Grant MC/U12266B (to A.R), a Wellcome Trust Institutional Strategic Support
739 Fund (to C.A.), and Idiap Research Institute (to R.L. and L.F.)

740

741 **AUTHORS CONTRIBUTION**

742 R.L., C.A. and A.R. conceived and designed the study. C.A. performed the screen and all
743 experiments presented in the study. R.L. designed the computational framework, performed
744 data analysis and interpretation of results, designed the figures and derived the model. L.F.
745 developed machine learning frameworks for combinatorial analysis of RBPome underlying

axonal localisation. R. L. and A.R. wrote the manuscript with the critical input from C.A and L.F. All authors discussed the results and contributed to the final manuscript.

REFERENCES

Allen M, Bird C, Feng W, Liu G, Li W, Perrone-Bizzozero NI, Feng Y. 2013. HuD promotes BDNF expression in brain neurons via selective stabilization of the BDNF long 3' UTR mRNA. *PLoS One* **8**: e55718.

Andreassi C, Crerar H, Riccio A. 2018. Post-transcriptional Processing of mRNA in Neurons: The Vestiges of the RNA World Drive Transcriptome Diversity. *Front Mol Neurosci* **11**: 304.

Andreassi C, Luisier R, Crerar H, Darsinou M, Blokzijl-Franke S, Lenn T, Luscombe NM, Cuda G, Gaspari M, Saiardi A, et al. 2021. Cytoplasmic cleavage of IMPA1 3' UTR is necessary for maintaining axon integrity. *Cell Rep* **34**: 108778.

Andreassi C, Riccio A. 2009. To localize or not to localize: mRNA fate is in 3'UTR ends. *Trends Cell Biol* **19**: 465–474.

Andreassi C, Zimmermann C, Mitter R, Fusco S, De Vita S, Devita S, Saiardi A, Riccio A. 2010. An NGF-responsive element targets myo-inositol monophosphatase-1 mRNA to sympathetic neuron axons. *Nat Neurosci* **13**: 291–301.

Antar LN, Afroz R, Dictenberg JB, Carroll RC, Bassell GJ. 2004. Metabotropic glutamate receptor activation regulates fragile x mental retardation protein and FMR1 mRNA localization differentially in dendrites and at synapses. *J Neurosci* **24**: 2648–2655.

Ascano M, Bodmer D, Kuruvilla R. 2012. Endocytic trafficking of neurotrophins in neural development. *Trends Cell Biol* **22**: 266–273.

Attig J, Agostini F, Gooding C, Chakrabarti AM, Singh A, Haberman N, Zagalak JA, Emmett W, Smith CWJ, Luscombe NM, et al. 2018. Heteromeric RNP Assembly at LINEs Controls Lineage-Specific RNA Processing. *Cell* **174**: 1067–1081.e17.

Barbacid M. 1994. The Trk family of neurotrophin receptors. *J Neurobiol* **25**: 1386–1403.

Bhattacharyya A, Watson FL, Bradlee TA, Pomeroy SL, Stiles CD, Segal RA. 1997. Trk receptors function as rapid retrograde signal carriers in the adult nervous system. *J Neurosci* **17**: 7007–7016.

Cosker KE, Fenstermacher SJ, Pazyra-Murphy MF, Elliott HL, Segal RA. 2016. The RNA-binding protein SFPQ orchestrates an RNA regulon to promote axon viability. *Nat Neurosci* **19**: 690–696.

Dalla Costa I, Buchanan CN, Zdradzinski MD, Sahoo PK, Smith TP, Thames E, Kar AN, Twiss JL. 2020. The functional organization of axonal mRNA transport and translation. *Nat Rev Neurosci*. http://dx.doi.org/10.1038/s41583-020-00407-7.

Das S, Singer RH, Yoon YJ. 2019. The travels of mRNAs in neurons: do they know where they are going? *Curr Opin Neurobiol* **57**: 110–116.

785 Dictenberg JB, Swanger SA, Antar LN, Singer RH, Bassell GJ. 2008. A direct role for FMRP in
786 activity-dependent dendritic mRNA transport links filopodial-spine morphogenesis to
787 fragile X syndrome. *Dev Cell* **14**: 926–939.

788 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M,
789 Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.

790 Dorskind JM, Kolodkin AL. 2021. Revisiting and refining roles of neural guidance cues in circuit
791 assembly. *Curr Opin Neurobiol* **66**: 10–21.

792 Doyle M, Kiebler MA. 2011. Mechanisms of dendritic mRNA transport and its role in synaptic
793 tagging. *EMBO J* **30**: 3540–3552.

794 Dutang M. Fitdistrplus: an R package for fitting distributions. *J Stat Softw*.

795 Elshamy WM, Ernfors P. 1996. Requirement of neurotrophin-3 for the survival of proliferating
796 trigeminal ganglion progenitor cells. *Development* **122**: 2405–2414.

797 Engel KL, Lo H-YG, Goering R, Li Y, Spitale RC, Taliaferro JM. 2022. Analysis of subcellular
798 transcriptomes by RNA proximity labeling with Halo-seq. *Nucleic Acids Res* **50**: e24.

799 Erson-Bensan AE. 2016. Alternative polyadenylation and RNA-binding proteins. *J Mol*
800 *Endocrinol* **57**: F29–34.

801 Foundation for Statistical Computing RR. R: A language and environment for statistical
802 computing. *RA Lang Environ Stat Comput*.

803 Fraley C, Raftery AE. 2002. MCLUST: Software for Model-Based Clustering, Density
804 Estimation and Discriminant Analysis. http://dx.doi.org/10.21236/ada459792.

805 Harrington AW, St Hillaire C, Zweifel LS, Glebova NO, Philippidou P, Halegoua S, Ginty DD.
806 2011. Recruitment of actin modifiers to TrkA endosomes governs retrograde NGF
807 signaling and survival. *Cell* **146**: 421–434.

808 Hentze MW, Castello A, Schwarzl T, Preiss T. 2018. A brave new world of RNA-binding
809 proteins. *Nat Rev Mol Cell Biol*. http://dx.doi.org/10.1038/nrm.2017.130.

810 Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H, Diekhans M, Furey
811 TS, Harte RA, Hsu F, et al. 2006. The UCSC Genome Browser Database: update 2006.
812 *Nucleic Acids Res* **34**: D590–8.

813 Holt CE, Martin KC, Schuman EM. 2019. Local translation in neurons: visualization and
814 function. *Nat Struct Mol Biol* **26**: 557–566.

815 Holt CE, Schuman EM. 2013. The central dogma decentralized: new perspectives on RNA
816 function and local translation in neurons. *Neuron* **80**: 648–657.

817 Kanai Y, Dohmae N, Hirokawa N. 2004. Kinesin transports RNA: isolation and characterization
818 of an RNA-transporting granule. *Neuron* **43**: 513–525.

819 Kar AN, MacGibeny MA, Gervasi NM, Gioio AE, Kaplan BB. 2013. Intra-axonal synthesis of
820 eukaryotic translation initiation factors regulates local protein synthesis and axon growth in
821 rat sympathetic neurons. *J Neurosci* **33**: 7165–7174.

822 Kislauskis EH, Zhu X, Singer RH. 1994. Sequences responsible for intracellular localization of
823 beta-actin messenger RNA also affect cell phenotype. *J Cell Biol* **127**: 441–451.

824 Krichevsky AM, Kosik KS. 2001. Neuronal RNA granules: a link between RNA localization and
825 stimulation-dependent translation. *Neuron* **32**: 683–696.

826 Kuechler ER, Jacobson M, Mayor T, Gsponer J. 2022. GraPES: The Granule Protein
827 Enrichment Server for prediction of biological condensate constituents. *Nucleic Acids Res*.
828 http://dx.doi.org/10.1093/nar/gkac279.

829 Kuruvilla R, Zweifel LS, Glebova NO, Lonze BE, Valdez G, Ye H, Ginty DD. 2004. A
830 neurotrophin signaling cascade coordinates sympathetic neuron development through
831 differential control of TrkA trafficking and retrograde signaling. *Cell* **118**: 243–255.

832 Laishram RS, Anderson RA. 2010. The poly A polymerase Star-PAP controls 3'-end cleavage
833 by promoting CPSF interaction and specificity toward the pre-mRNA. *EMBO J* **29**: 4132–
834 4145.

835 Loedige I, Baranovskii A, Mendonsa S, Dantsuji S, Popitsch N, Breimann L, Zerna N,
836 Cherepanov V, Milek M, Ameres S, et al. 2023. mRNA stability and m6A are major
837 determinants of subcellular mRNA localization in neurons. *Mol Cell*.
838 http://dx.doi.org/10.1016/j.molcel.2023.06.021.

839 Makita T, Sucov HM, Gariepy CE, Yanagisawa M, Ginty DD. 2008. Endothelins are vascular-
840 derived axonal guidance cues for developing sympathetic neurons. *Nature* **452**: 759–763.

841 Martínez JC, Randolph LK, Iascone DM, Pernice HF, Polleux F, Hengst U. 2019. Pum2
842 Shapes the Transcriptome in Developing Axons through Retention of Target mRNAs in
843 the Cell Body. *Neuron* **104**: 931–946.e5.

844 Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing
845 reads. *EMBnet.journal* **17**: 10–12.

846 Mayr C. 2017. Regulation by 3′-Untranslated Regions. *Annual Review of Genetics* **51**: 171–
847 194. http://dx.doi.org/10.1146/annurev-genet-120116-024704.

848 Mitschka S, Mayr C. 2022. Context-specific regulation and function of mRNA alternative
849 polyadenylation. *Nat Rev Mol Cell Biol*. http://dx.doi.org/10.1038/s41580-022-00507-5.

850 Nagano S, Jinno J, Abdelhamid RF, Jin Y, Shibata M, Watanabe S, Hirokawa S, Nishizawa M,
851 Sakimura K, Onodera O, et al. 2020. TDP-43 transports ribosomal protein mRNA to
852 regulate axonal local translation in neuronal axons. *Acta Neuropathol* **140**: 695–713.

853 Nussbacher JK, Tabet R, Yeo GW, Lagier-Tourenne C. 2019. Disruption of RNA Metabolism in
854 Neurological Diseases and Emerging Therapeutic Interventions. *Neuron* **102**: 294–320.

855 Olguin SL, Patel P, Buchanan CN, Dell'Orco M, Gardiner AS, Cole R, Vaughn LS,
856 Sundararajan A, Mudge J, Allan AM, et al. 2022. KHSRP loss increases neuronal growth
857 and synaptic transmission and alters memory consolidation through RNA stabilization.
858 *Commun Biol* **5**: 672.

859 Patel P, Buchanan CN, Zdradzinski MD, Sahoo PK, Kar AN, Lee SJ, Vaughn LS, Urisman A,
860 Oses-Prieto J, Dell'Orco M, et al. 2022. Intra-axonal translation of Khsrp mRNA slows
861 axon regeneration by destabilizing localized mRNAs. *Nucleic Acids Res* **50**: 5772–5792.

862 Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer
863 P, Weiss R, Dubourg V, et al. 2011. Scikit-learn: Machine learning in Python. *the Journal
864 of machine Learning research* **12**: 2825–2830.

865 Riccio A, Ahn S, Davenport CM, Blendy JA, Ginty DD. 1999. Mediation by a CREB family

transcription factor of NGF-dependent survival of sympathetic neurons. *Science* **286**: 2358–2361.

Riccio A, Pierchala BA, Ciarallo CL, Ginty DD. 1997. An NGF-TrkA-mediated retrograde signal to transcription factor CREB in sympathetic neurons. *Science* **277**: 1097–1100.

Ross RA, Lazarova DL, Manley GT, Smitt PS, Spengler BA, Posner JB, Biedler JL. 1997. HuD, a neuronal-specific RNA-binding protein, is a potential regulator of MYCN expression in human neuroblastoma cells. *Eur J Cancer* **33**: 2071–2074.

Scott-Solomon E, Boehm E, Kuruvilla R. 2021. The sympathetic nervous system in development and disease. *Nat Rev Neurosci* **22**: 685–702.

Scott-Solomon E, Kuruvilla R. 2018. Mechanisms of neurotrophin trafficking via Trk receptors. *Mol Cell Neurosci* **91**: 25–33.

Shav-Tal Y, Singer RH. 2005. RNA localization. *J Cell Sci* **118**: 4077–4081.

Sloan CA, Chan ET, Davidson JM, Malladi VS, Strattan JS, Hitz BC, Gabdank I, Narayanan AK, Ho M, Lee BT, et al. 2016. ENCODE data at the ENCODE portal. *Nucleic Acids Res* **44**: D726–32.

Snee M, Kidd GJ, Munro TP, Smith R. 2002. RNA trafficking and stabilization elements associate with multiple brain proteins. *J Cell Sci* **115**: 4661–4669.

Štalekar M, Yin X, Rebolj K, Darovic S, Troakes C, Mayr M, Shaw CE, Rogelj B. 2015. Proteomic analyses reveal that loss of TDP-43 affects RNA processing and intracellular transport. *Neuroscience* **293**: 157–170.

St Johnston D. 2005. Moving messages: the intracellular localization of mRNAs. *Nat Rev Mol Cell Biol* **6**: 363–375.

Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, et al. 2017. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* **45**: D362–D368.

Terenzio M, Koley S, Samra N, Rishal I, Zhao Q, Sahoo PK, Urisman A, Marvaldi L, Oses-Prieto JA, Forester C, et al. 2018. Locally translated mTOR controls axonal local translation in nerve injury. *Science* **359**: 1416–1421.

Terenzio M, Schiavo G, Fainzilber M. 2017. Compartmentalized Signaling in Neurons: From Cell Biology to Neuroscience. *Neuron* **96**: 667–679.

Thelen MP, Kye MJ. 2019. The Role of RNA Binding Proteins for Local mRNA Translation: Implications in Neurological Disorders. *Front Mol Biosci* **6**: 161.

Tian B, Manley JL. 2013. Alternative cleavage and polyadenylation: the long and short of it. *Trends Biochem Sci* **38**: 312–320.

Tian B, Manley JL. 2017. Alternative polyadenylation of mRNA precursors. *Nat Rev Mol Cell Biol* **18**: 18–30.

Tushev G, Glock C, Heumüller M, Biever A, Jovanovic M, Schuman EM. 2018. Alternative 3' UTRs Modify the Localization, Regulatory Potential, Stability, and Plasticity of mRNAs in Neuronal Compartments. *Neuron* **98**: 495–511.e6.

Van Nostrand EL, Freese P, Pratt GA, Wang X, Wei X, Xiao R, Blue SM, Chen J-Y, Cody NAL, Dominguez D, et al. 2020. A large-scale binding and functional map of human RNA-binding proteins. *Nature* **583**: 711–719.

Wagnon JL, Briese M, Sun W, Mahaffey CL, Curk T, Rot G, Ule J, Frankel WN. 2012. CELF4 regulates translation and local abundance of a vast set of mRNAs, including genes associated with regulation of synaptic function. *PLoS Genet* **8**: e1003067.

Wang W, van Niekerk E, Willis DE, Twiss JL. 2007. RNA transport and localized protein synthesis in neurological disorders and neural repair. *Dev Neurobiol* **67**: 1166–1182.

Willis DE, van Niekerk EA, Sasaki Y, Mesngon M, Merianda TT, Williams GG, Kendall M, Smith DS, Bassell GJ, Twiss JL. 2007. Extracellular stimuli specifically regulate localized levels of individual neuronal mRNAs. *J Cell Biol* **178**: 965–980.

Zhu H, Zhou H-L, Hasman RA, Lou H. 2007. Hu proteins regulate polyadenylation by blocking sites containing U-rich sequences. *J Biol Chem* **282**: 2203–2210.

Figure Legends

Figure 1. NGF and NT-3 are equally capable of regulating transcriptional changes when applied to distal axons. *(A)* Schematic of the experimental set-up. *(B)* Volcano plot representing $\log_2$ fold-change ($\log_2$FC) in gene expression values between neurons whose axons were exposed to NGF or NT-3, and corresponding P-values ($-\log_{10}$). Blue dots = genes significantly up-regulated in NGF; purple dots = genes significantly up-regulated in NT-3 (FC>1.5 and P-value <0.01). *(C)* Enrichment scores of GO biological pathways associated with up-regulated genes in NGF- (*upper*) and NT-3- treated (*lower*) neurons. *(D)* Fisher enrichments in predicted binding sites motifs for 21 transcription transcription factors in the 1000 nucleotide promoter region of the up-regulated genes in NGF or NT-3 treated neurons. *(E,F)* Sequence logos of predicted motifs bound by transcription factors (*upper*) and fractions of promoter regions with these motifs (*lower*) in the total pool of expressed genes (grey bar), the genes up-regulated in NGF (blue bar) and the genes up-regulated in NT-3 (purple bar). Fisher enrichment test. *(G)* (*left*) Number of transcripts exhibiting significant promoter distal-to-proximal shifts in NGF compared to NT-3 (full bar) alongside their composition in terms of proximal shift in NGF versus distal shift in NT-3 (outlined bar). (*right*) Legend of the color outline schematizing the relative usage of the promoter-proximal (Ip) and promoter-distal (Id) 3' UTR isoforms underlying differential 3' UTR isoform usage between NGF and NT-3. *(H)* Same as *(G)* for changes in NT-3 compared to NGF. (**I**) Enrichment scores of GO biological pathways associated with significant increase in distal-to-proximal promoter 3' UTR usage in NGF compared to NT-3 (*upper*) and vice-versa (*lower*).

Figure 2. Predicted RBP regulome underlying alternative polyadenylation. *(A)* Distribution of significant positive association between the binding of 126 RBPs in defined regions along the 3' UTR and the relative usage of the promoter-proximal (*upper*) or promoter-distal (*lower*) 3' UTR isoform. Dark lines display the median significance and shaded areas indicate lower and upper quartiles. *(B,C)* Scatterplot of the extent of standardized significant association between CPSF1 *(B)* or CSTF2 *(C)* cross-link event in defined regions along the 3' UTR and the relative usage of the promoter-proximal (black) or promoter-distal (grey) 3' UTR isoform in the axons. *(D)* Negative association between RBP binding and the relative usage of the promoter-proximal (*upper*) or promoter-distal (*lower*) 3' UTR isoform. *(E)* Network of protein-protein interactions for 27 candidate regulators of APA predicted to prevent promoter-proximal usage when bound to the [0:150] nucleotide region down-stream the 3' end. Edges represent experimentally determined protein-protein interactions annotated in the STRING database (Szklarczyk et al. 2017). Nodes indicate proteins coloured according to biological

pathways they are enriched in. *(F)* Top five biological pathways and associated P-values over-represented in the 27 candidate negative regulators of short 3' UTR isoform. Fisher enrichment test. *(G)* Same as *(B,C)* for ELAVL1. *(H)* Distributions of the propensity scores of the 126 RBPs, the 17 positive regulators of APA (short and long 3' UTR isoform) and the 27 negative regulators of the short 3' UTR isoform to localize into cellular condensates as predicted by GraPES(Kuechler et al. 2022). Welch's *t*-test assessing the significant difference between the mean propensity scores.

Figure 3. Reduction of negative regulatory activity is predicted to underlie neurotrophin-driven APA. *(A-D)* Distribution of Fisher enrichment scores ($-\log_{10}$(*P*-value)) in cross-link events of 126 RBPs along the short 3' UTR isoforms of the pairs of isoforms exhibiting significant promoter-proximal shifts in NT-3 *(A)* or NGF *(B),* and promoter-distal shifts in NT-3 *(C)* or NGF *(D)*. *(E-H)* Networks of protein-protein interactions for the RBPs exhibiting significant enrichment in cross-link events in the [0:+150] nt regions downstream the 3' end of the short 3' UTR isoforms associated with significant promoter-proximal shifts in NT-3 *(A)* or NGF *(B)*, and promoter-distal shifts in NT-3 *(C)* or NGF *(D)*. Edges represent experimentally determined protein-protein interactions annotated in the STRING database. Nodes indicate proteins coloured according to biological pathways they are enriched in. Arrows indicate the predicted directions in changes in activity of the RBPs. *(I-L)* Genome browser views of 3' end sequencing profiles and CLIP crosslinking events for predicted positive (*green*) and negative (*gold*) regulators of APA for 3' UTR isoforms showing a marked shift toward increase in promoter-proximal 3' UTR usage (*Rbm41* and *Slc39a13*), or in promoter-distal 3' UTR usage (*Usp15* and *Ppfia4*) in either condition. Grey boxes highlight the location of promoter-proximal and promoter-distal 3' UTR isoforms.

Figure 4. Axons exposed to NGF or NT-3 contain distinct 3' UTR isoforms. *(A)* Number of Ensembl transcripts ID detected in cell bodies only (full bars), and cell bodies and distal axons (empty bars) in NGF (blue) and NT-3 (purple) culture conditions. *(B)* Overlap between the transcripts detected in the distal axons of neurons exposed to either NGF or NT-3. *(C)* Cell body (full bars) and axonal (empty bars) transcript IDs showing multiple 3′ UTRs in either NGF or NT-3 culture conditions. Two-sided Fisher's exact count test. *(D)* Maximum 3′ UTR lengths for existing annotations in Ensembl Rn5 (Ensembl) and for those newly identified by 3′ end RNA sequencing (Extended annotation) in either NGF or NT-3. Two-sided Wilcoxon rank-sum test. *(E)* Top five GO biological pathways which associated transcripts exhibiting the most significant increase in axonal localisation in NGF (*left)* compared to NT-3 (*right)*, as quantified with standardized scores comparing GO-annotated transcript with the full pool of detected transcripts (*Z*-score). *(F)* Distributions of the

localisation scores for the background genes and the genes belonging to the collagen catabolic biological pathway in NGF (*right*), in NT-3 (*center*), and the differences in localisation scores between the two conditions (*right*). Welch's *t*-test assessing the significant difference between the mean localisation scores. Boxplots display the five number summary of median, lower and upper quartiles, minimum and maximum values. *(G)* Same as *(E)* for the top five GO biological pathways exhibiting excess in axonal localisation in NT-3 compared to NGF. *(H)* Same as *(F)* for genes belonging to the regulation of VEGF production. *(I)* Average mRNA abundance across the four cell bodies samples and the difference in localisation scores between NGF and NT-3. 482 transcripts significantly more localized in NGF as compared to NT-3 (blue dots; *Z*-score>1.96) and 348 transcripts significantly more localized in NT-3 compared to NGF (purple dots; *Z*-score<-1.96). *(J)* (*Top*) Genome browser view of a representative transcript with significant detection in the axonal compartment in NT-3-treated culture condition and residual detection in NGF condition (*Atf3*). (*Bottom*) eCLIP cross linking events along the gene for NT-3-specific pairs of positive regulators of axonal transport (DDX24:SLTM; EFTUD2:SRSF1, green) and negative pairs regulators of axonal transport specific to NGF or NT-3 (PUM2:HNRNPC and CELF4:CELF2, respectively, orange). *(K)* Lower levels of *Atf3* localized to NGF-treated axons (*upper*) compared to NT-3 treated axons (*lower*). Black arrows point to cell body signal, white arrowheads point to axonal signal. mRNA puncta were not subject to pixel dilation. (*Insets*) Magnification (2×) of the boxed areas. Scale bar, 10μm.

Figure 5. The RBP regulome underlying transcripts axonal localisation in NGF and NT-3. *(A)* Distribution of the significance in the difference between the mean localisation scores in NGF (*left*) and NT-3 (*right*) of the isoforms exhibiting or not a cross-link event in 50 nucleotide- long regions along the 3' UTR for 126 RBPs. Dark line displays the median and shaded areas indicate lower and upper quartiles. *(B)* Increasing localisation scores in NGF (*left*) and NT-3 (*right*) conditions as a function of the running average of detected number of cross-link events across the 332 CLIP-seq data in the [-150:-100] nucleotides upstream the 3' end. *(C)* Heatmap showing the significance in the difference between the mean localisation scores of the 3' UTR isoforms exhibiting or not a cross-link event in 50 nucleotide long regions along the 3' UTR for the top 32 RBPs positively associated with axonal localisation. *(D)* Scatterplot of the extent of significant association between FMR1 cross-link events and the axonal localisation along the 3' UTR (*top)*. Blue = NGF; purple = NT-3. **(***bottom***)** Barplots showing the fraction of [-250:-50] nucleotide regions upstream the 3' ends exhibiting FMR1 cross-link event in the full set of detected transcript (grey bar), and the pools of over- and under-transported transcripts in NGF (blue bars) and NT-3 (purple bars), respectively. Fisher enrichment test. *(E)* Logistic regression classifier performances trained

with different feature sets: all regulators (43 RBPs; M1), positive regulators only (32 RBPs; M2), negative regulators only (11 RBPs; M3) and interaction terms (pairs of RBPs; M4) for NGF dataset (*top*) and NT-3 dataset (*bottom*). *(F)* Protein-protein interactions for 16 predicted negative RBPs exhibiting significant enrichment in cross-link events in the [-250:-50] nt regions upstream the 3' end of the under-transported transcripts. Edges represent experimentally determined protein-protein interactions annotated in the STRING database. Nodes indicate proteins coloured according to biological pathways they are enriched in. (*inset*) Top 4 biological pathways and associated enrichment P-values as obtained from Fisher enrichment test. *(G, H)* Genome browser view of representative transcripts either restricted *(G)* or with significant detection in the axonal compartment *(H)* in NT-3- and NGF-treated culture conditions and CLIP crosslinking events for most negative (orange) and most positive (green) RBP regulators of axonal transport common to NGF and NT-3 conditions. *(I, J)* Barplot showing the significance in features importance to logistic regression classifier performance of pairs of RBPs that are specific to NGF *(I)* or NT-3 *(J)*. *(K-N)* Similar to *(G)* for transcripts over- and under-transported in NGF condition (K,M) and over- and under-transported in NT-3 condition *(L,N;* see also Supplemental Fig. S5J).

Figure 6. Negative regulators of APA in the cell body are candidate 3' UTR cleavage factors in the axons. *(A)* (*left*) Number of 3' UTR isoforms with proximal shift uniquely detected in axons in NGF (blue) and NT-3 (purple) culture condition. (*right*) Venn diagram showing the overlap between the candidate of axonal remodeling in NGF and NT-3. *(B,C)* Representative transcripts with a marked shift toward axonal increase in promoter-proximal 3' UTR uniquely detected either in axons treated with NGF (*Kif26b,* B) or in axons treated with NT-3 (*Parm1, C*). *(D)* Distribution of the extent of significant enrichment in 126 RBP cross-link events in defined regions along the 3' UTR of the 80 *(left)* and 60 *(right)* candidate isoforms of axonal remodeling in NGF and NT-3 condition, respectively. Dark lines display the median significance and shaded areas indicate lower and upper quartiles. *(E)* Number of candidate RBPs regulators of axonal remodeling identified in the [-250:+150] nt region surrounding the 3' end of the promoter-proximal isoform in both conditions, in NGF only and NT-3 only (grey, blue and purple bar, respectively). *(F)* Fraction of promoter-proximal 3' UTR isoforms exhibiting cross-link events for UPF1 in [0:50] nucleotide region down-stream of the cleavage site. Grey bar = all promoter-proximal 3' UTR; blue bars = 80 candidates of axonal remodeling in NGF; purple bars = 60 candidates of axonal remodeling in NT-3. *(G)* Fraction of candidate RBPs regulators of axonal remodeling in NGF and NT-3 at specific positions along the 3' UTR where they exhibit the most significant enrichment. Fisher enrichment test. *(H)* Same as *(G)* for AKAP8L and PTBP1. *(I)* (*upper)* Distribution of the extent of significant positive or negative association between the cross-link events in defined regions along the

148 short (*left*) and long (*right*) 3' UTR isoforms of 65 candidate regulators for axonal
149 remodeling, and their relative usage of the short and the long isoform in the cell body. Dark
150 lines display the median significance and shaded areas indicate lower and upper quartiles.
151 (*bottom*) Heatmaps showing the individual significance association between cross-link
152 events of individual RBPs along the short (*left*) or long (*right*) 3' UTR isoform and their
153 relative usage in the axons.
154
155 **Figure 7.** Proposed model where the local translation and neurotrophic-mediated axon
156 growth are linked to co-transcriptional APA, RNA localization, axonal 3' UTR remodeling
157 through release of mRNAs and cleavage factors from transport granules.

Figure 1

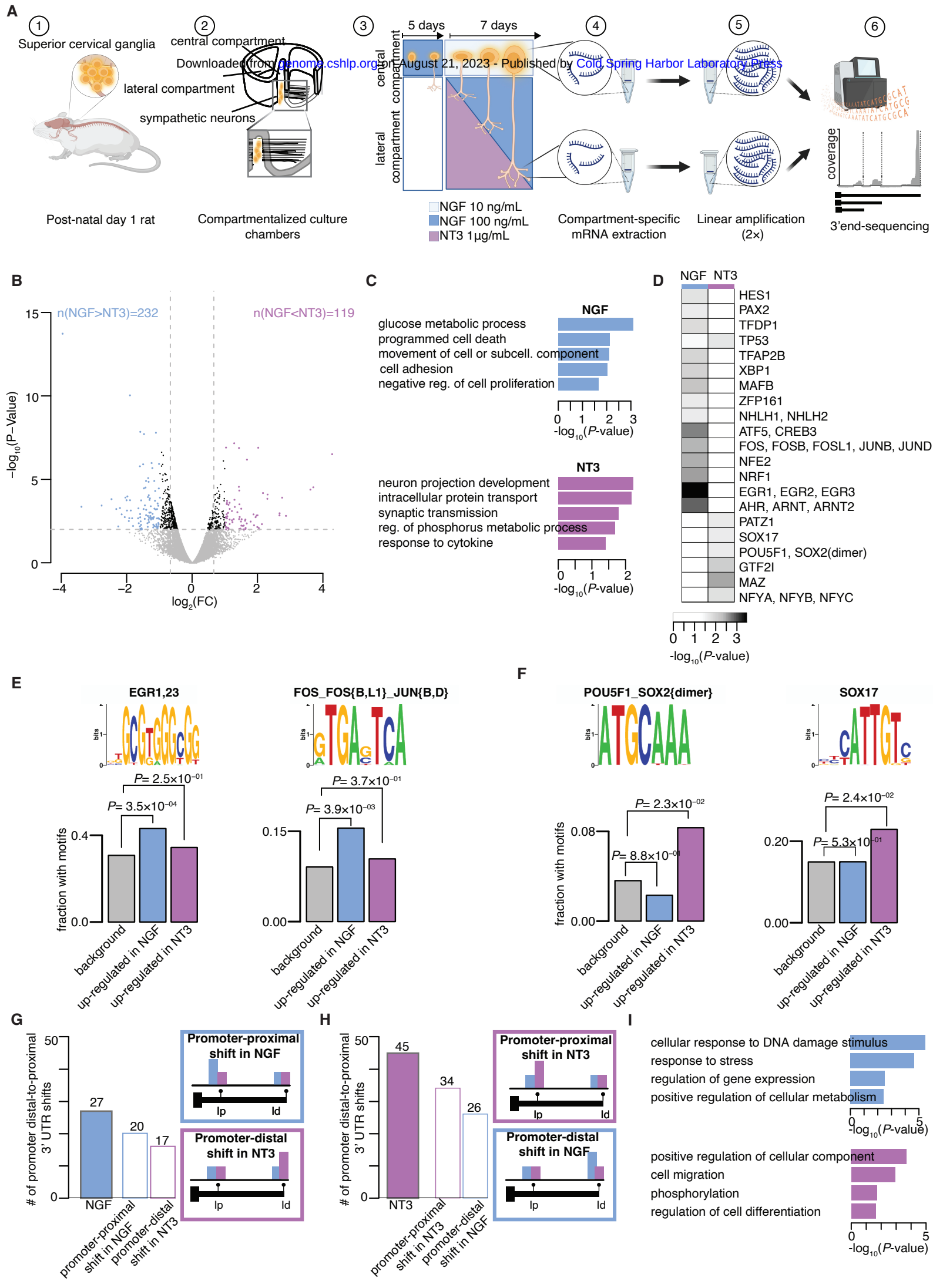


Figure 2

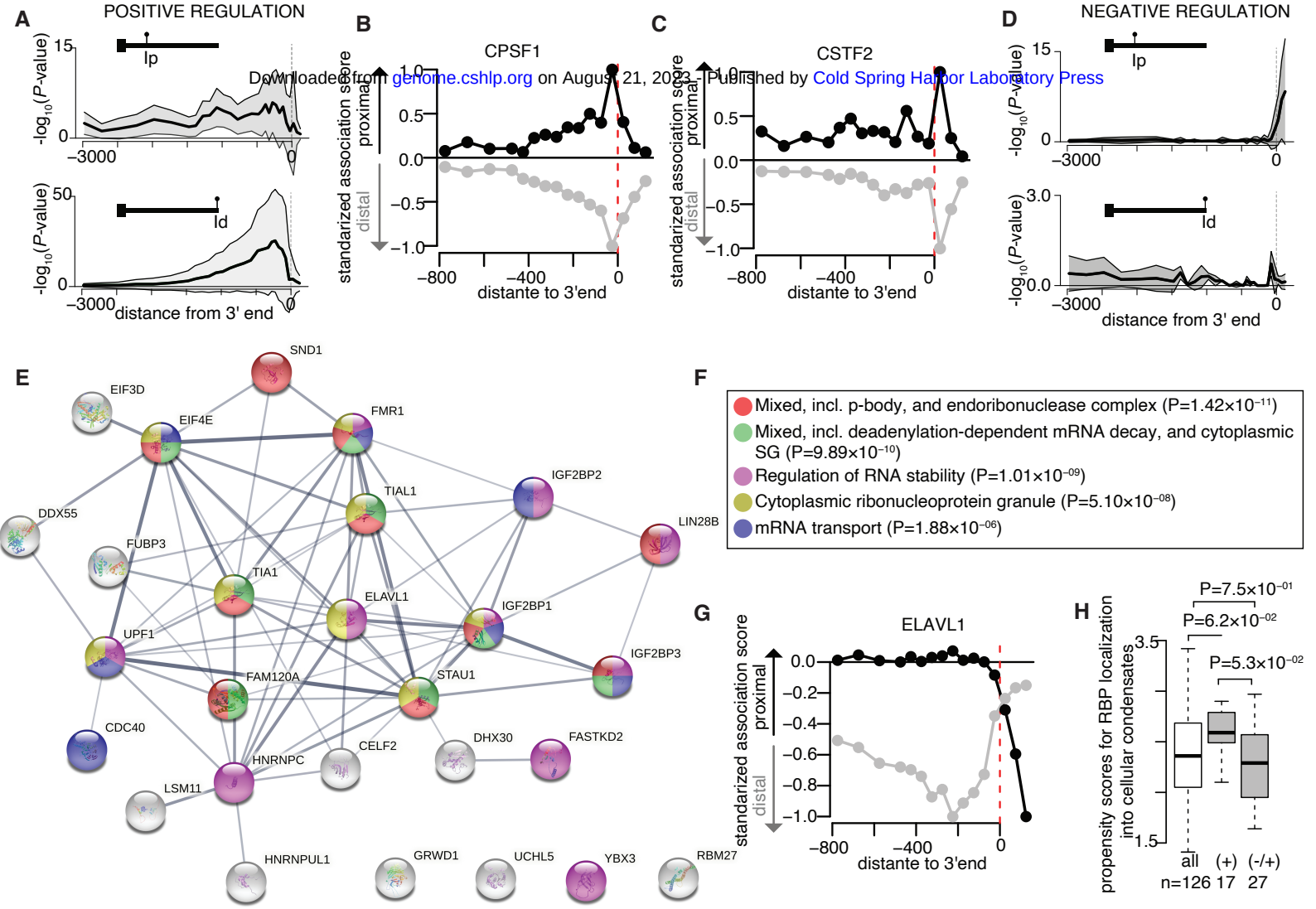


Figure 3

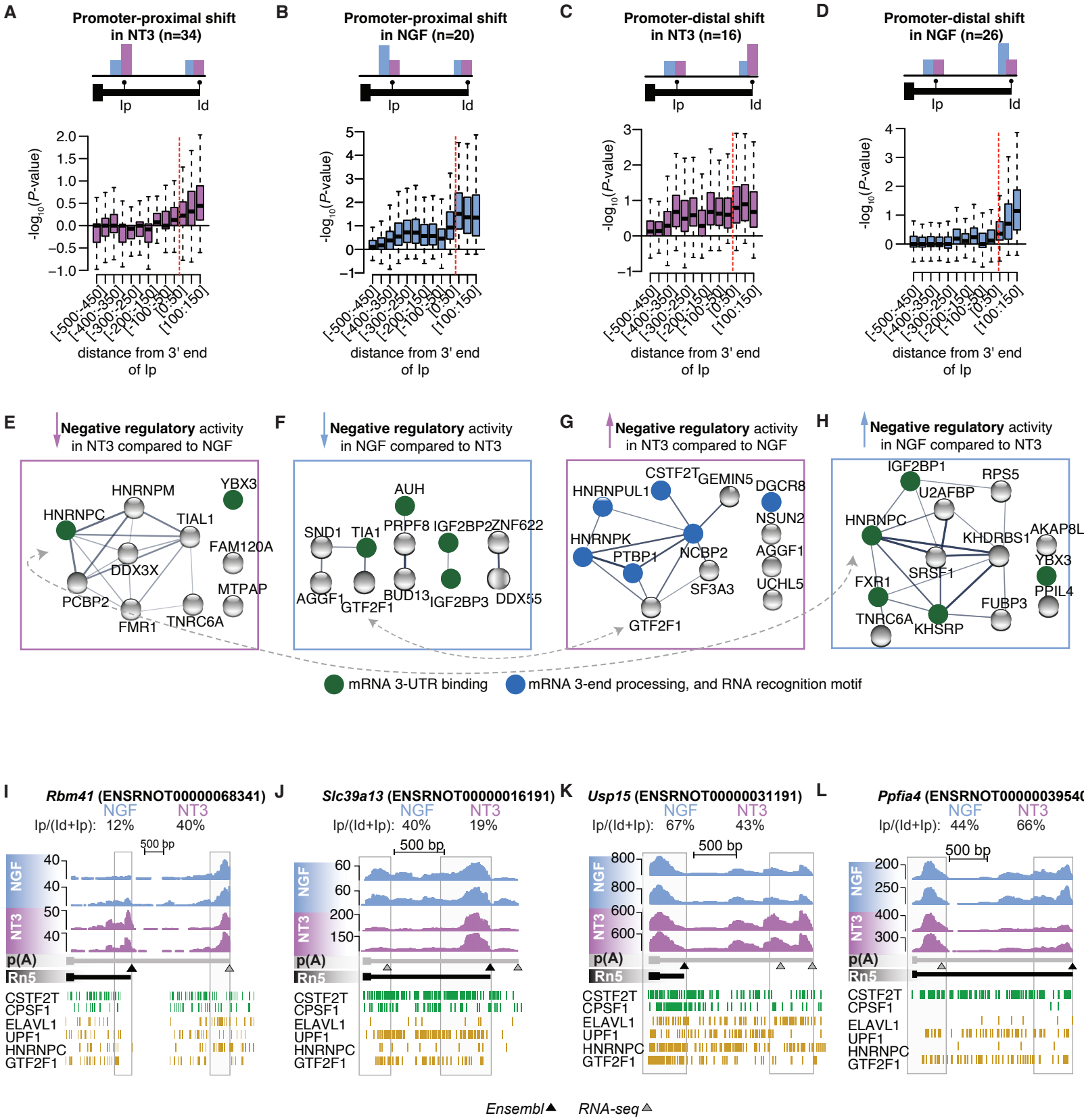


Figure 4

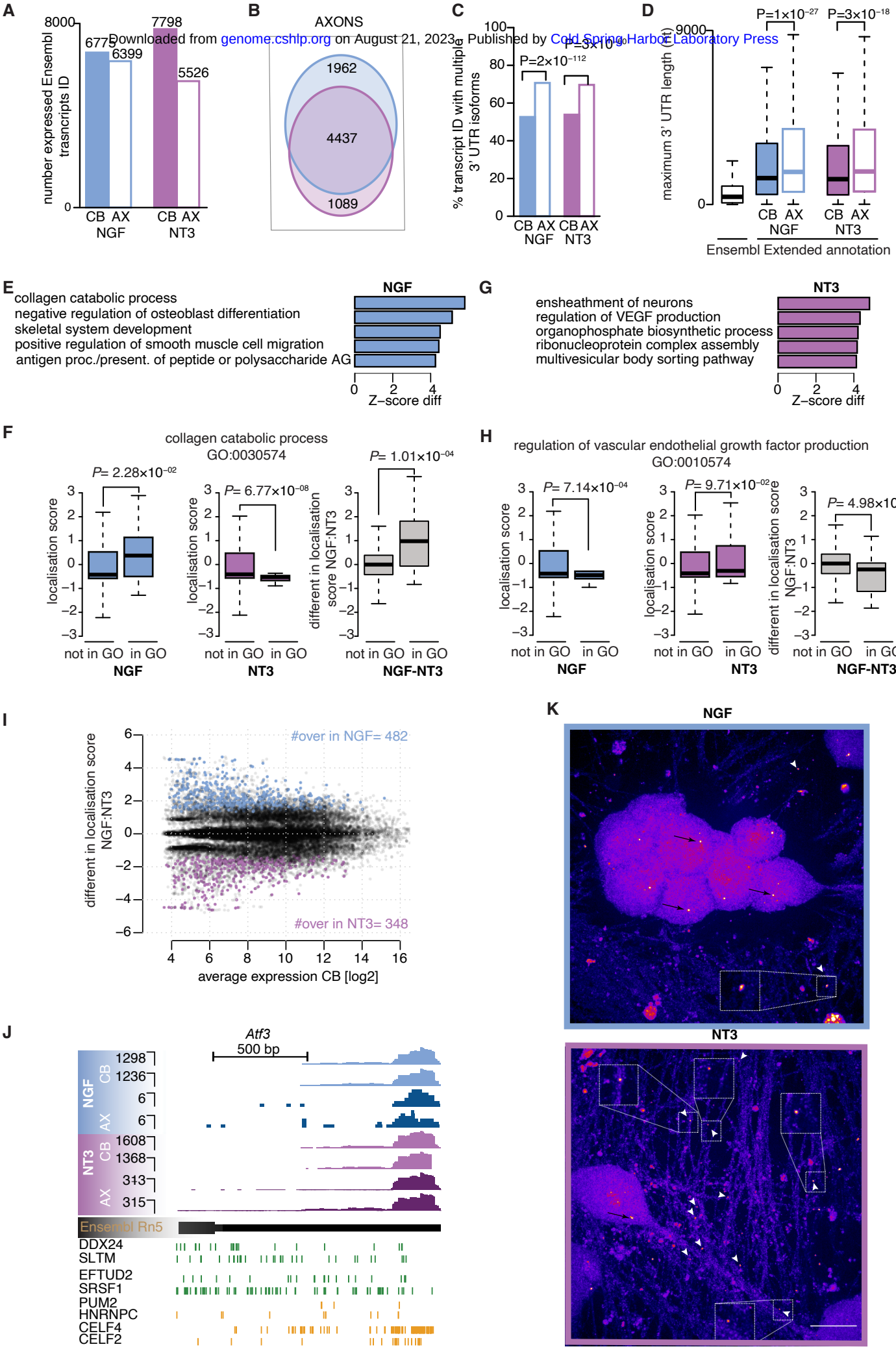
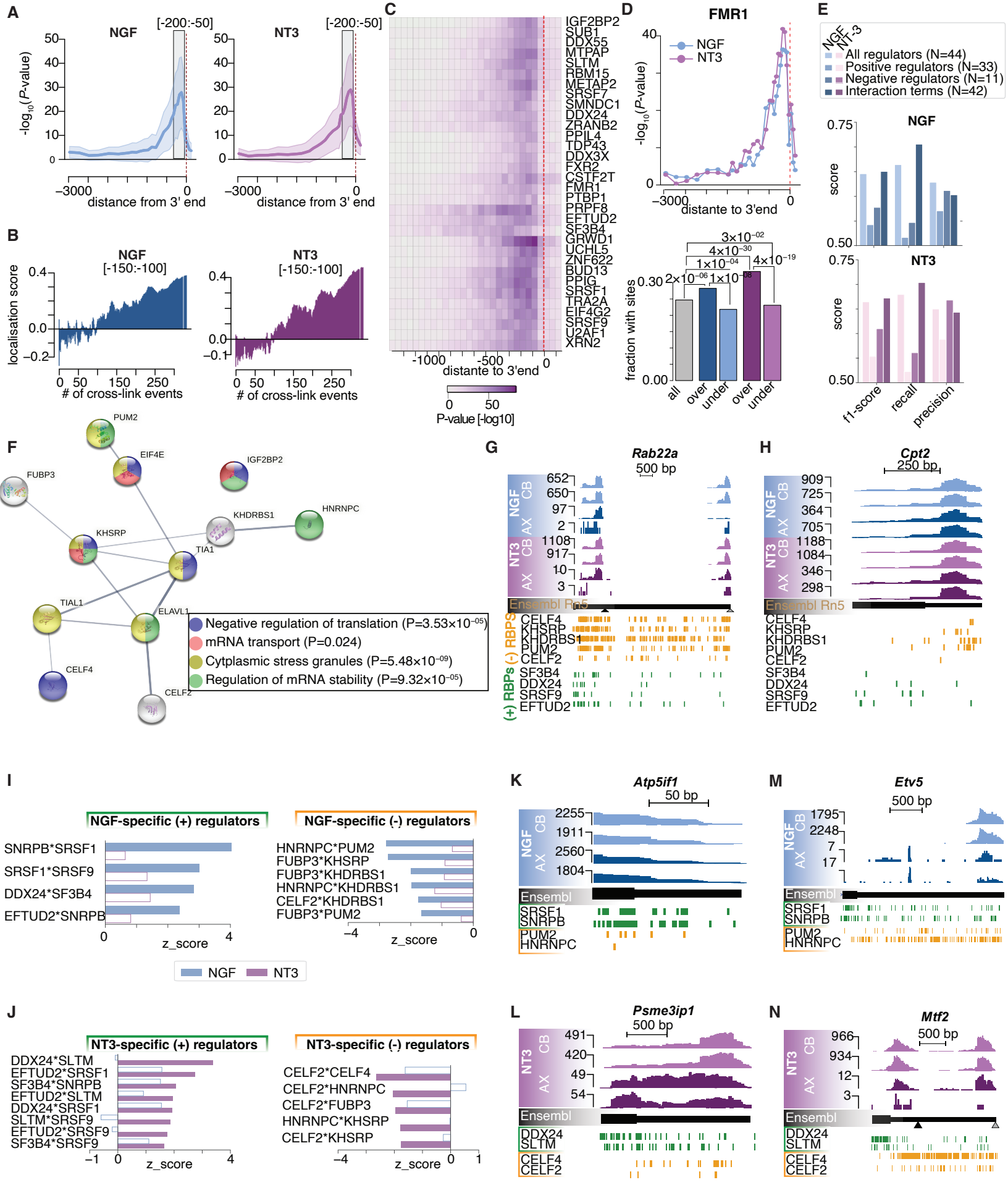


Figure 5



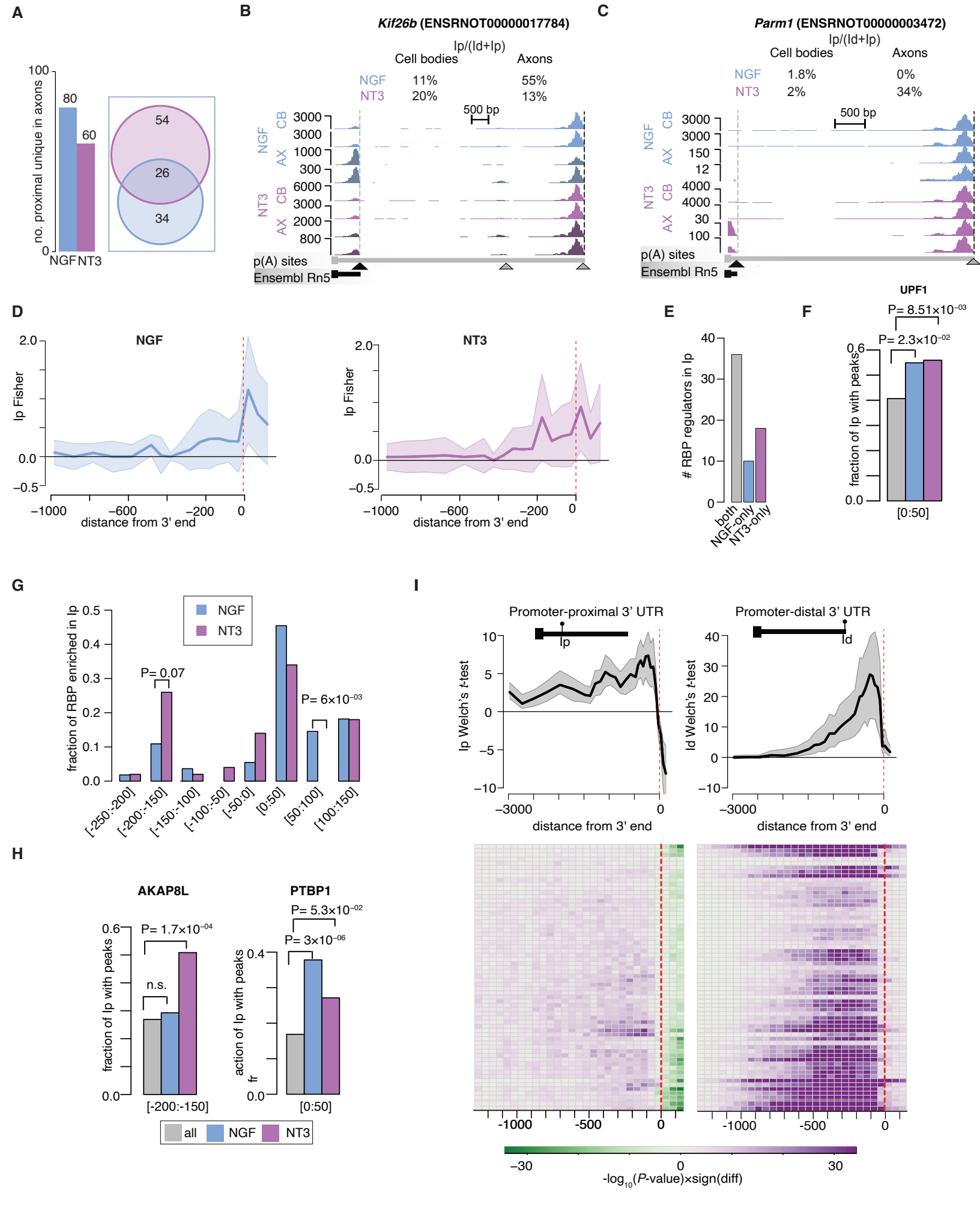
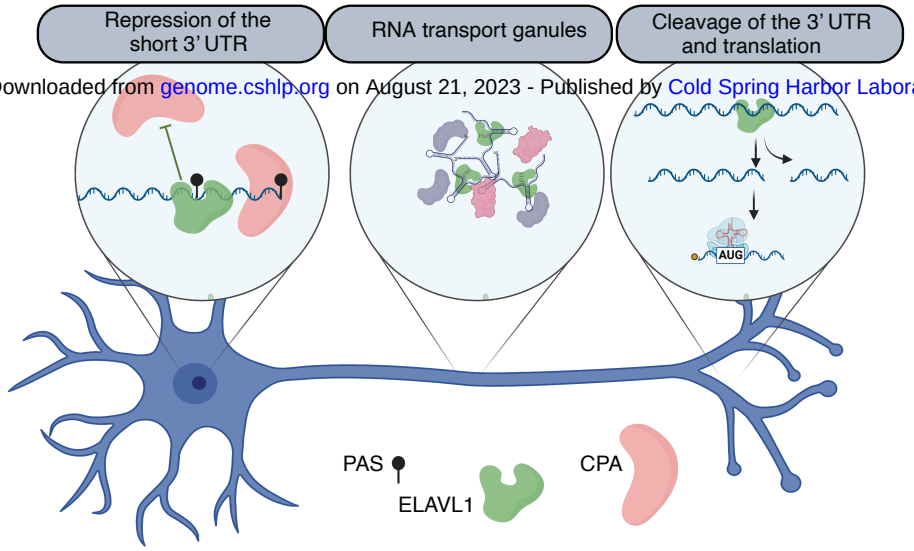


Figure 7



The predicted RNA-binding protein regulome of axonal mRNAs

Raphaelle Luisier, Catia Andreassi, Lisa Mathilde Fournier, et al.

Genome Res. published online August 15, 2023
Access the most recent version at doi:10.1101/gr.277804.123