The novel Rab5 effector FERRY links early endosomes with the translation machinery

J. S. Schuhmacher[1], S. tom Dieck[2], S. Christoforidis[3,4], C. Landerer[1], L. Hersemann[1], S. Seifert[1], A. Giner[1], A. Toth-Petroczy[1], Y. Kalaidzidis[1,5], E. M. Schuman[2], M. Zerial[1*].

[1]Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstrasse 108, 01307, Dresden, Germany.

[2]Max Planck Institute for Brain Research, Max-von-Laue-Str. 4, 60438 Frankfurt am Main, Germany.

[3]Institute of Molecular Biology and Biotechnology-Biomedical Research, Foundation for Research and Technology, 45110 Ioannina, Greece.

[4]Laboratory of Biological Chemistry, Department of Medicine, School of Health Sciences, University of Ioannina, 45110 Ioannina, Greece.

[5]Faculty of Bioengineering and Bioinformatics, Moscow State University, Moscow, Russia.

*Correspondence: zerial@mpi-cbg.de

Abstract

Local translation is vital to polarized cells such as neurons and requires a precise and robust distribution of different mRNAs and the translation machinery across the entire cell. The underlying mechanisms are poorly understood and important players are still to be identified. Here, we discovered a novel Rab5 effector complex which leads to mental retardation when genetically disrupted. The Five-subunit Endosomal Rab5 and RNA/ribosome intermediarY, FERRY complex localizes to early endosomes and associates with the translation machinery and a subset of mRNAs including mRNAs for mitochondrial proteins. It directly interacts with mRNA, thereby exhibiting different binding efficacies. Deletion of FERRY subunits reduces the endosomal localization of transcripts, indicating a role in mRNA distribution. Accordingly, FERRY-positive early endosomes harboring mRNA encoding mitochondrial proteins were observed in close proximity to mitochondria in neurons. Therefore, the FERRY complex plays a role for mRNA localization by linking early endosomes with the translation machinery.

Introduction

Subcellular mRNA localization is a widespread phenomenon in biology. The correct positioning of mRNA transcripts is vital for fundamental biological processes comprising embryonic development, cellular homeostasis, neuronal plasticity and adaptive response to environmental cues (reviewed in: (Cioni et al., 2018; Das et al., 2021; Glock et al., 2017; Martin and Ephrussi, 2009; Turner-Bridger et al., 2020)). A prime example is the asymmetric localization of specific mRNAs during oogenesis that determines the body axes and patterns of the future embryo (reviewed in: (Becalska and Gavis, 2009; Riechmann and Ephrussi, 2001)). While the oocyte represents a morphologically rather simple example, a completely different scenario unfolds in the brain, where neurons are spanning long distances with their axonal and dendritic processes. Not only are these compartments highly specialized in their function but they also respond to external cues on a millisecond timescale at the distal end of their network, far away from the cell body. Neurons handle these challenges by producing many proteins locally at their site of action, in a process called local translation, which is involved in axon outgrowth, branching synaptogenesis, regeneration and neuronal plasticity (Cioni et al., 2018; Jung et al., 2014; Kim and Jung, 2020; Rangaraju et al., 2017). This however requires the availability of the respective mRNAs at the sites of local translation, and hence the precise subcellular localization of a plethora of mRNAs (Glock et al., 2017; Turner-Bridger et al., 2020).

Transcriptomic studies have identified thousands of different mRNAs in neuronal sub-compartments, such as axons, dendrites or the neuropil (Andreassi et al., 2010; Briese et al., 2016; Cajigas et al., 2012). Furthermore, these transcripts are distributed heterogeneously, with mRNAs showing distinct localization patterns, for example, being restricted to axons or dendrites or even smaller sub-compartments. While mRNAs with synaptic function are enriched in the somatodendritic region, in axons mRNAs for proteins connected to translation, the cytoskeleton as well as mitochondrial proteins are significantly enriched compared to the somatodendritic region (Andreassi et al., 2010; Briese et al., 2016). These findings reveal a complex mRNA distribution plan, where thousands of mRNAs have to find their correct location.

Such a complex task requires active mRNA transport along the cytoskeleton. A direct connection between RNA binding proteins (RBPs) and motors proteins has been observed in various forms, for example the targeting of mRNAs by RBPs which recognize *cis*-regulatory elements on the respective mRNA, including the so called ‘zipcoes’ (reviewed in: (Buxbaum et al., 2015; Das et al., 2021)). Recently, different compartments of the endolysosomal system have been associated in situ with the spatial organization of components of the translation machinery, including mRNA, mRNP granules as well as ribosomes (Cioni et al., 2019; Higuchi et al., 2014; Liao et al., 2019). In general, the endolysosomal system is the central logistic system of eukaryotic cells, comprising multiple membrane-enclosed organelles, such as early and late endosomes as well as the lysosome, is responsible for the trafficking and sorting of a

large variety of cargos, including membrane receptors, lipids, extracellular fluids and signaling proteins. Within this network, the early endosome is the first station for cargo coming from the plasma membrane, that is subsequently sorted either into the recycling pathway via recycling endosomes, or into the degradation pathway, via late endosomes, multivesicular bodies and the lysosomes. The identity of endosomes is determined by an intricate interplay between their unique protein residents and specific lipids which are intimately linked to Rab GTPases (Cezanne et al., 2020; Pfeffer, 2013). Along the endosomal pathway, different Rab GTPases characterize different organelles, such as Rab4 and Rab11 for recycling endosomes and Rab7 for late endosomes (reviewed in: (Wandinger-Ness and Zerial, 2014)). Rab5 is the hallmark GTPase of the early endosome and a molecular switch with two distinct states, a GDP bound or nucleotide free, inactive or a GTP loaded, active state. On the early endosome Rab5 gets activated by a Guanosine exchange factor (GEF) and is then able to recruit a plethora of Rab5 effector proteins, such as the molecular tether EEA1 or Rabankyrin-5, thereby not only regulating the lipid content of the early endosome, but also orchestrating different functionalities of the organelle, by forming distinct domains on the membrane (Cezanne et al., 2020; Franke et al., 2019; Lauer et al., 2019; Lippe et al., 2001; Murray et al., 2016; Schnatwinkel et al., 2004; Zhang et al., 2012).

As a well-established multifunctional transport system, the endosomal system is ideally suited to regulate mRNA localization, especially in morphologically complex compartments like the hyphae of fungi or the processes of neurons. In the fungus *U. maydis*, a special adaptor system enables the long-distance travel of mRNA and polysomes on early endosomes (Higuchi et al., 2014). In higher eukaryotes lysosomes and late endosomes are involved in RNA transport. While lysosomes serve as an Annexin A11-mediated mRNP granule transport vehicle, late endosomes were identified as translation platforms for mitochondrial proteins in neurons (Cioni et al., 2019; Liao et al., 2019). A recent study showed the co-localization of mRNA with early endosomes, indicating that also early endosomes might be part of an mRNA distribution machinery (Popovic et al., 2020). While late endosomal motility is primarily retrograde, early endosomes show bidirectional motility in neurons (Goto-Silva et al., 2019). Therefore, early endosomes appear more suitable to support directional mRNA transport than late endosomes. However, a molecular mechanism describing the connection between early endosomes and mRNA or the entire translation machinery is still missing. To date, none of the mRNA-associated proteins appears to localize on early endosomes nor do no endosomal proteins have RNA binding motifs. The identification of a molecular connection between early endosomes and the translation machinery is especially crucial as it sets the basis for addressing the questions about the transcript specificity of mRNA localization and the distribution to the correct location. Considering the large number of transcripts that need to be precisely localized, one would envision a molecular machinery that transports specific mRNA subsets, discriminates between different mRNAs and thereby using a limited number of vesicular carriers.

Closing this gap, we report the discovery of a novel five-subunit Rab5 effector complex, which we named Five-subunit Endosomal Rab5 and RNA/ribosome intermediarY (FERRY) complex. Through direct interaction with Rab5 and mRNA, it connects the early endosome with the translation machinery. Furthermore, the FERRY complex is able to bind specific transcripts with high efficacy including mRNAs of mitochondrial proteins for fundamental processes, such as the respiratory chain, the TCA cycle and the mitochondrial protein synthesis.

Results

Identification of a novel Rab5 effector complex

In previous studies we isolated the complete set of Rab5 effectors from bovine brain cytosol, using a Rab5 affinity chromatography (Christoforidis et al., 1999). Upon further purification of this intricate set of proteins, we were surprised to observe that as many as five proteins co-fractionated in size exclusion chromatography, in fractions 22 to 25 (Figure 1A, left panel). To purify further these effectors, the above fractions were pooled and subjected to anion exchange chromatography, which, in comparison to size exclusion chromatography, separates proteins on a different principle (*i.e.* based on ionic charges). Interestingly, the same set of five proteins co-eluted from the anion exchange column in fractions 41 to 43 (Figure 1A, right panel). The co-fractionation of these proteins, both in size exclusion and anion exchange chromatography, indicated that they exist in a complex, which raised a great interest regarding its identity and function. Mass spectrometry revealed the five proteins as Tbck (101 kDa), Ppp1r21 (88 kDa), C12orf4 (64 kDa), Cryzl1 (39 kDa) and Gatd1 (23 kDa) (Figure 1B). For 3 of the 5 proteins human gene mutations have been reported (Beck-Wodl et al., 2018; Bhoj et al., 2016; Chong et al., 2016; Guerreiro et al., 2016; Hancarova et al., 2019; Loddo et al., 2020; Ortiz-Gonzalez et al., 2018; Philips et al., 2017; Rehman et al., 2019; Suleiman et al., 2018; Zapata-Aldana et al., 2019). For clarity, we will refer to the novel complex as the Five-subunit Endosomal Rab5 and RNA/ribosome intermediarY (FERRY) complex, with the individual subunits being designated Fy-1 – Fy-5 (Figure 1B).

To show that the five co-fractionating proteins form a complex, we co-expressed Fy-1 – Fy-3 in baculovirus-infected insect cells and incubated the lysate with individually purified Fy-4 and Fy-5 to reconstitute the FERRY complex *in vitro* (see Methods: Protein purification). Figure 1C shows that we obtained a stable complex comprising all five proteins eluting as a monodisperse peak from SEC. In order to estimate the stoichiometry of the components in the FERRY complex, we compared the intensity of the corresponding signals of a Coomassie stained SDS PAGE, which suggested a ratio of 1:2:1:2:4 for Fy-1:Fy-2:Fy-3:Fy-4:Fy-5, respectively. Using mass photometry, we obtained a molecular weight of 525 ± 41 kDa for the fully assembled complex which fits very well with the estimated ratios and a calculated molecular weight of 521 kDa. This was further corroborated by a cryoEM structure which showed a ratio of 2:2:4 for Fy-2, Fy-4 and Fy-5 (Quentin et al., 2021). With the FERRY

149 complex in hand, we tested whether it fulfills the typical criterion of Rab5 effectors and binds
150 predominantly to the activated, GTP-loaded Rab5, by performing a Glutathione-S-transferase
151 (GST) pulldown assay. The FERRY complex bound much stronger to Rab5:GTP than
152 Rab5:GDP, while no interaction was observed with GST, (Figure 1D).

153 We next validated our findings and examined the specificity of the Rab GTPase interaction.
154 We *in vitro* translated Fy-1 to Fy-5 incorporating ^{35}S methionine and performed *in vitro* binding
155 assays against different Rab GTPases of the endosomal system, including Rab5, Rab4 and
156 Rab11 (recycling endosome) and Rab7 (late endosome) (Figure 1E). In this experimental set
157 up, the binding of each component of the complex was tested individually, in the absence of
158 the other subunits, thereby allowing identification of the subunit(s) of the complex that mediate
159 binding between the FERRY complex and Rab5. Out of the five subunits, only Fy-2 bound to
160 Rab5:GTP, but not Rab5:GDP (Figure 1E). These results indicate that Fy-2 mediates the
161 interaction between the FERRY complex and Rab5:GTP. This was also confirmed by hydrogen
162 deuterium exchange mass spectrometry (HDX-MS), which identified the Rab5 binding site of
163 the FERRY complex near the C-terminus of Fy-2 (Quentin et al., 2021). In addition, no
164 interaction was observed between the FERRY complex and the other Rab GTPases, neither in
165 the GDP- nor GTP-bound form. These results indicate that the FERRY complex is indeed a
166 Rab5 effector complex and very likely localizes on Rab5-positive and thus early endosomes.
167 To validate this prediction, we raised an antibody against Fy-2 which is suitable for
168 immunofluorescence (Figure S1, see also Methods: Antibody validation). The fluorescence
169 signal revealed a punctate localization pattern in HeLa cells, that matches the localization of
170 the early endosomal marker EEA1 (Figure 1F). This finding shows that the FERRY complex
171 localizes to early endosomes.

172 The properties of the five FERRY subunits exhibit a substantial variability in size, domain
173 composition and structural features. Indeed, the FERRY complex does not resemble any known
174 endosomal complexes (*e.g.* HOPS, CORVET, or the ESCRT complexes) (Figure 1B).
175 Searching for traces of the FERRY complex in the course of evolution, we performed a
176 phylogenetic analysis of the subunits of the FERRY complex. While Fy-1 is the most ancestral
177 subunit with homologues in some fungi, we also found an assembly of Fy-1, Fy-3 and a shorter
178 version of Fy-2 in insects and some nematodes, that lacks the Fy-4 and the Fy-5 binding sites
179 that were mapped based on structural information (Quentin et al., 2021). With the evolution of
180 the Chordata, we observed a transition from the reduced 3 component assembly to the five-
181 subunit complex, via the co-occurrence of two novel proteins, Fy-4 and Fy-5 and the extension
182 of Fy-2 with the Fy-4 and Fy-5 binding sites (Figure 1G and Figure S2). This co-evolution
183 further supports that the FERRY subunits form a stable protein complex that is evolutionary
184 conserved.

185

The FERRY complex associates with the translation machinery

Even though the FERRY complex has not previously been identified, it may play an important role in brain function. Clinical studies on patients, with a mutation in the *fy-1 (tbck)* or *fy-2 (ppp1r21)* gene, show that loss of either of these proteins severely impairs brain development and function, causing symptoms such as a mental retardation, intellectual disability, hypotonia, epilepsy, and dysmorphic facial features resulting in a premature death of the patients (Bhoj et al., 2016; Chong et al., 2016; Guerreiro et al., 2016; Hancarova et al., 2019; Loddo et al., 2020; Ortiz-Gonzalez et al., 2018; Philips et al., 2017; Suleiman et al., 2018; Zapata-Aldana et al., 2019). Different studies report the accumulation of lipofuscin the human brain and further indicate disturbances in the endocytic system (Beck-Wodl et al., 2018; Rehman et al., 2019). These results suggest that FERRY complex carries out an endocytic function which is essential for brain development and neuronal function.

To get insights into the cellular role of the FERRY complex, we examined the interaction network of the novel complex using a GST pulldown approach (Figure 2A). In a first step, we generated a GST fusion variant of the FERRY complex (GST-FERRY). As observed for the native complex, GST-FERRY eluted from SEC as a monodispersed peak yielding pure complex (Figure 2B). Subsequently, GST-FERRY was immobilized on resin, incubated with fresh HEK 293 cell detergent lysate (see Methods: HEK 293 lysate preparation), stringently washed and eluted from the resin (Figure 2A). Mass spectrometry of the elution fractions revealed 34 proteins as potential interaction partners of the FERRY complex (Figure 2C, Table S1). Almost Three-quarters of the candidates (73.5 %) represent ribosomal proteins of the both the large and the small subunit (Figure 2D), which suggests that complete ribosomes and hence the translation machinery were pulled down by the FERRY complex. These results provide evidence that the FERRY complex may associate with ribosomes, and thereby link the translation machinery with the endosomal system. However, ribosomal and mitochondrial proteins could also be considered contaminants in our biochemical assay.

To test for the specificity of the above interactions we asked whether RNAs accompany the ribosomes and RNA-binding proteins as FERRY interactors. To identify transcripts co-eluting with the FERRY complex, we modified the protocol of the GST-FERRY pulldown assay to obtain RNA, instead of proteins, which was subsequently analyzed by sequencing (Figure 2A). The RNA sequencing reads were then mapped against the human genome identifying more than 17 000 different mRNAs. A comparison of the FERRY complex with a GST control, applying a stringent cut-off (adjusted p-value < 0.01), provided 252 mRNAs significantly associated with the FERRY complex (Figure 2E, Table S1). Among these candidates, the largest group of mRNAs (66 transcripts/ 26.2 %) constitute nuclear-encoded mitochondrial proteins (Figure 2F). Furthermore, we also identified 13 components (5.2 %) of the endosomal system, for example *vps8* mRNA. As part of this group, we also identified *fy-1* and *fy-3* mRNA, which might suggest that the FERRY complex can associate with mRNAs of its subunits. A third group of transcripts was classified as nucleosome components (22 transcripts/ 8.7 %).

Even though we used the lysate of HEK 293 cells, we found transcripts (12/ 4.8 %) that are either enriched (*e.g. begain* mRNA) (Deguchi et al., 1998) or play an important role in the brain (*e.g. pafah1b3* mRNA) (Nothwang et al., 2001). To further characterize the mRNA candidates, we performed an enrichment analysis against a gene set collection (MSigDB C5 collection: ontology gene sets). Among others, various gene sets connected to mitochondria were significantly enriched (adjusted p-value < 0.01), including mitochondrial matrix genes (gene set #1714), genes connected to the mitochondrial ribosome (gene set #2354), to cellular respiration (gene set #480) as well as to the tricarboxylic acid cycle (gene set #4413). In summary, these results suggest that the FERRY complex interacts with specific mRNAs, especially those encoding mitochondrial proteins. The molecular nature of the interaction between the FERRY complex and the translation machinery however, cannot be derived solely from such an assay.

The FERRY complex interacts directly and selectively with mRNA

If mRNA is a critical link between the FERRY complex and the mitochondrial translation machinery, then we should demonstrate that the FERRY complex binds directly to specific mRNAs. To test this hypothesis, we performed an electrophoretic mobility shift assay (EMSA) with *in vitro* transcribed mRNAs and the FERRY complex. We chose *mrpl41* a top candidate of the RNA screen and included the 5' UTR, the coding region, the 3' UTR and a short stretch of 50 adenines, yielding a 660-nucleotide, artificially poly-adenylated mRNA. With increasing amounts of FERRY complex an additional signal at a higher molecular weight appeared in the EMSA, indicating a direct interaction between the FERRY complex and *mrpl41* mRNA (Figure 3A).

We next tested whether the RNA binding to the FERRY complex is Rab5-dependent. We performed the EMSA with a fixed FERRY/RNA ratio and added increasing amounts of Rab5:GTPγS to the assay. This did not have a visible effect on the FERRY-RNA interaction (Figure 3B). We further tested whether individual FERRY subunits are sufficient to bind mRNA or if the full FERRY complex is required. To do so, we compared the FERRY complex with its subunits Fy-4 and Fy-5, in the ratios that were observed in the CryoEM structure (Quentin et al., 2021). However, the results showed that neither of the two small subunits is able to interact with *mrpl41* (Figure S3).

The enrichment of specific groups of mRNAs in the RNA screen point towards the ability of the FERRY complex to discriminate between different mRNAs. To examine the specificity of mRNA binding we chose 8 mRNAs out (of the 252 found in the screen) comprising different mitochondrial functionalities, such as the respiratory chain (*cox6b* and *cox8a*), the ATP Synthase (*atp5f1b*), the mitochondrial stress response (*gstp1* and *prdx5*), the mitochondrial ribosome (*mrpl41*), the TCA cycle (*mdh2*) and the mitochondrial ubiquitination machinery (*uchl1*), and tested the direct interaction with the FERRY complex using an EMSA. While

mrpl41, *mdh2* and *atp5f1b* exhibited a strong interaction with the FERRY complex, the interaction with the other five candidates was much weaker (Figure 3C). These results demonstrate that the FERRY complex binds transcripts with different efficacy *in vitro*.

The FERRY complex influences mRNA localization in HeLa cells

The biochemical data suggest that FERRY mediates the association of mRNA or the translation machinery with early endosomes. To test this prediction, we designed an experiment to compare the localization of early endosomes (marked by EEA1), the FERRY complex (Fy-2) and different mRNAs, (*i.e. mrpl41*, *mdh2* and *atp5f1b*) upon knock-out of different FERRY subunits in HeLa cells (Figure 4A). The mRNA candidates that show a clear interaction with the FERRY complex *in vitro* binding assay were chosen. Furthermore, the mRNA candidates also represent different important mitochondrial pathways or functionalities, such as the mitochondrial ribosome (*mrpl41*), the TCA cycle (*mdh2*) and the respiratory chain (*atp5f1b*). In this assay, mRNA localization was determined by single molecule fluorescence *in situ* hybridization (smFISH). In a second step, using the CRISPR/Cas9 technology, we generated *fy-2*, *fy-4* and *fy-5* knock-out HeLa cell lines and confirmed the loss of the respective protein by Western blot (Figure S4). Immunostaining of Fy-2 in the *fy-2* knock-out cell line showed a strong reduction but no complete loss of fluorescent signal (Figure S1). Since the knock-out cell lines were generated by indel formations (see Methods: Generation of Generation of HeLa knockout (KO) cell lines), the remaining signal might either be caused by residual Fy-2 protein in the cells or by the recognition of an additional protein by the antibody. To avoid misinterpretations, we excluded the *fy-2* KO cell line from the experiment and used the signal from the antibody against EEA1 for image analysis and quantification (also see Methods: Antibody validation).

Using automated microscopy, we acquired images visualizing Fy-2, EEA1 and the mRNA candidates in HeLa wildtype and the *fy-5* and *fy-4* KO cell lines. As seen before, we observed a strong co-localization between EEA1 and the Fy-2 in the wildtype, but also in the KO cell lines (Figure 4B, Figure S5A, B). This indicates that the co-localization of Fy-2 and EEA1, does not require Fy-4 or Fy-5. In wildtype HeLa cells we observed 10.2 %, 7.1 % and 10.1 % co-localization of mRNA and EEA1 positive early endosomes for *mdh2*, *atp5f1b* and *mrpl41*, respectively. We often additionally observe the presence of fluorescence signal for Fy-2 at these co-localization events (Figure 4B boxes, Figure S5A, B boxes). While the loss of Fy-4 had no measurable effect on the co-localization of the mRNAs with early endosomes, a significant decrease in co-localization was observed upon knock-out of *fy-5* (Figure 4C-F). The loss of Fy-5 caused a decrease in co-localization of 27 %, 25 % and 20 % for *atp5f1b*, *mdh2* and *mrpl41* mRNAs, respectively (Figure 4F). This clearly indicates that Fy-5 but not Fy-4 affects the ability of the FERRY complex to interact with mRNA. This is in agreement with biochemical and structural data, showing that Fy-5 is an integral part of one of the interaction sites of the FERRY complex with *mrpl41* mRNA. The fact that the interaction of the FERRY

complex with mRNA comprises two main interfaces, involving Fy-1, Fy-2 and Fy-5, may explain why a loss of Fy-5 decreases mRNA early endosome co-localization but does not abolish it (Quentin et al., 2021). The observation that a loss of Fy-4 does not affect mRNA-early endosome co-localization might be explained by its location right at the center of the core particle of the FERRY complex embraced by a Fy-2 dimer. Cross-linking experiments also showed that Fy-4 is not directly involved in mRNA binding (Quentin et al., 2021).

In summary, we found that the FERRY complex contributes to the localization of specific mRNAs encoding mitochondrial proteins to early endosomes. The FERRY-mediated connection between the endosomal system and the translation machinery, might generate an mRNA transport platform, that seems to be crucial for morphologically complex cell types.

The FERRY complex localizes to axons as well as to the somatodendritic region

With their long processes, neurons are a prime example of morphologically complex cells. Furthermore, the genetic loss of certain FERRY subunits has major impact on brain development and function. Therefore, we assessed the localization of the FERRY complex in primary rat hippocampal neurons. To determine its distribution, we compared the FERRY localization to the endosomal markers EEA1 and Rabankyrin-5. EEA1 and Rabankyrin-5 differ in their localization in neurons, since EEA1 is restricted to the somatodendritic region (Wilson et al., 2000), while Rabankyrin-5 is also found in axons (Goto-Silva et al., 2019). Immunofluorescence staining of Fy-2 revealed a punctate pattern of fluorescent foci dispersed across the neuron (Figure 5A, overview), as also observed in HeLa cells (Figure 1F). Indeed, the fluorescent signal strongly co-localized with the endosomal markers EEA1 and Rabankyrin-5. We observed many triple positive (Fy-2, EEA1, Rabankyrin-5) endosomes (Figure 5A, details, white arrowheads), but also fluorescent foci that were only positive for Fy-2 and Rabankyrin-5, mainly in thin structures, where EEA1 signal was absent (Figure 5A, blue, yellow arrowheads). These results suggest that the FERRY complex is present in both the somatodendritic region as well as axons.

In order to validate this hypothesis, we performed immunofluorescence against Map2 and the phosphorylated neurofilament-1 (pNF) as markers of the somatodendritic region and axons, respectively (Figure 5B, overview). As our previous experiments suggested, we again observed Fy-2 and Rabankyrin-5 positive early endosomes in thin structures positive for the axonal marker pNF (Figure 5B, box). In summary, the FERRY complex resides on early endosomes, distributed across the neuronal soma, dendrites and axons, immediately raising the question about possible mRNA localization on these endosomes.

The FERRY complex co-localizes with mRNA on early endosomes in neurons

In order to address the question about the RNA load of FERRY positive endosomes in neurons, we again combined immunofluorescence with smFISH. To visualize total mRNA distribution in the cell, we used an *in-situ* hybridization probe directed against the polyA tail of RNA. We focused our imaging on dendrites and axons of the neuron excluding the soma, since the cell body has a high protein and mRNA density which may impede quantification. Furthermore, we were especially interested in the distribution of mRNA in distal regions. While the mRNA density in major dendrites is still high, it decreases in thinner processes and forms clusters at nodes. Overall, we observed that 6.1 % of mRNA foci co-localize with the FERRY complex (Figure 5C). Often, these events also co-localize with EEA1, suggesting that mRNA is located on early endosomes (Figure 5C, light blue box). In other cases, a larger endosome is surrounded by several mRNA foci, with the fluorescent signals being in close proximity rather than co-localizing (Figure 5C, white box). Taking into account the molecular dimensions of the FERRY complex, mRNAs and the labelling methods we used, we estimated that the fluorescent signals of the FERRY complex and the mRNA can have a distance of around 250 nm or more and still represent a mRNA-FERRY complex (Figure S6). Our findings corroborate the notion that early endosomes are active players in the organization and transport of mRNAs in neurons.

We next tested the co-localization of the FERRY complex with specific transcripts in neurons choosing *mdh2* and *uchl1* mRNA based on the initial mRNA binding screen (Figure 2E) and the Fy-5-dependent co-localization with early endosomes of *mdh2* (Figure 4F). Compared to the fluorescent signal of the entire mRNA population, the signal for individual mRNAs was weak and we observed only scarce co-localization with the FERRY complex. More often the fluorescent signals were in close proximity, rather than overlapping (Figure 5C-E). Given the methodological circumstances (Figure S6), we assumed that fluorescence signals of FERRY and mRNA within a range of 250 nm still represent a FERRY-mRNA complex (Figure 5D, E boxes). Counting the number of events, we found 13.2 % of *mdh2* transcripts and 10.3 % of *uchl1* mRNAs in contact with the FERRY complex. A proper quantification is impeded by a substantial variability between cells and a different transcript density in the neurons.

The interaction between the FERRY complex and different transcripts encoding mitochondrial proteins suggests that FERRY-positive early endosomes loaded with mRNA destined for mitochondria might be observed on mitochondria. To examine this, we additionally stained neurons with TOMM70 as a marker for mitochondria. When visualizing the entire mRNA population, we regularly found co-localization of the FERRY complex with mRNA on mitochondria (Figure 6A). Since mRNA and mitochondria are quite abundant in neurons, it is difficult to estimate how specific this co-localization is. Therefore, we also assessed the co-localization of the FERRY complex with *mdh2* mRNA and mitochondria (Figure 6B). Even though these events were infrequent, we indeed observed examples where the fluorescence signal of the FERRY complex, *mdh2* mRNA and mitochondria were in close proximity

(Figure 6B, blue box) or even co-localizing (Figure 6B, grey box). These findings support the notion that the FERRY complex is involved in the localization and the distribution of specific mRNAs such as transcripts encoding for mitochondrial proteins (e. g. *mdh2* mRNA), most probably through mediating their endosomal transport (Figure 6C).

Discussion

A novel link between the endosomal system and the translation machinery

In this study we identify and characterize a novel Rab5 effector complex, named the FERRY complex, which is composed of five subunits, named Fy-1 to Fy-5. It is able to bind activated Rab5 on early endosomes, while it simultaneously associates with ribosomes and mRNA through direct interaction with mRNA (Figure 6C). Thereby, we discovered a new link between early endosomes and the translation machinery in higher eukaryotes, providing molecular insights into the molecular mechanisms regulating the association of mRNA with endosomes. Furthermore, a screen for FERRY-associated mRNAs revealed a strong enrichment for specific groups of transcripts (*e.g.* mRNA for mitochondrial proteins), already indicating that the FERRY complex can selectively bind to RNA (Figure 6C). This selectivity seems to originate from its ability to exhibit a different binding mode for different transcripts, through a composite binding interface comprising several subunits (Quentin et al., 2021). These features allow the FERRY complex to transform the early endosome into a transport vehicle for mRNA distribution. Unlike EGF or transferrin, which reside inside the endosome, the RNA is transported on the outside.

Specificity in mRNA distribution in neurons

Local translation in neurons requires active transport of thousands of mRNAs over long distances to the far processes of axons or dendrites (Das et al., 2021). It is impossible that every transcript has a specialized transport vehicle, raising the question as to how mRNA distribution is organized and how many different vehicles are involved in this process. Furthermore, it is unknown how mRNAs are targeted to the right location, for example to mitochondria. From the pool of FERRY-associated mRNA candidates that were identified by RNA sequencing, we observed that the complex binds different transcripts with different efficacies *in vitro*. Given the intricate mRNA binding interface of the FERRY complex, one can envision how the complex might be able to discriminate between different mRNAs (Quentin et al., 2021). However, mechanisms of specific interaction, localization and translation are much more complex and include regulatory elements for dynamic mRNA interaction, the recognition of post-transcriptional mRNA modifications, ways to recognize the correct target location as well as translational regulation. The FERRY complex may serve as a molecular tool to address these different aspects of specificity in mRNA trafficking in future studies.

The FERRY complex from an evolutionary perspective

The presence of mRNA or ribosomes on endosomes seems to be a common feature of eukaryotes, which has been observed in different organisms ranging from fungi to humans (Cioni et al., 2019; Das et al., 2021; Higuchi et al., 2014). To which degree this link between two fundamental functionalities, such as the endosomal system and the translation machinery, relies on common molecular mechanisms is less clear. Our phylogenetic analysis indicated that the FERRY complex has developed from an ancestral remnant in some fungi, via a three-protein assembly in insects and some nematodes, to its full extent in the Chordata (Figure 1G). The reduced version of the FERRY complex lacks Fy-4 and Fy-5 and comprises a shorter variant of Fy-2, which lacks the middle domain between the two, terminal coiled-coils (Figure 1B). While the interface on Fy-2 for Fy-1 and Fy-3 is located in the conserved C-terminal coiled-coil region, the binding interfaces for Fy-4 and Fy-5 reside in the middle domain of Fy-2, which is absent in the three-subunit version (Quentin et al., 2021). Hence, the locations of the interfaces indicate that the three-subunit version of the FERRY complex is still able to form a complex. However, whether this complex still links the early endosome with the translation machinery is not yet understood. With the nervous system becoming more and more complex during the course of evolution and since the loss of the FERRY complex has detrimental effects on the brain, it would be interesting to see whether the transition from the three-subunit to the five-subunit complex established new functionalities or additional layers of regulation.

Attachment of mRNA on endosomes

Recent studies have highlighted the vital role of different endosomal compartments, for mRNA transport and localization. While Annexin 11A mediates the binding and transport of mRNP granules on lysosomes, late endosomes were not only identified as mRNA transport vehicles but also serve as platform for translation in neurons (Cioni et al., 2019; Liao et al., 2019). With the occurrence of different connections between the endosomal system and the translation machinery, questions arise as to how many different transcripts bind to an endosome, how many mRNA binding sites can endosomes offer and whether these are provided by different RBPs. The observation of several, up to four mRNA foci, on a single endosome (Figure 5C) indicates that endosomes may be able to accommodate multiple mRNA binding events. However, it does not answer the question whether these originate from the same RBP or from different mRNA attachment systems. The presence of multiple different physical contacts between endosomes and mRNA is supported by a recent study, showing that transcripts can interact with early endosomes in a translation-dependent or translation independent fashion, which points towards different mechanisms (Popovic et al., 2020). However, the molecular mechanism of these binding modes remains elusive. Having multiple RBPs or complexes on the endosomes, immediately raises the question of how these systems are connected and whether they share tasks, either by function (mRNA transport, translational regulation), mRNA specificity (mitochondrial transcripts, synaptic mRNAs, mRNAs for stress response) or

location (mitochondria, synapse) or they have rather redundant functions, which might increase the robustness of mRNA localization. Furthermore, how the individual proteins or protein complexes communicate with each other will be an interesting field of future discoveries.

The role of mitochondrial transcripts

The transport of mRNA for mitochondrial proteins on late endosomes was recently reported (Cioni et al., 2019). Now we propose the FERRY-mediated association of mitochondrial transcripts to the early endosome. Indeed, the specific late endosome cargo transcripts *lmnb2* and *vdac2* were also enriched in our screen for FERRY-associated transcripts, which raises the question about the purpose of different mRNA localization systems for a group of transcripts or even a single mRNA. Furthermore, a phylogenetic analysis also revealed the presence of the FERRY complex in *X. laevis* ruling out an explanation by genomic differences (Figure S2). However, nuclear-encoded messages for mitochondrial proteins form a large group of high abundant mRNAs, that need to be specifically localized to very distal sub-compartments of neurons, such as axonal growth cones or synapses. This opens up a variety of possible explanation, ranging from simple redundancy, to a division of labor in delivering to different neuronal sub-compartments, transporting different cargo mRNAs, or providing a different regulatory impact on the mitochondria, to even a scenario where one system is more responsible for transport while the other predominantly organizes storage or translation. Given the complex morphology of neurons and their energy requirement in various sub-compartments, an intricate system to maintain mitochondrial integrity and secure energy supplies is not surprising.

Connection between mRNA localization and neurodegeneration

A disruption of the connection between the endosomal system and the translation machinery by mutation or genetic loss of a protein is often attended by neurological defects, such as epilepsy and neurodegeneration through the loss of the FERRY complex or in case of Marie-Charcot-Tooth disease by a mutation in Rab7 (Cioni et al., 2019). Other neurodegenerative diseases are also linked to mitochondrial dysfunction, such as Parkinson's, Huntington's or Alzheimer's disease (reviewed in: (Abou-Sleiman et al., 2006; Moreira et al., 2010; Park et al., 2018; Reddy et al., 2009)). Coincidentally, two of the three genetic diseases mentioned before seem to have an impact on the localization of transcripts for mitochondrial proteins, which might in turn affect mitochondrial function. While under normal conditions mitochondria might be able to compensate for the impairment of the supply chain for some proteins, a compensatory mechanism might fail, if too many proteins are affected by the lack of supply or if external or internal stresses overcharge the compensatory mechanisms. The extreme morphology of neurons, with their long and thin processes, offers an explanation as to why neurons are predominantly affected when mRNA transport is disrupted, since in cells with shorter dimensions diffusion might be able to compensate for that loss. In this scenario, it is also possible that the aberrant localization of mitochondrial transcripts, for example, can cause

a multitude of symptoms, depending on the neuronal sub-compartment in which the transcripts are missing, which might be determined by the respective mRNA distribution system.

Acknowledgements

Firstly, we thank R. Schäfer for her support with cell culture and cloning and I. Bartnik for excellent technical support. We also acknowledge S. Raunser and D. Quentin for valuable feedback regarding the manuscript and the members of the cluster of excellence "Physics of Life" for stimulating discussion. Especially, we would like to thank the following Services and Facilities of the MPI-CBG for their support: The antibody Facility, the light microscopy facility, the mass spectrometry facility, the genome engineering facility and protein expression and purification facility. We also thank the DRESDEN-concept Genome Center (DcGC at CMBC at the TU Dresden) supported by DFG (INST 269/768-1) for technical support. Furthermore, we would like to thank Refeyn Ltd (Oxford, UK) for the use of their Mass Photometer. Part of this work and J.S.S. were funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project Number 112927078 - TRR 83.

Author contributions

Conceptualization, J.S.S. and M.Z.; Software, L.H.; Formal Analysis, J.S.S., C.L., L.H., Y.K., A.T.-P. and M.Z.; Investigation, J.S.S., S.t.D., A.G., S.S., S.C. and M.Z.; Data Curation, L.H.; Writing – Original Draft, J.S.S. and M.Z.; Writing – Review & Editing, all authors.; Visualization, J.S.S., C.L. and L.H.; Supervision, J.S.S. and M.Z.; Project Administration, J.S.S. and M.Z.; Funding Acquisition, E.M.S and M.Z.

Competing interests:

The authors declare no competing financial interests.

Data availability:

RNA Sequencing (RNA-Seq) data and the respective scripts for the analysis of the RNA-Seq and proteomics data are available in a public repository (https://dx.doi.org/21.11101/0000-0007-EEE3-D).

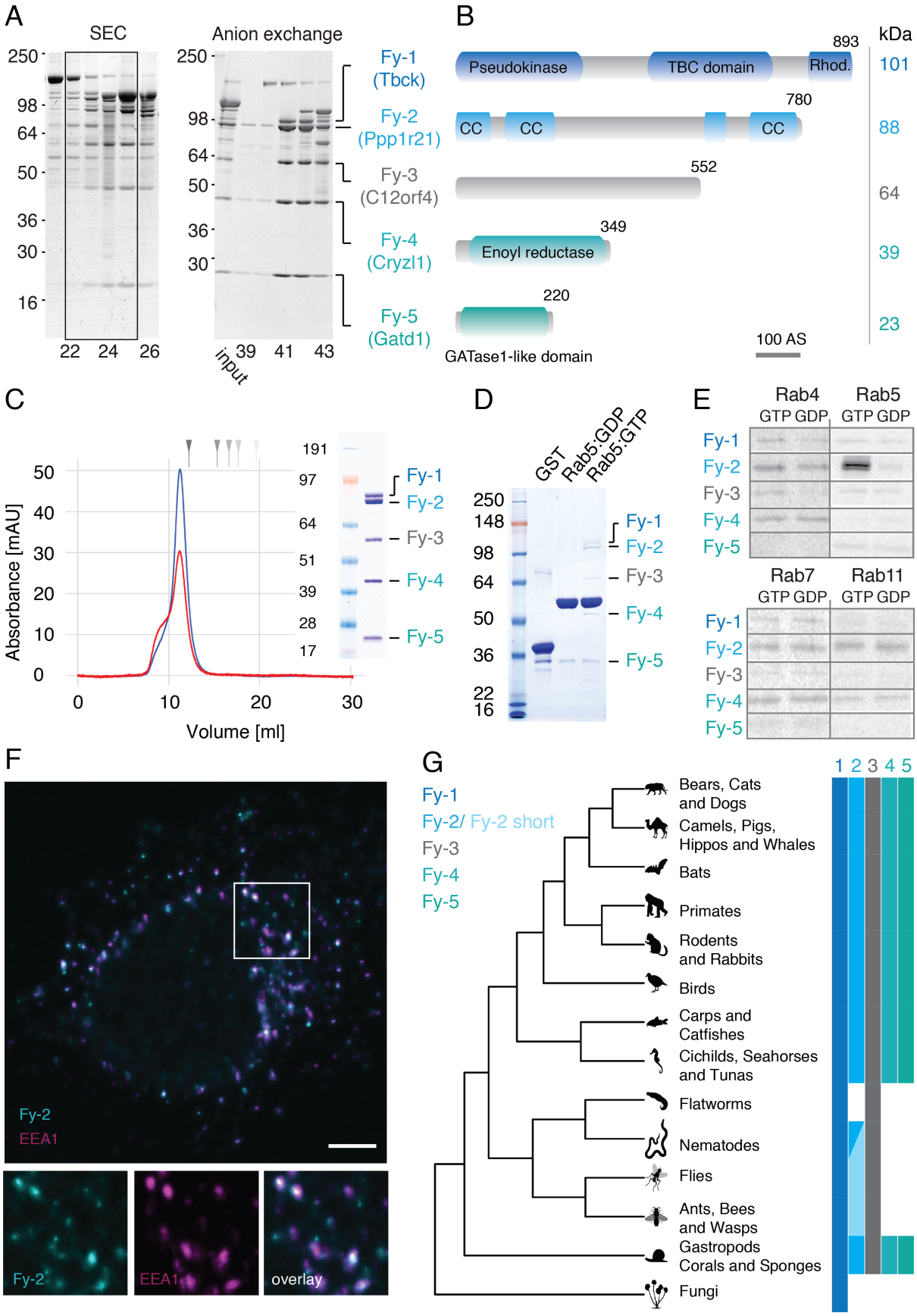


Figure 1: **A)** The entirety of Rab5 effectors, obtained by Rab5 affinity chromatography, was fractionated by two chromatographic techniques, i) size exclusion chromatography (SEC, left) and ii) anion exchange chromatography (right). The fractions were analyzed by SDS-PAGE and Coomassie staining. The five proteins of the new FERRY

complex co-eluted in fractions 22 to 25 from SEC (left gel). Fractions 22 to 25 were combined and subjected to an anion exchange chromatography. The fractions obtained were analyzed by SDS-PAGE and Coomassie staining (right gel). The input (loaded material) and fractions 39 to 43 are shown. **B)** Scheme of the domain architecture of the components of the FERRY complex drawn to scale (Rhod.: Rhodanese domain, CC: coiled-coil). **C)** SEC profile of the FERRY complex (blue: 280 nm, red: 254 nm) with a Coomassie-stained SDS PAGE of the peak fraction as inset. The grey arrows represent a molecular weight standard (670, 158, 44, 17, 1.35 kDa). **D)** Coomassie-stained SDS PAGE of an *in vitro* pulldown assay using GST, GST-Rab5:GDP and GST-Rab5:GTP to probe the interaction with the FERRY complex. **E)** Fluorographic analysis of GST binding assays using different Rab GTPases in the active and inactive state against *in vitro* translated ^{35}S methionine-containing components of the FERRY complex. **F)** Immunostaining of HeLa cells against EEA1 and Fy-2 (Scale bar: 5 µm). The individual channels of the boxed region is shown below in higher magnification. **G)** Phylogenetic analysis of the subunits of the FERRY complex (a full tree including individual species is given in Figure S2).

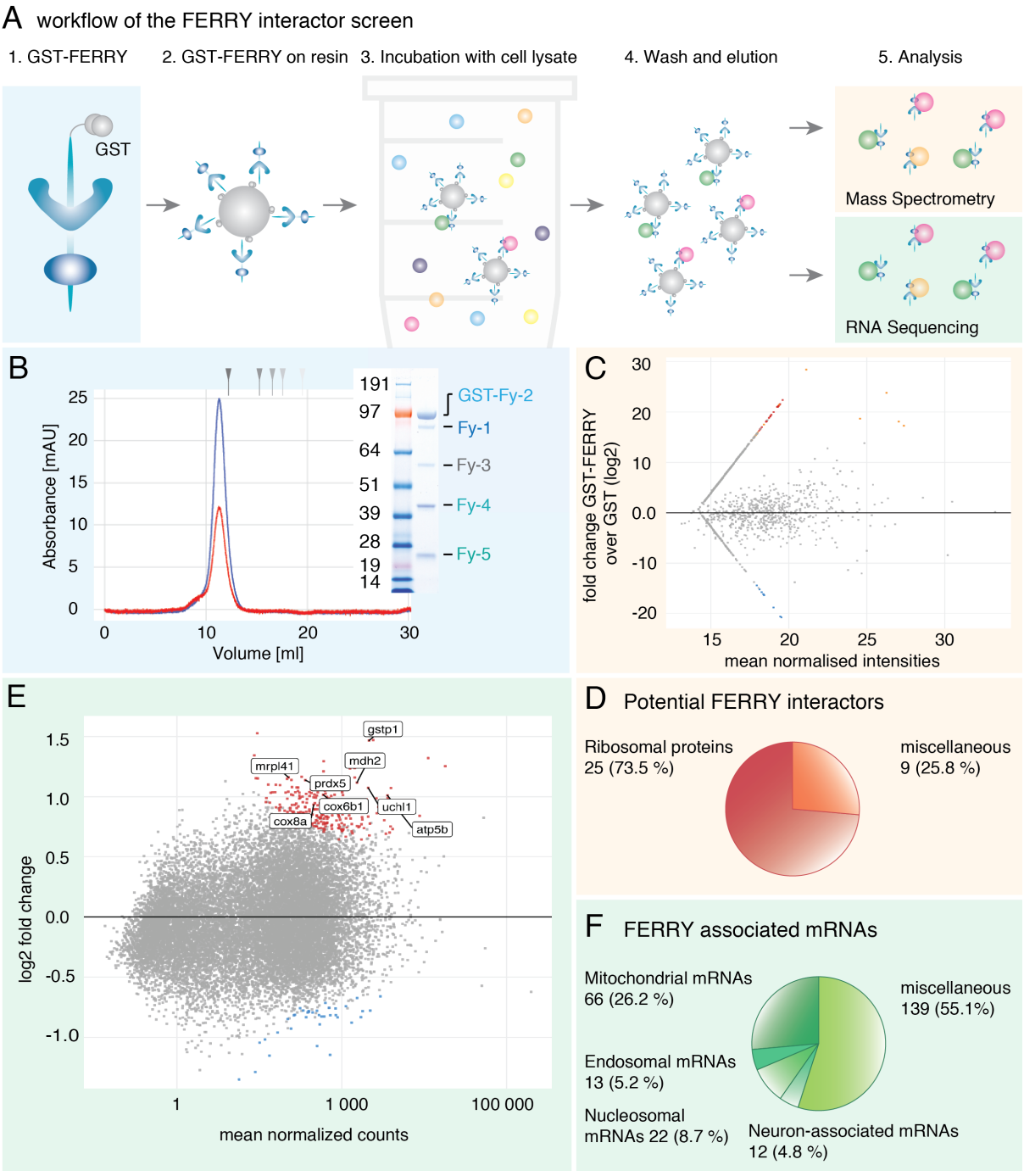


Figure 2: A) Scheme of the workflow of the *in vitro* GST-FERRY interactor screen. **B)** SEC profile of GST-FERRY (blue: 280 nm, red: 254 nm) with an SDS PAGE of the peak fraction as inset. The grey arrows represent a molecular weight standard (670, 158, 44, 17, 1.35 kDa). **C)** MA blot of the mass spectrometry results of the GST-FERRY interactor screen, with a grey dot for each protein. Candidates enriched in GST-FERRY and GST are indicated in red and blue, respectively. **D)** The pie chart visualizes different groups of the 34 potential FERRY interactors **E)** MA blot of the RNA sequencing of potential FERRY-associated mRNAs. Each dot represents a specific transcript. mRNA candidates associated with GST-FERRY and GST are highlighted in red and blue, respectively, with some candidates labeled **F)** The pie chart shows the different groups of the 252 FERRY-associated mRNAs.

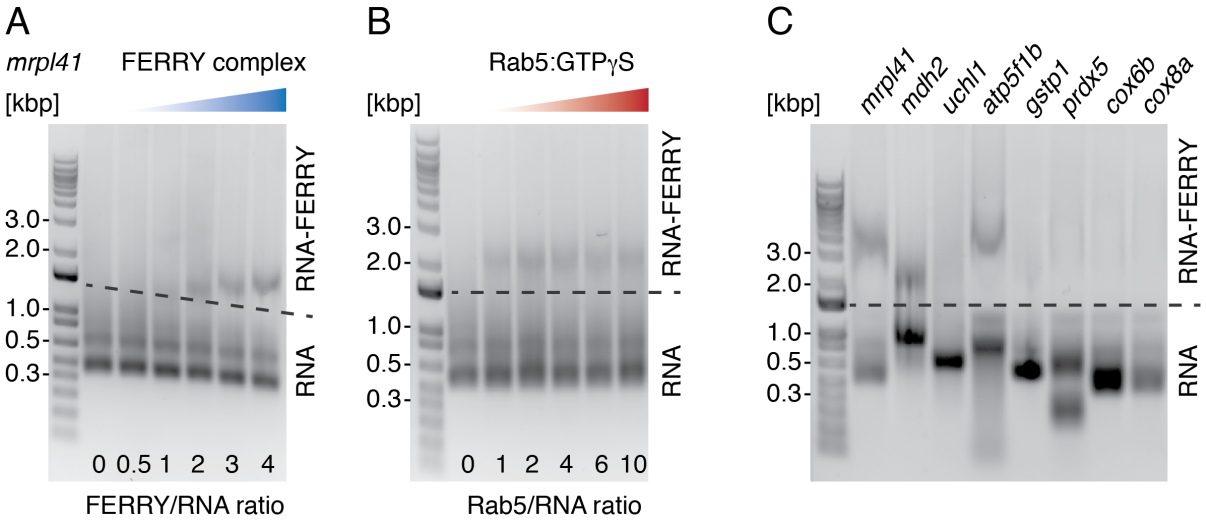


Figure 3: A) Results of an electrophoretic mobility shift assay (EMSA) to test the interaction between the FERRY complex and *mrpl41* mRNA with increasing ratios of FERRY complex to RNA. **B)** EMSA to probe the interaction between the FERRY complex and *mrpl4* in the presence of Rab5:GTPγS. For this assay a fixed ratio of FERRY complex to RNA of 3 was used, while the amounts of Rab5:GTPγS were successively increased as indicated. **C)** EMSA to assess the interaction of the FERRY complex with different mRNAs. This assay was performed at a fixed FERRY/mRNA ratio of 5.

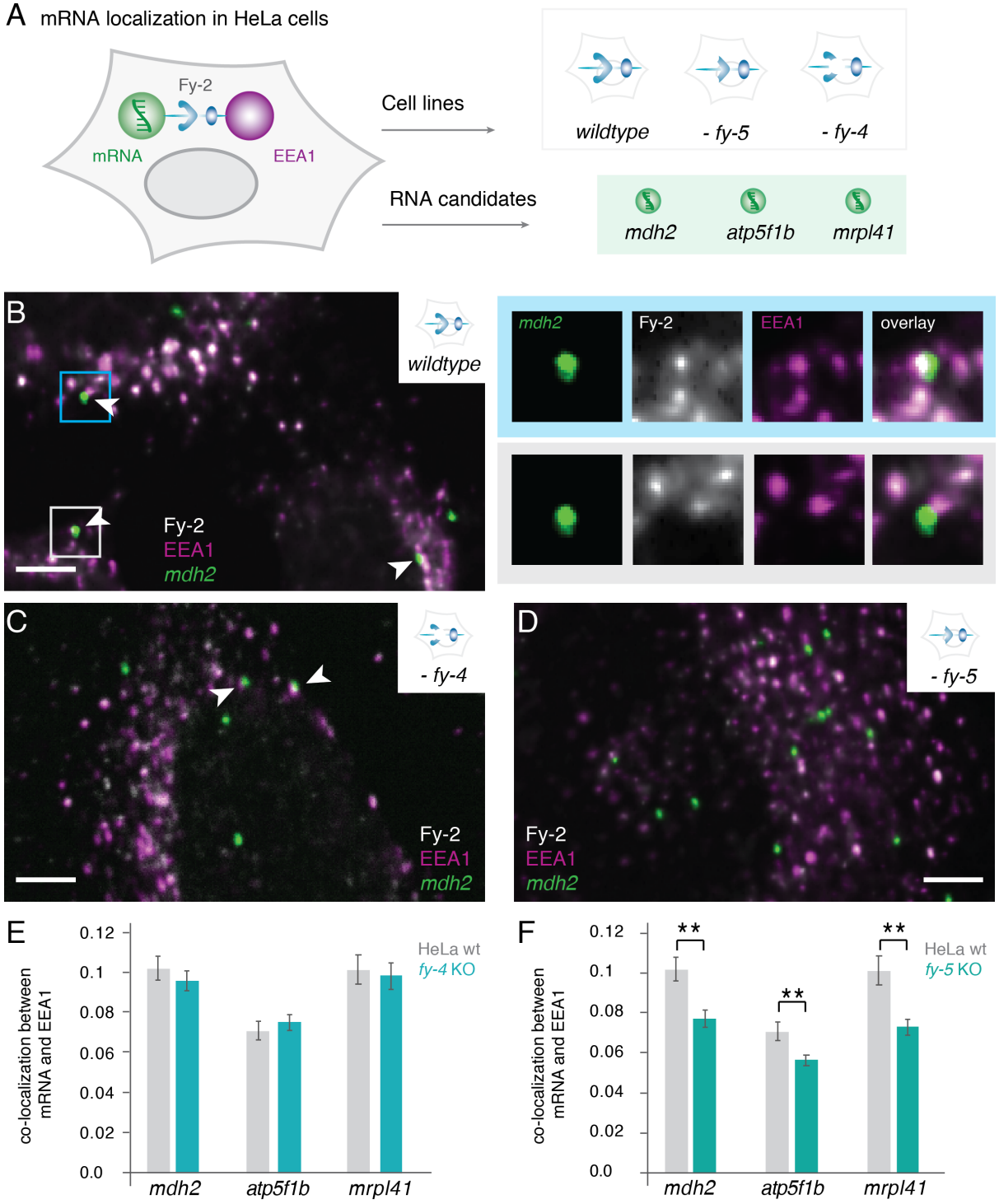


Figure 4: A) Scheme of the mRNA localization/ immunofluorescence experiment, showing the different markers (mRNA: smFISH, EEA1: antibody and Fy-2: antibody), mRNAs (in the green box) and cell lines (grey box) involved. **B)** Exemplary image of the combined visualization of Fy-2, EEA1 and *mdh2* mRNA in wildtype HeLa cells (Scale bar: 5 µm). Events of co-localization of mRNA with Fy-2 and EEA1 are indicated with white arrow heads. Magnified images of the individual channels and the overlay of the two regions boxed in blue and grey are given on the right side. (images: 3.9 x 3.9 µm). **C)** Exemplary image visualizing Fy-2, EEA1 and *mdh2* mRNA in the absence of Fy-4 (Scale bar: 5 µm). Events of co-localization of mRNA with Fy-2 and EEA1 are indicated with white arrow heads. **D)** Exemplary image visualizing Fy-2, EEA1 and *mdh2* mRNA in the absence of Fy-5 (Scale bar: 5 µm). **E)** Quantification of co-localization of the different mRNAs and EEA1 in HeLa wildtype cells

571 and upon knock-out of Fy-4. **F)** Quantification of co-localization of the different mRNAs and EEA1 in HeLa
572 wildtype cells and upon knock-out of Fy-5. The asterisk indicates a statistically significant difference (p-value
573 < 0.01)

574

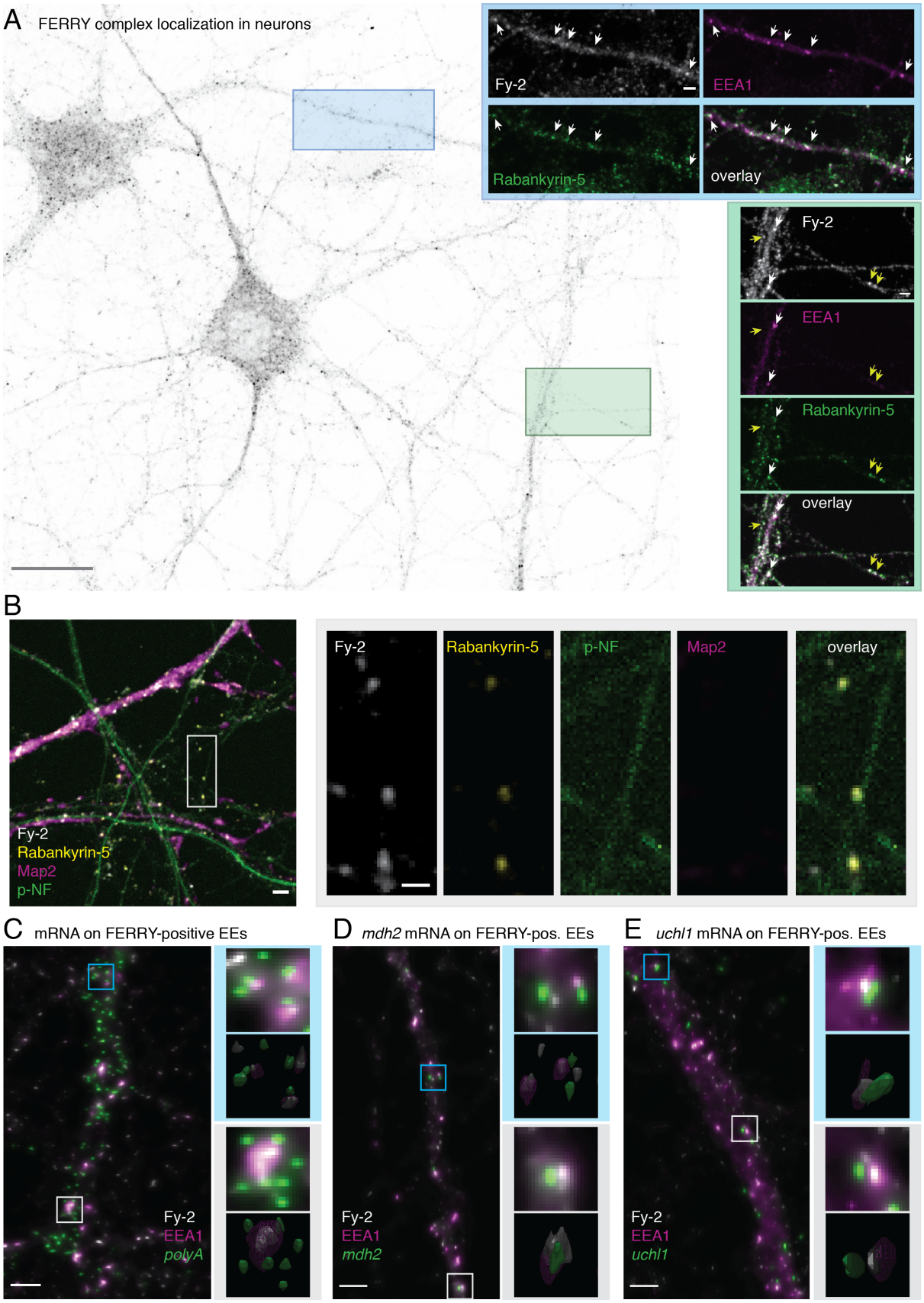
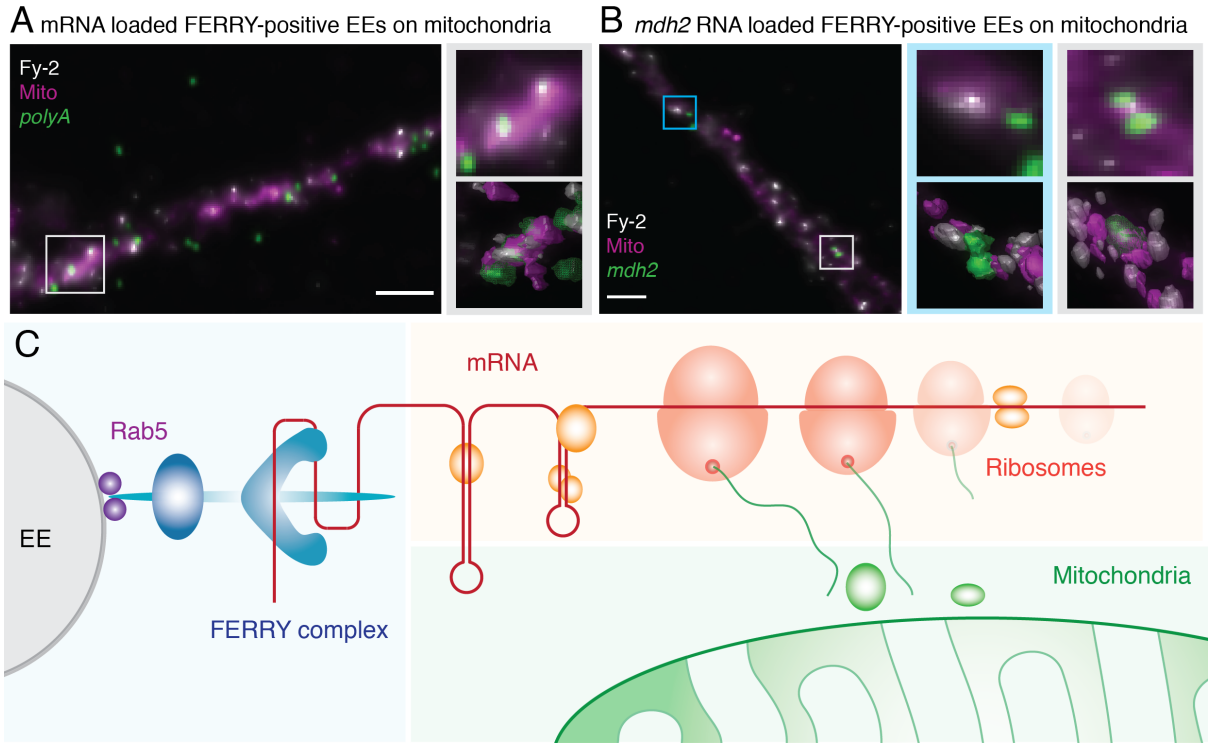


Figure 5: A) localization of the FERRY complex in neurons. Primary rat hippocampal neurons were grown at low density supported by an astrocyte feeder layer. After fixation, the Rab5 effectors Fy-2, EEA1 and Rabankyrin-5 were visualized by immunostaining. The overview image shows the localization of the FERRY complex on a

larger scale (Scale bar: 20 µm). The insets highlighted in green and blue show two boxed regions in higher magnification and additionally the localization of EEA1 and Rabankyrin-5 and the triple overlay (Scale bar: 2 µm). Endosomes with a co-localization of all three markers are indicated with white arrowheads, while endosomes where only Fy-2 and Rabankyrin-5 co-localize are marked with yellow arrowheads. **B)** Primary rat hippocampal neurons were stained for Fy-2, Rabankyrin-5, Map2 and a phosphorylated neurofilament (pNF). The overview image shows an overlay of all four markers (Scale bar: 2 µm). A magnification of the region in the white box is given on the right, showing the single channels as well as an overlay (Scale bar: 1 µm). **C)** Combination of immunostaining of Fy-2 and EEA1 with smFISH against the polyA tail of mRNA in primary hippocampal neurons. Magnified images of the regions indicated in light grey and blue are given on the right in combination with a 3D representation. **D)** and **E)** show images of a combination of immunostaining of Fy-2 and EEA1 with smFISH against *mdh2* and *uchl1* mRNA in primary hippocampal neurons (Scale bar: 2 µm). Magnified images of the regions indicated in light grey and blue are given on the right in combination with a 3D representation.



592

593 **Figure 6: A)** Combination of immunostaining of Fy-2 and Mitochondria (Tomm70a) and smFISH against polyA
594 (Scale bar: 2 µm). The boxed region is presented in higher magnification along with a 3D representation of the
595 fluorescence signal. **B)** Immunostaining of Fy-2 and Mitochondria (Tomm70a) combined with smFISH against
596 *mdh2* (Scale bar: 2 µm). The boxed region is presented in higher magnification along with a 3D representation of
597 the fluorescence signal **C)** Scheme of the current understanding of the cellular role of the FERRY complex. This
598 novel Rab5 effector complex connects the early endosomes (blue colors) with the translation machinery (red
599 colors) and predominantly associates with transcript for mitochondrial proteins (green).

600

Methods

Molecular Cloning

Human Fy-1 (Tbck, ENSG00000145348, Q8TEA7), Fy-2 (Ppp1r21, ENSG00000162869, Q6ZMI0), Fy-3 (C12orf4, ENSG00000047621, Q9NQ89), Fy-4 (Cryzl1, ENSG00000205758, O95825), Fy-5 (Gatd1, ENSG00000177225, Q8NB37) and Rab5a (ENSG00000144566, P20339), were amplified by polymerase chain reaction (PCR) using Q5 High-Fidelity DNA polymerase (NEB) and digested using NotI, NcoI, AscI, XhoI, PciI (NEB) according to the manufacturer's protocol. Fy-5 was cloned into a pET based bacterial expression vector as an N-terminally hexahistidine (His_6) tagged variant without cleavage site. Fy-4 was cloned into an expression vector for expression in SF9 cells also carrying a non-cleavable N-terminal His_6 tag. Fy-1, Fy-2 and Fy-3 were cloned into a multi gene construct based on a pBLA vector. For the purification of the FERRY complex Fy-1 carried a cleavable N-temrinal His_6 tag, the other 2 genes were untagged. To obtain GST-FERRY, Fy-2 carried a cleavable Gultathione-S-transferase (GST) tag, while Fy-1 and Fy-3 remained untagged. Rab5 was used as GST fusion variant in the bacterial expression vectors pGAT2 for GST pulldown assays and pGEX-6P-3 for electrophoretic mobility shift assays (EMSA). Plasmids and primers used in this study are listed in the resources table (Table S2).

Virus production and insect cell expression

SF9 cells growing in ESF921 media (Expression Systems) were co-transfected with linearized viral genome and the expression plasmid, and selected for high infectivity. P1 and P2 viruses were generated according to the manufacturer's protocol. Best viruses were used to infect SF9 cells at 10^6 cells/mL at 1% vol/vol and routinely harvested after around 48 hours at about $1.5x10^6$ cells/ml. The pellet was suspended in lysis buffer (20 mM HEPES (pH 7.5), 250 mM NaCl, 20 mM KCl, 20 mM $MgCl_2$ and 40 mM imidazole) or SEC buffer (20mM HEPES, pH 7.5, 250mM NaCl, 20mM KCl, 20mM $MgCl_2$) supplemented with a protease inhibitor cocktail, flash frozen in liquid nitrogen and stored at -80 degrees.

Protein purification

Fy-5 and GST-Rab5:

For expression of Fy-5 and GST-Rab5, *E. coli* BL21 (DE3) (company) were grown in LB medium under autoinduction conditions using D-(+)-lactose monohydrate at 1.75% (w/v), supplemented with the respective antibiotic (50 µg/mL kanamycin or100 µg/mL ampicillin) at 30 °C under constant shaking (165 rpm). Bacteria were harvested by centrifugation (4000 x g, 20 min, 4 °C), suspended in lysis buffer and subsequently lysed or stored at -80 °C. After sonication the lysate was clarified by centrifugation (22 500 rpm/61 236 x *g*, 20 min, 4 °C) and

applied to a HisTrap FF column (GE Healthcare) equilibrated with 10 column volumes (CV) of lysis buffer. After extensive washing with lysis buffer, the proteins were eluted in 10-13 ml elution buffer (20 mM HEPES (pH 7.5), 250 mM NaCl, 20 mM KCl, 20 mM $MgCl_2$ and 500 mM imidazole). Elution fractions containing protein were concentrated using Amicon Ultracel-10K/ Ultracel-30K (Millipore) centrifuge filters and subsequently applied to size exclusion chromatography (SEC) using a Superdex 200 column (HiLoad 16/600 Superdex 200 pg, GE Healthcare) equilibrated in SEC buffer. Fractions were analysed using SDS-PAGE. Protein containing fractions were pooled and concentrated to fit experimental requirements. Protein concentrations were determined by spectrophotometer (NanoDrop Lite, Thermo Scientific).

Fy-4:

For expression of Fy-4, insect cell suspensions were lysed using sonication, the lysate subsequently clarified by centrifugation (22 500 rpm/61 236 x *g*, 20 min, 4 °C), filtrated using Millex® HV membrane filter units with a pore size of 0.45 µm (Merck Millipore) and applied to a HisTrap FF column (GE Healthcare) equilibrated with 10 CV of lysis buffer. After washing with lysis buffer, the protein was eluted in 10-13 ml elution buffer and concentrated with a centrifuge filter, (Amicon Ultracel-30K, Millipore). Thereafter, the protein was applied to SEC using a Superdex 200 column (HiLoad 16/600 Superdex 200 pg, GE Healthcare) equilibrated in SEC buffer. The fractions were analysed by SDS-PAGE. Protein containing fractions were pooled and concentrated according to experimental requirements. The protein concentration was determined by spectrophotometer (NanoDrop Lite, Thermo Scientific).

FERRY complex:

SF9 cell pellets prior infected with a virus containing Fy-1, Fy-2 and Fy-3 were melted and immediately supplemented with an excess of purified Fy-4 and Fy-5 before lysis. Subsequently, the cells were lysed using a Microfluidizer (LM20, Microfluidics). The lysate was clarified by centrifugation (22 500 rpm/61 236 x *g*, 20 min, 4 °C) and filtrated using membrane filters with a pore size of 0.45 µm (Millex® HV membrane filter units, Merck Millipore). The clarified lysate was supplemented with Ni-NTA agarose (1.3 ml resin/ 1 l insect cell pellet, Qiagen) and incubated for 30 mins at 4 °C on a rotating wheel. Subsequently, the resin was transferred into gravity flow chromatography columns (Poly-Prep® Chromatography Column, Bio-Rad) and washed 3 times with i) 8 CV lysis buffer, ii) 8 CV wash buffer (20 mM HEPES, pH 7.5, 250 mM NaCl, 20 mM KCl, 20 mM $MgCl_2$ and 80 mM imidazole), and iii) 8 CV lysis buffer. The protein was eluted in 1 ml fractions with elution buffer and protein containing fractions were applied to SEC without further concentration, using either a Superdex 200 (HiLoad 16/600 Superdex 200 pg, GE Healthcare) or a Superose 6 increase (Superose 6 Increase 10/300 GL, GE Healthcare) which were equilibrated in SEC buffer. Protein containing fractions were pooled and concentrated according to experimental requirements. Concentration was determined by a spectrophotometer (NanoDrop Lite, Thermo Scientific)

676 **GST-FERRY complex:**

677 SF9 cell pellets prior infected with a virus containing Fy-1, GST-Fy-2 and Fy-3 were melted
678 and immediately supplemented with an excess of purified Fy-4 and Fy-5. The cells were lysed
679 using a Microfluidizer (LM20, Microfluidics), the lysate was clarified by centrifugation
680 (22 500 rpm/61 236 x *g*, 20 min, 4 °C) and subsequently filtrated using membrane filters with
681 a pore size of 0.45 µm (Millex® HV membrane filter units, Merck Millipore). The clarified
682 lysate was supplemented with Glutathione Sepharose 4B (Cytiva, 2.2 ml resin/1 l insect cell
683 pellet) and incubated for 1.5 h at 4 °C on a rotating wheel. The beads were washed once with
684 10 ml SEC buffer supplemented with purified Fy-4 and 5 and 2 times with 10 ml SEC buffer.
685 To elute the GST-FERRY complex, the beads were incubated with GSH buffer (20 mM
686 HEPES (pH 7.5), 250 mM NaCl, 20 mM KCl, 20 mM $MgCl_2$, 20 mM GSH) for 1.5 h at 4 °C
687 on a rotating wheel and the beads were removed using filter columns (MoBiTec). The protein
688 complex was concentrated using centrifuge filters (Amicon Ultracel-30K, Millipore) and
689 subjected to SEC using a Superdex 200 column (HiLoad 16/600 Superdex 200 pg, GE
690 Healthcare) equilibrated in SEC buffer. Protein containing fractions were pooled and
691 concentrated according to experimental requirements. Concentration was determined by a
692 spectrophotometer (NanoDrop Lite, Thermo Scientific)

693 **Rab5:GTPγS:**

694 Expression of Rab5a was performed under autoinduction conditions as described before (Fy-5
695 and GST-Rab5). Harvested bacterial pellets were resuspended in SEC buffer and lysed using
696 sonication. Glutathione Sepharose 4B (Cytiva) was added to the clarified lysate and incubated
697 for 1.5 h at 4 °C. The resin was washed 3 times with SEC buffer and the protein cleaved off
698 the resin using HRV 3C protease (produced in house) at 4 °C over night on a rotating wheel.
699 Afterwards, the protein was concentrated using Amicon Ultracel-30K (Millipore) centrifuge
700 filters and subsequently applied to SEC using a Superdex 200 column (HiLoad 16/600
701 Superdex 200 pg, GE Healthcare) equilibrated in SEC buffer. Fractions were analyzed using
702 SDS-PAGE. Protein containing fractions were pooled and concentrated according to
703 experimental requirements. The protein concentration was determined by a spectrophotometer
704 (NanoDrop Lite, Thermo Scientific).

705 For the nucleotide loading, Rab5 was concentrated using an Amicon Ultracel-30K (Millipore)
706 centrifuge filter, subsequently supplemented with 2.5 mM GTPγS and 250 nM of a GST fusion
707 of the Rab5 GEF domain of Rabex5 (GST-Vps9) and incubated for 1 h on ice. To remove the
708 Rab5 GEF domain, Glutathione Sepharose 4B (Cytiva) was added to the mixture and incubated
709 for 1.5 h at 4 °C. The resin was pelleted by centrifugation (12 000 rpm/ 15 300 x g, 10 min,
710 4 °C) and the supernatant containing the GTPγS loaded Rab5 was flash frozen and stored at -
711 80 °C. The protein concentration was determined using a BCA assay (Pierce™ BCA Protein
712 Assay Kit, Thermo Scientific).

713

GST pulldown assay

1.1 nmol of purified GST or GST-Rab5 was incubated with 12 µl Glutathione Sepharose 4B (Cytiva) in 400 µl SEC buffer in small filter columns (MoBiTec) on a rotating wheel for 30 min at room temperature (rt). Subsequent centrifugation (4000 rpm/3500 x *g*, 1 min, 4 °C) removed unbound protein and the resin was washed once with 400 µl SEC buffer. For nucleotide exchange, 2 mM nucleotide (GDP or GTP) and 235 nM of GST-Vps9 was added to the columns in 400 µl SEC buffer and incubated for 10 min at rt. After centrifugation (4000 rpm/3500 x *g*, 1 min, 4 °C), 0.8 nmol FERRY complex was added to the columns in 400 µl SEC buffer and incubated for 10 min at rt. Again, unbound protein was removed by centrifugation (4000 rpm/3500 x *g*, 1 min, 4 °C) and the columns were washed 3 times with 400 µl SEC buffer. Proteins were eluted with 40 µl of GSH buffer (SEC buffer with 20 mM GSH) for 20 min at rt and analysed by SDS-PAGE.

Identifying orthologous sequences

We downloaded all eukaryotic reference proteomes from uniport (last accessed: March 2nd 2020) (UniProt, 2019). We used PorthoMCL (Tabari and Su, 2017) to identify orthologous clusters containing human FERRY components (GALD1_HUMAN, QORL1_HUMAN, CL004_HUMAN, PPR21_HUMAN, TBCK_HUMAN). Sequences deviating strongly in length (Figure S1B) from their human homolog were removed (Table S2). We further distinguished PPR21_HUMAN orthologs between sequences which contain a Fy-4 and a Fy-5 binding site and sequences which do not. For the detection of the presence of the Fy-4 and the Fy-5 binding sites, we aligned all identified Fy-2 sequences. We considered the binding sites present if all of the regions aligned to the PPR21_HUMAN binding regions contained less than 20% gaps (ignoring gapped sites in PPR21_HUMAN).

Phylogenetic tree estimation

All orthologous clusters were scanned for species which contain at least 80% of identified species with FERRY proteins (custom R script; R 3.6.1; (R Core Team, 2019)). Sequences belonging to FERRY containing species were extracted and aligned using MAFFT with default settings (Rozewicki et al., 2019). Each alignment was trimmed using trimAL (Capella-Gutierrez et al., 2009). The maximum likelihood (ML) tree was estimated using IQTree (Nguyen et al., 2015) whereby each protein was represented as a partition (Chernomor et al., 2016). The Whelan and Goldman matrix (Whelan and Goldman, 2001) with ML optimized amino acid frequencies (WAG+FO) was used as common model for all partitions. Branch support was calculated by IQTree via ultra-fast bootstrapping (UFBoot, 10,000) (Hoang et al., 2018). The consensus tree with the presents/absence information was visualized using the R package ggtree (Version 2.0.4) (Yu et al., 2018; Yu et al., 2017).

FERRY evolution and ancestral state reconstruction

The identified orthologous genes were used to estimate the ancestral composition of the FERRY complex. The probability for each protein's presence at each internal node was estimated using Pagel's algorithm (Pagel, 1994) implemented in the R package ape (Version 5.3) (Paradis and Schliep, 2019).

Antibody production

Rabbit polyclonal antibodies against Fy-4 were raised in NZW rabbits using standard procedures. 200 ug of recombinant protein emulsified in Complete Freund's adjuvant was used for immunization. Three boosts were done at 4-week intervals using 200 ug of recombinant protein emulsified in Incomplete Freund's adjuvant. The final bleed was harvested 10 days after the last boost. Antibodies were affinity-purified on Fy-4 immobilized on a HiTrap NHS-activated HP column (GE Healthcare). Antibodies were eluted using Pierce Gentle Ag/Ab Elution Buffer (ThermoFisher).

Mouse monoclonal antibodies against different components of the FERRY complex were raised in Balb/c mice after subtractive immunization (Sleister and Rao, 2001) with Fy-5. Mice were injected with recombinant Fy-5 in the presence of the immunosuppression drug cyclophosphamide in order to preferentially eliminate Fy-5-reactive B and T lymphocytes. Thereafter the mice were immunized with the entire FERRY complex. Hybridoma were generated using PEG fusions following standard protocols. Clones reacting with individual components of the FERRY complex were selected in a multiplex electrochemiluminescence assay on the MSD platform (Mesoscale Discovery, Rockville, MD). Antibodies were purified from hybridoma supernatant using HiTrap Protein G columns (GE Healthcare).

Antibody validation

Validation of in-house produced antibodies against components of the FERRY complex for western blot (WB) were tested against 100 ng, 10 ng and 1 ng of recombinant FERRY complex. Candidates with high sensitivity (detection of 1 ng) and good selectivity (preferably no or no interfering additional signal) were chosen.

To validate the mouse monoclonal Fy-2 antibody for immunofluorescence (IF), we generated a *fy-2* knock-out HeLa cell line making use of the CRISPR/Cas9 technology. Even though, western blot analysis showed the disappearance of the Fy-2 signal (Figure S4), we observed residual signal by immunofluorescence with the same antibody (Figure S1). Given the fact, that the western blot signal is already weak in the wildtype (wt), a small residual fraction of protein might be below the detection limit. The remaining fluorescence signal may either be

caused by residual, maybe truncated Fy-2 protein in the *fy-2* KO cell line or the recognition of an additional protein by the antibody. Nevertheless, we observed a clear difference in signal intensity between wt and the *fy-2* KO condition in immunofluorescence (Figure S1) and therefore concluded that the antibody recognizes Fy-2 and is suitable for immunofluorescence. To further control for the specificity, we co-stained with EEA1 whenever possible and checked the fluorescence signal by visual inspection. We also excluded the Fy-2 antibody signal from automated image analysis, especially automated object detection, since the residual signal interferes with finding general parameters for object identification.

Antibodies

The following primary antibodies were used for IF or WB experiments at the concentrations or dilutions indicated: anti-Rab5 (mouse, monoclonal, BD Bioscience, 610725, IF 1:100), anti-EEA1 (rabbit, polyclonal, laboratory-made, IF 1:1000), anti-Rabankyrin-5 (rat, monoclonal, laboratory-made, IF 1:2000), anti-Map2 (rabbit, polyclonal, Chemicon, IF 1:1000), anti-pNF-H (mouse, monoclonal, Biolegend, IF 1:5000), anti-Fy-1 (rabbit, polyclonal, Sigma Aldrich, HPA039951, WB 1:1000) anti-Fy-2 (mouse, monoclonal, laboratory-made, IF 1:1000, WB 0.5 µg/µl), anti-Fy-3 (rabbit, polyclonal, Sigma Aldrich, HPA037871, WB 1:1000), anti-Fy-4 (rabbit, polyclonal, laboratory-made, IF 1:1000, WB 0.5 µg/µl), anti-Fy-5 (mouse, monoclonal, laboratory-made) WB (0.5 µg/µl) and anti-GAPDH (rabbit, monoclonal, Sigma Aldrich, G8795, WB 1:5000).

The following fluorescent secondary antibodies for immunostainings were purchased from Invitrogen and used in a 1:1000 dilution: Goat anti-Rat IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488, Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 568, Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 405, Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 647, F(ab')2-Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 647, Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488. For Western blot horseradish peroxidase (HRP) secondary antibodies were supplied from Jackson ImmunoResearch and used at a 1:10 000 dilution.

HEK 293 lysate preparation

FreeStyle™ 293-F Cells (Thermo Fisher Scientific) were grown in suspension culture in FreeStyle™ 293 Expression Medium (Thermo Fisher Scientific) to density of 4 x 10^6 cells/ml and harvested by centrifugation (500 x g, 10 min, 20 °C). The cell pellets were suspended in lysate buffer (6 ml/ liter cell culture, 50 mM HEPES (pH 7.5), 100 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT, 0.1% Tween 20), supplemented with a protease inhibitor cocktail and immediately

flash frozen in liquid nitrogen. For lysate preparation the pellets were melted, lysed using a microfluidizer (LM20, microfluidics). The lysate was subsequently clarified by a two-step centrifugation (4000 rpm/ 1935 x g, 10 min, 4 °C and 22 500 rpm/ 61 236 x *g*, 25 min, 4 °C), yielding around 15 ml cells lysate per liter cell culture.

GST-FERRY interactor screens

The GST-FERRY interactor screen was performed at 4 °C in gravity flow filter columns (Poly-Prep® Chromatography Column, Bio-Rad). 500 µl Glutathione Sepharose 4B (GE Healthcare) was added to 0.8 µmol of GST or 7 mg of GST-FERRY complex in 9 ml SEC buffer and incubated for 2.5 h on a rotating wheel. The solution was let run through and the resulting bed of beads was washed 3 x 2 ml SEC buffer. 10 ml of freshly prepared HEK 293 lysate was added to each column and incubated for 1.5 h on a rotating wheel. The lysate was allowed to flow through and another 5 ml of cell lysate was added to each column and also run through the column. The columns were extensively washed with 4 ml lysis buffer and 2 x 5 ml and 2 x 7 ml SEC+ buffer (20 mM HEPES, pH 7.5, 250 mM NaCl, 20 mM KCl, 20 mM $MgCl_2$, 1 mM DTT and 0.1% Tween 20). For the elution of the proteins the columns were incubated with 500 µl of GSH buffer for 40 min on a rotating wheel. The elution fractions were visualized by SDS PAGE and further analysed by mass spectrometry.

To isolate FERRY-associated RNA, the GST-FERRY interactor experiment was performed as described with slight modifications. For the elution of the proteins and the associated RNA, RLT buffer from the AllPrep DNA/RNA/miRNA Universal Kit (Qiagen) was supplemented with 1% β-Mercaptoethanol and 20 mM GSH and the pH adjusted to 7.5. The subsequent isolation of nucleic acids was performed using the AllPrep DNA/RNA/miRNA Universal Kit (Qiagen) according to the manufacturer's protocol. The obtained RNA samples were flash frozen and stored at -80 °C. Prior sequencing, the concentration of the samples was determined by spectrophotometer (NanoDrop Lite, Thermo Scientific) and the samples were analyzed using a 2100 Bioanalyzer (Agilent).

Mass spectrometry

Samples were separated on SDS PAGE, visualized with Coomassie staining and entire gel lanes cut in 10 pieces each of which was processed individually. Proteins were in-gel reduced by dithiothreitol (DTT), alkylated by iodoacetamide and digested overnight with trypsin (Promega). The resulting peptide mixtures were extracted twice by exchange of 5% formic acid (FA) and acetonitrile, extracts pulled together and dried in a vacuum centrifuge. Peptides were re-suspended in 25µl of 5% FA and 5µl aliquot was analysed by LC-MS/MS on a nanoUPLC system interfaced on-line to a Q Exactive HF Orbitrap mass spectrometer (both Thermo Fischer Scientific). The nanoUPLC was equipped with an Acclaim PepMap100 C18

75 µm i.d. x 20 mm trap column and 75 µm x 50 cm analytical column (3µm/100A, Thermo Fisher Scientific). Peptides were separated using a 80 min linear gradient; solvent A - 0.1% aqueous FA, solvent B - 0.1% FA in acetonitrile. Blank runs were introduced after each sample analysis to minimize carryover. Instrument performance was monitored with QCloud system (Chiva et al., 2018). Data were acquired using a Top 20 approach; precursor m/z range was 350-1600 and dynamic exclusion time was 20 s. The lock-mass function was set on the background ion (Si(CH3)2O)6 at m/z 445.12. Acquired spectra were converted into the .mgf format and merged into a single file for each sample.

Acquired data were processed with the MaxQuant software package (v.1.6.10.43) using default setting iBAC options, with Match-Between-Runs (MBR) disabled. Enzyme specificity was trypsin, number of allowed miscleavages – two; variable modification – cysteine carbamidomethyl, propionamide; methionine oxidation; protein N-terminus acetylated.

Mass photometry

Mass Photometry (MP, iSCAMS) of the FERRY complex was performed on a One^{MP} instrument (Refeyn, Oxford, UK) at room temperature. High precision 24 x 50 mm coverslips (Thorlabs CG15KH) were cleaned with ultrasound, rinsed with isopropanol and water and dried with clean nitrogen gas (Young et al., 2018). 20 µl diluted FERRY complex (43 and 34 nM, in PBS) was spotted into a reusable culture well gasket with 3 mm diameter and 1mm depth (Grace Bio-Labs). MP signals were recorded for 60 s at a suitable concentration in order to detect a sufficient set of target particles (>500). Raw MP data were processed in the DiscoverMP software (Refeyn, Oxford, UK).

Library preparation and Sequencing

mRNA was enriched from 100ng DNAse treated total RNA using the NEBNext rRNA depletion Kit (human, mouse, rat, NEB) according to the manufacturer's instructions. Final elution was done in 5 µl nuclease free water. Samples were then directly subjected to the workflow for strand specific RNA-Seq library preparation (NEBNext Ultra II Directional RNA Library Prep, NEB). 0.15 µM NEB Adaptor were used for ligation. Non-ligated adaptors were removed by adding XP beads (Beckmann Coulter) in a ratio of 1:0.9. Dual indexing was done during the following PCR enrichment (12 cycles, 65°C) using custom amplification primers carrying the index sequence indicated with 'NNNNNNN'. (Primer1: AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC T, primer2: CAA GCA GAA GAC GGC ATA CGA GAT NNNNNNNN GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC T). After two more XP bead purifications (1:0.9) libraries were quantified using the Fragment Analyzer (Agilent). Libraries were equimolarly pooled before sequencing them with a length of 75 bp in single end mode on an Illumina NextSeq 500 system to a depth of at least 2 x 10^7 reads.

Analysis of the mass spectrometry data

From the MaxQuant proteinGroups.txt file only protein groups with at least 1 unique peptide and which were identified in at least two out of three biological replicates in at least one condition were considered for differential abundance analysis using DEP v1.4.0 (Zhang et al., 2018). After variance stabilizing normalization (Huber et al., 2002) of iBAQ intensities, missing values were imputed applying the nearest neighbor averaging imputation method (KNN) to missing at random (MAR) and left-censored imputation using a deterministic minimal value approach (MinDet) to missing not at random (MNAR) protein groups (Gatto et al., 2021). MNARs refer to those protein groups with missing values in all replicates of one of the two conditions while all other missing values are considered as MAR. The application of empirical Bayes statistics on protein group-wise linear models was done using limma (Ritchie et al., 2015) and differentially abundant proteins were identified by applying a log2 fold change threshold of 1 and an adjusted p-value cutoff of 0.05.

Analysis of the RNA sequencing data

Raw reads were checked for their overall quality using FastQC v0.11.2 (Andrews, 2010). Read mapping to the human genome reference assembly (GRCh38_p13) and genes counts estimation based on Ensembl release v99 (Yates et al., 2020) were done using STAR v2.5.2b (--outFilterMultimapNmax 1 --outSJfilterCountUniqueMin 8 3 3 3 --quantMode GeneCounts; (Dobin et al., 2013) by taking read strandedness into account. Count data were filtered genes with more than 10 counts in any sample and served as input for differential gene expression analysis using DESeq2 v1.22.1 (Love et al., 2014). A log2-fold change threshold of 1 and an adjusted p-value cutoff of 0.01 was applied to FDRs obtained by using IHW v1.10.1 (Ignatiadis et al., 2016). Results summary in form of a MA plot was done using ggplot2 v3.2.1 (Wickham, 2016) following layout settings from the ggpubr package v0.2.5 (Kassambara, 2020).

Rab5 affinity chromatography

GST-Rab5 affinity chromatography was carried out as described before (Christoforidis et al., 1999). In summary, GST-Rab5:GDP or GST-Rab5:GTPγS loaded glutathione Sepharose was incubated with bovine brain cytosol, the beads extensively washed and the bound proteins subsequently eluted. The resulting mixture of Rab5 effector proteins was further purified by SEC and anion exchange chromatography. Fractions were analyzed using silver stained SDS PAGE.

In vitro translation binding assay

Binding assays with *in vitro* translated proteins were essentially performed as described (Nielsen et al., 2000). Briefly, [^{35}S]-methionine-labelled proteins were transcribed and translated using a TnT™ coupled transcription–translation kit (Promega) according to the manufacturer's protocol. Resulting proteins were incubated with GST-Rab5:GDP or GST-Rab5:GTPγS loaded Glutathione Sepharose for 2 h at 4 °C. Subsequently, the beads were washed and Rab5-bound proteins were eluted and analyzed by SDS PAGE and fluorography as described (Christoforidis et al., 1999).

mRNA production and electrophoretic motility shift assays (EMSA)

mRNA sequences for *mrpl41*, *mdh2*, *uchl1*, *atp5f1b*, *gstp1*, *prdx5*, *cox6b* and *cox8a* comprise the coding region, the 3' and 5' untranslated regions (UTRs) and an additional polyA appendix of 50 adenines (Table S2). The mRNAs were produced by *in vitro* transcription using the T7 RiboMAX™ Express Large Scale RNA Production System (Promega) according to the manufacturer's protocol. Resulting RNA was purified using a Phenol:Chloroform extraction and an isopropanol precipitation as described in the manual of the mMESSAGE mMACHINE™ T7 Transcription kit (Thermo Fisher). In brief, the *in vitro* transcription reactions were quenched with Ammonium acetate stop solution from the mMESSAGE mMACHINE™ T7 Transcription Kit (Thermo Fisher) and supplemented with Phenol:Chloroform:Isoamyl Alcohol 25:24:1 (Sigma Aldrich). The aqueous phase was recovered and RNA precipitated by adding equal amounts of isopropanol. After chilling at - 20 °C for at least 15 min, the precipitated RNA was pelleted by centrifugation (20 800 x g, 15 min, 4 °C), the supernatant removed and the pellet resuspended in RNAse-free water. RNA concentrations were determined by spectrophotometer (NanoDrop Lite, Thermo Scientific)' and the RNA was stored at - 80 °C until usage.

For direct RNA-protein interaction assays, 15 pmol of mRNA were mixed with protein (FERRY complex, Fy-4, Fy-5, Rab5:GTPγS or combinations) in varying protein/RNA ratios in SEC buffer in a total volume of 35 µl and subsequently incubated for 80 min at 37 °C. The samples were analyzed by ethidium bromide-stained gel electrophoresis using 1% agarose gels.

Generation of HeLa knockout (KO) cell lines

To generate gene knockouts in HeLa, we used CRISPR/Cas9 cleavage induced random (NHEJ mediated) mutations using guide RNAs targeted 5' end of the coding sequence of the genes of interest. We used electroporation of Cas9 protein complexed with crRNA and trRNAs (altR, IDT), using the Neon electroporator device and kits (Invitrogen) with concentrations and electroporation settings as previously described (Spiegel et al., 2019). For list of crRNA protospacers used for each gene, see the resources table (Table S2). The success of the gene

disruption was initially assessed by western blot of single cell derived clones. The disruption of the target alleles was further confirmed by fluorescent PCR and Sanger sequencing of PCR amplicons (For the genotyping primers used and description of the alleles, see the resources table (Table S2).

HeLa cell culture

Hela Kyoto and FERRY subunit knock-out cells were cultured in DMEM media supplemented with 10% FBS Superior (Merck) and 50 µg/ml streptomycin (P/S) (Gibco) at 37°C with 5% CO_2. For smFISH studies, cells were seeded into 384 well plates at a density of 3000 cells/well in 50 µl using the drop dispenser (Multidrop, Thermo Fischer Scientific) and cultured for 24h.

Single molecule fluorescence in situ hybridization (smFISH) and immunostaining

Endosomes and endogenous mRNAs were stained by using the ViewRNA® Cell Plus Assay kit (Invitrogen, 88-19000). The kit consists of 16 solutions that are used to perform an immuno-fluorescence staining followed by a single molecule fluorescence in situ hybridization (smFISH) using the sequential branched-DNA amplification technique. The manufactures protocol for 96 well plates was adapted to a 384 well plate format by down-scaling to 12.5 µl/well for steps containing staining solutions and to 25 µl/well for steps containing washing/fixing solutions (96 well protocol: 50 µl and 100 µl, respectively). For details see the manufactures protocol (https://assets.thermofisher.com/TFS-Assets/LSG/manuals/88-19000.pdf).

In brief, all steps were performed manually using an 8-channel aspirator for removal and automated multi-channel pipettes for addition of liquids. All wash steps following fixation and immunostaining were done 3 times with PBS including RNase inhibitor solution, whereas all wash steps following smFISH were done 5 times with RNA wash buffer solution. Cells were fixed and permeabilized using the provided solutions of the kit. After washing with PBS, cells were incubated with blocking buffer, primary antibody solution (including EEA1 and Fy-2 antibodies at a dilution of 1:2000 and 1:1000, respectively) and secondary solutions (including antibodies against rabbit and mouse IgG labelled with Alexa 488 or Alexa 568 (Alexa 647 for probe HPRT1), respectively, at a dilution of 1:500). After immunostaining cells were fixed and ready for smFISH. Different probes were used to label different mRNAs (Invitrogen, all probes were of type 6 (647nm), except the house-keeping gene HPRT1 (type 1, 546nm); *atp5f1b*: VA6-3168504, *mdh2*: VA6-3172506, *mrpl41*: VA6-3169863, *hprt1*: VA1-11124). Cells were incubated for 2h at 40°C with a diluted probe. After washing the cells with RNA wash buffer solution, the protocol was continued the next day with the smFISH branched-DNA amplification technique steps. Subsequently, cells were incubated with pre-amplifier, amplifier and label solution each for 1h at 40°C. Finally, the cells were stored in PBS containing DAPI (1µg/mL) to stain the nuclei and CellMaskBlue (CMB) (0.25µg/mL) to stain the cytoplasm.

Preparation of hippocampal cultures

Primary rat hippocampal neurons used in this study were obtained and cultured in two different ways. For initial Fy-2 localization experiments, the protocol for culturing hippocampal neurons was adapted from (Goto-Silva et al., 2019) with slight modifications. In brief, neurons were isolated from rat embryos at E17. The rat hippocampi from embryos of either sex were dissected in PBS (25 mM Na-phosphate buffer, pH 7.4, 110 mM NaCl, 1 mM EDTA) and dissociated in digestion solution (100 mg/ml DNAse I and 200 Units Papain in PBS) for 20 min. After two washes of the tissue with plating medium (DMEM containing 10% FCS, 2 mM glutamine, 50 mg/ml penicillin/streptomycin, Invitrogen), it was triturated in plating medium and subsequently cells counted. The neurons were plated on glass cover slips coated with 1 mg/ml poly-L-lysine (Sigma-Aldrich) at a density of 25 000 cells/ml in the presence of a mouse astrocyte feeder layer, derived from the mouse cortex from mice of age P0-P3 of either sex (Kaech and Banker, 2006).

Primary neurons for mRNA localization experiments were obtained and cultured according to the following protocol. Neuronal cultures were prepared from dissociated hippocampi of P0/P1 SD rats as previously described (Cajigas et al., 2012). Hippocampi were collected in Dissociation Medium on ice (DM with 1 mM HEPES, 82 mM Na_2SO_4, 30 mM K_2SO_4, 5.8 mM $MgCl_2$, 0.252 mM $CaCl_2$, 20 mM Glucose, 0.001% Phenol Red) and treated with cysteine-activated papain solution in DM (10 ml DM, 3.2 mg Cysteine, 300 µl Papain Sigma P3125, pH readjusted to 7, filtered sterile) two times 15 min at 37°C before several washes with cold DM and Neuronal growth medium (NGM: Neurobasal A supplemented with B27 and Glutamax). Dissociation of the tissue was achieved by trituration through a 10 ml pipette for 10 times. Before counting in a Neubauer chamber, cells were pelleted by centrifugation for 5 min, 67 x g at 4 °C, resuspended in cold NGM and 30 000 cells were seeded in 250 µl NGM on poly-D-Lys coated 14 mm MatTek glass bottom dishes. After attachment of the cells (2-3 h later) 0.7 ml conditioned NGM (80% NGM, 15% glia-conditioned NGM, 5% cortical neuron-conditioned NGM) was added and regular feeding by addition of NGM was performed thereafter. The neurons were kept in an incubator at 37°C in a humidified atmosphere with 5% CO_2.

Animals

The rat pups were used without gender determination. Timed pregnant rats were purchased from either Janvier (RjHan:WI - Wistar rats) or Charles River Laboratories, maintained under food and water ad libitum in a 12h-12h light dark cycle until pups were delivered, pups were sacrificed shortly after birth by decapitation with sharp scissors before dissection of the tissue. The procedures involving animal treatment and care were conducted in conformity with the institutional guidelines that are in compliance with the national and international laws and

policies (DIRECTIVE2010/63/EU; German animal welfare law, FELASA guidelines) and approved by and reported to the local governmental supervising authorities (Regierungspräsidium Darmstadt and Landesdirektion Sachsen). The animals were euthanized according to annex 2 of §2 Abs. 2 Tierschutz-Versuchstier-Verordnung.

Immunostaining of neurons

Immunostaining was performed at room temperature and the plates were subsequently stored at 4 °C if necessary. After adhesion, cells were washed once with PBS and fixed using 3% Paraformaldehyde (PFA) for 15 min. After washing with PBS, residual PFA was quenched using 500 mM Ammonium chloride in PBS for 10 min and the cells were washed 3 times with PBS. For permeabilization the cells were treated with 0.1% Triton X-100 in PBS for 3 min and subsequently washed three times with PBS. After blocking with 10% FBS for 20 min, the cells were incubated with the primary antibody for 2 h. Before and after the application of the secondary antibody for 1 h, the cells were washed 3 times with PBS.

High sensitivity FISH and immunostaining in neurons

In situ hybridization was performed using the ViewRNA ISH Cell Assay Kit (Thermo Fisher) according to the manufacturer's protocol with the modifications described previously (Cajigas et al., 2012). Probe sets targeting the respective mRNAs were purchased from Thermo Fisher. In brief, rat hippocampal neuron cultures grown for two weeks on MatTek glass bottom dishes were fixed for 20 min with PBS containing 1mM MgCl2, 0.1 mM CaCl2, 4% Sucrose and 4% PFA, pH 7.4 at room temperature, washed and subsequently permeabilized for 3 min with the provided detergent solution. Gene specific type1 (Uchl1) and type6 (Mdh2, polyA) probe sets were applied in 1:100 dilution for 3 h at 40°C. After several washes signal amplification steps with PreAmp/Amp and Label Probe reagents coupled to a 550 nm dye were all performed for 1 h at 40°C followed by washes at room temperature after each step. All probe sets and branched DNA reagents were diluted in the provided solutions 1:100. Immunostaining for Fy-2, endosome and mitochondria markers was performed after completion of the FISH protocol. FISH-stained cells were blocked for 30 min in blocking buffer (BB) at room temperature (BB: PBS with 4% goat serum) and incubated with primary antibodies in BB for 1 h at room temperature. After washing, secondary antibodies in BB were applied for 30 min, cells were washed and nuclei stained by a 3 min incubation with 1 µg/µl DAPI in PBS. Cells were washed in PBS and mounted with Aquapolymount (Polysciences).

Microscopy

automated HeLa imaging:

Confocal imaging was performed on an automated spinning disc confocal microscope (Yokogawa CV7000) using a 60x 1.2NA objective. DAPI and CMB was acquired with a laser excitation at 405 nm and an emission band pass filter BP445/45, Alexa 488 with a 488 nm laser and BP525/50 filter, Alexa 568 with a 561 nm laser and BP600/37 filter, Alexa 647 with a 640 nm laser and a BP676/29 filter. 9 fields were acquired per well as a stack with 4 z-planes and 1 µm distance. Each condition was done in duplicate wells and three independent experiments.

Spinning disk neuron imaging:

Neurons were imaged on a Nikon TiE spinning disk microscope equipped with a 100x/ 1.45NA Plan Apochromat, DIC oil immersion objective, Yokogawa CSU-X1 scan head and a Andor DU-897 back-illuminated CCD detector. Images were acquired with 600 ms exposure, while the laser intensities were adapted to the respective antibodies and requirements. Overview images of almost entire neurons were taken as a set of individual small images (6 x 6 images) with an overlap of 5% and combined using the Fiji implemented Grid/Collection Stitching tool (Preibisch et al., 2009) without overlap computation.

confocal neuron imaging:

Images were acquired with a LSM780 confocal microscope (Zeiss) equipped with Zen10 software using a 63x/1.46-NA oil objective (alpha Plan Apochromat 63×/1.46 oil DIC M27) and Argon 488, DPSS 561 and HeNe 633 laser lines for excitation in single tracks and a MBS488/561/633 beam splitter. Images were acquired in 12-bit mode as z-stacks and a time series with 4x Zoom, 512px x 512 px resolution and 0.1 µm Tetraspec beads (ThermoFisher) imaged under the same conditions. The laser power and detector gain in each channel was set to cover the full dynamic range but avoid saturated pixels.

Image analysis

HeLa cell images

Microscopy images for the localization of Fy-2, EEA1 and different mRNAs in HeLa cells were processed using the stand-alone freely available software MotionTracking (MT) (http://motiontracking.mpi-cbg.de). Images of were imported into MT and subsequently corrected for the chromatic shift of individual channels based on images of Tetraspec beads. For quantification, fluorescent foci of EEA1 and mRNA were detected using automated object detection and the co-localization was calculated based on 0.35 overlap threshold (Collinet et al., 2010; Kalaidzidis et al., 2015).

Neuron images

Microscopy images for the localization of Fy-2, EEA1, mRNA and mitochondria in neurons were also processed with MT. Image sequences of fixed neurons were imported into MT and drift corrected and deconvoluted by algorithms implemented in MT. In a last step, images were corrected for the chromatic shift of individual channels based on images of Tetraspec beads before and after the imaging. Motion Tracking implemented object detection was used to determine the mRNA foci while subsequent image analysis and quantification was performed by visual inspection. Given the possible distance between the fluorescence signals of EEA1 and mRNA or Fy-2 and mRNA (Figure S6), automated object detection followed by a co-localization analysis was not suitable for this purpose.

Western blotting

Cells were collected from a 10 cm cell culture dish, washed with cold PBS and subsequently lysed in PBS supplemented with 1% Triton X-100. HeLa cell lysates were clarified by centrifugation (14 000 rpm/ 20 800 x *g*, 15 min, 4 °C) and the concentration determined using a BCA assay (Pierce™ BCA Protein Assay Kit, Thermo Scientific). After running an SDS PAGE (12%), the gel was subsequently transferred onto a nitrocellulose membrane (Amersham). Blots were washed with PBST (PBS supplemented with 0.1% Tween 20) and then incubated with WB blocking buffer (5% non-fat milk powder in PBST) over night at 4 °C. After washing with PBST blots were then incubated with the primary antibodies (anti-Fy-1 to anti-Fy-5 and anti-GAPDH as a loading control) at the dilutions indicated earlier for 1 h at room temperature. After washing the secondary HRP antibody was applied to the blot for 1 h at room temperature. All antibodies were added in PBST with 5% milk. The blots were developed using ECL™ Western Blotting Reagents (Cytiva) on respective films (Amersham) in a Kodak X-OMAT 200 Processor.

References

Abou-Sleiman, P.M., Muqit, M.M., and Wood, N.W. (2006). Expanding insights of mitochondrial dysfunction in Parkinson's disease. Nat Rev Neurosci *7*, 207-219.

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

Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data http://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

Becalska, A.N., and Gavis, E.R. (2009). Lighting up mRNA localization in Drosophila oogenesis. Development *136*, 2493-2503.

1154 Beck-Wodl, S., Harzer, K., Sturm, M., Buchert, R., Riess, O., Mennel, H.D., Latta, E.,
1155 Pagenstecher, A., and Keber, U. (2018). Homozygous TBC1 domain-containing kinase
1156 (TBCK) mutation causes a novel lysosomal storage disease - a new type of neuronal ceroid
1157 lipofuscinosis (CLN15)? Acta Neuropathol Commun *6*, 145.

1158 Bhoj, E.J., Li, D., Harr, M., Edvardson, S., Elpeleg, O., Chisholm, E., Juusola, J., Douglas, G.,
1159 Guillen Sacoto, M.J., Siquier-Pernet, K.*, et al.* (2016). Mutations in TBCK, Encoding TBC1-
1160 Domain-Containing Kinase, Lead to a Recognizable Syndrome of Intellectual Disability and
1161 Hypotonia. Am J Hum Genet *98*, 782-788.

1162 Briese, M., Saal, L., Appenzeller, S., Moradi, M., Baluapuri, A., and Sendtner, M. (2016).
1163 Whole transcriptome profiling reveals the RNA content of motor axons. Nucleic Acids Res *44*,
1164 e33.

1165 Buxbaum, A.R., Haimovich, G., and Singer, R.H. (2015). In the right place at the right time:
1166 visualizing and understanding mRNA localization. Nat Rev Mol Cell Biol *16*, 95-109.

1167 Cajigas, I.J., Tushev, G., Will, T.J., tom Dieck, S., Fuerst, N., and Schuman, E.M. (2012). The
1168 local transcriptome in the synaptic neuropil revealed by deep sequencing and high-resolution
1169 imaging. Neuron *74*, 453-466.

1170 Capella-Gutierrez, S., Silla-Martinez, J.M., and Gabaldon, T. (2009). trimAl: a tool for
1171 automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics *25*, 1972-
1172 1973.

1173 Cezanne, A., Lauer, J., Solomatina, A., Sbalzarini, I.F., and Zerial, M. (2020). A non-linear
1174 system patterns Rab5 GTPase on the membrane. Elife *9*.

1175 Chernomor, O., von Haeseler, A., and Minh, B.Q. (2016). Terrace Aware Data Structure for
1176 Phylogenomic Inference from Supermatrices. Syst Biol *65*, 997-1008.

1177 Chiva, C., Olivella, R., Borras, E., Espadas, G., Pastor, O., Sole, A., and Sabido, E. (2018).
1178 QCloud: A cloud-based quality control system for mass spectrometry-based proteomics
1179 laboratories. PLoS One *13*, e0189209.

1180 Chong, J.X., Caputo, V., Phelps, I.G., Stella, L., Worgan, L., Dempsey, J.C., Nguyen, A.,
1181 Leuzzi, V., Webster, R., Pizzuti, A.*, et al.* (2016). Recessive Inactivating Mutations in TBCK,
1182 Encoding a Rab GTPase-Activating Protein, Cause Severe Infantile Syndromic
1183 Encephalopathy. Am J Hum Genet *98*, 772-781.

1184 Christoforidis, S., McBride, H.M., Burgoyne, R.D., and Zerial, M. (1999). The Rab5 effector
1185 EEA1 is a core component of endosome docking. Nature *397*, 621-625.

1186 Cioni, J.M., Koppers, M., and Holt, C.E. (2018). Molecular control of local translation in axon
1187 development and maintenance. Curr Opin Neurobiol *51*, 86-94.

1188 Cioni, J.M., Lin, J.Q., Holtermann, A.V., Koppers, M., Jakobs, M.A.H., Azizi, A., Turner-
1189 Bridger, B., Shigeoka, T., Franze, K., Harris, W.A.*, et al.* (2019). Late Endosomes Act as
1190 mRNA Translation Platforms and Sustain Mitochondria in Axons. Cell *176*, 56-72 e15.

1191 Collinet, C., Stoter, M., Bradshaw, C.R., Samusik, N., Rink, J.C., Kenski, D., Habermann, B.,
1192 Buchholz, F., Henschel, R., Mueller, M.S.*, et al.* (2010). Systems survey of endocytosis by
1193 multiparametric image analysis. Nature *464*, 243-249.

1194 Das, S., Vera, M., Gandin, V., Singer, R.H., and Tutucci, E. (2021). Intracellular mRNA
1195 transport and localized translation. Nat Rev Mol Cell Biol.

1196 Deguchi, M., Hata, Y., Takeuchi, M., Ide, N., Hirao, K., Yao, I., Irie, M., Toyoda, A., and
1197 Takai, Y. (1998). BEGAIN (brain-enriched guanylate kinase-associated protein), a novel
1198 neuronal PSD-95/SAP90-binding protein. J Biol Chem *273*, 26269-26272.

1199 Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson,
1200 M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*,
1201 15-21.

1202 Franke, C., Repnik, U., Segeletz, S., Brouilly, N., Kalaidzidis, Y., Verbavatz, J.M., and Zerial,
1203 M. (2019). Correlative single-molecule localization microscopy and electron tomography
1204 reveals endosome nanoscale domains. Traffic *20*, 601-617.

1205 Gatto, L., Gibb, S., and Rainer, J. (2021). MSnbase, Efficient and Elegant R-Based Processing
1206 and Visualization of Raw Mass Spectrometry Data. J Proteome Res *20*, 1063-1069.

1207 Glock, C., Heumuller, M., and Schuman, E.M. (2017). mRNA transport & local translation in
1208 neurons. Curr Opin Neurobiol *45*, 169-177.

1209 Goto-Silva, L., McShane, M.P., Salinas, S., Kalaidzidis, Y., Schiavo, G., and Zerial, M. (2019).
1210 Retrograde transport of Akt by a neuronal Rab5-APPL1 endosome. Sci Rep *9*, 2433.

1211 Guerreiro, R.J., Brown, R., Dian, D., de Goede, C., Bras, J., and Mole, S.E. (2016). Mutation
1212 of TBCK causes a rare recessive developmental disorder. Neurol Genet *2*, e76.

1213 Hancarova, M., Babikyan, D., Bendova, S., Midyan, S., Prchalova, D., Shahsuvaryan, G.,
1214 Stranecky, V., Sarkisian, T., and Sedlacek, Z. (2019). A novel variant of C12orf4 in a
1215 consanguineous Armenian family confirms the etiology of autosomal recessive intellectual
1216 disability type 66 with delineation of the phenotype. Mol Genet Genomic Med *7*, e865.

1217 Higuchi, Y., Ashwin, P., Roger, Y., and Steinberg, G. (2014). Early endosome motility
1218 spatially organizes polysome distribution. J Cell Biol *204*, 343-357.

1219 Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., and Vinh, L.S. (2018). UFBoot2:
1220 Improving the Ultrafast Bootstrap Approximation. Mol Biol Evol *35*, 518-522.

1221 Huber, W., von Heydebreck, A., Sultmann, H., Poustka, A., and Vingron, M. (2002). Variance
1222 stabilization applied to microarray data calibration and to the quantification of differential
1223 expression. Bioinformatics *18 Suppl 1*, S96-104.

1224 Ignatiadis, N., Klaus, B., Zaugg, J.B., and Huber, W. (2016). Data-driven hypothesis weighting
1225 increases detection power in genome-scale multiple testing. Nat Methods *13*, 577-580.

1226 Jung, H., Gkogkas, C.G., Sonenberg, N., and Holt, C.E. (2014). Remote control of gene
1227 function by local translation. Cell *157*, 26-40.

1228 Kaech, S., and Banker, G. (2006). Culturing hippocampal neurons. Nat Protoc *1*, 2406-2415.

1229 Kalaidzidis, Y., Kalaidzidis, I., and Zerial, M. (2015). A probabilistic method to quantify the
1230 colocalization of markers on intracellular vesicular structures visualized by light microscopy.

1231 Kassambara, A. (2020). ggpubr: 'ggplot2' Based Publication Ready Plots (https://cran.r-
1232 project.org/web/packages/ggpubr/index.html).

1233 Kim, E., and Jung, H. (2020). Local mRNA translation in long-term maintenance of axon
1234 health and function. Curr Opin Neurobiol *63*, 15-22.

1235 Lauer, J., Segeletz, S., Cezanne, A., Guaitoli, G., Raimondi, F., Gentzel, M., Alva, V., Habeck,
1236 M., Kalaidzidis, Y., Ueffing, M., *et al.* (2019). Auto-regulation of Rab5 GEF activity in Rabex5
1237 by allosteric structural changes, catalytic core dynamics and ubiquitin binding. Elife *8*.

1238 Liao, Y.C., Fernandopulle, M.S., Wang, G., Choi, H., Hao, L., Drerup, C.M., Patel, R., Qamar,
1239 S., Nixon-Abell, J., Shen, Y., *et al.* (2019). RNA Granules Hitchhike on Lysosomes for Long-
1240 Distance Transport, Using Annexin A11 as a Molecular Tether. Cell *179*, 147-164 e120.

1241 Lippe, R., Miaczynska, M., Rybin, V., Runge, A., and Zerial, M. (2001). Functional synergy
1242 between Rab5 effector Rabaptin-5 and exchange factor Rabex-5 when physically associated in
1243 a complex. Mol Biol Cell *12*, 2219-2228.

1244 Loddo, S., Alesi, V., Radio, F.C., Genovese, S., Di Tommaso, S., Calvieri, G., Orlando, V.,
1245 Bertini, E., Dentici, M.L., Novelli, A., *et al.* (2020). PPP1R21-related syndromic intellectual
1246 disability: Report of an adult patient and review. Am J Med Genet A *182*, 3014-3022.

1247 Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and
1248 dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.

1249 Martin, K.C., and Ephrussi, A. (2009). mRNA localization: gene expression in the spatial
1250 dimension. Cell *136*, 719-730.

1251 Moreira, P.I., Carvalho, C., Zhu, X., Smith, M.A., and Perry, G. (2010). Mitochondrial
1252 dysfunction is a trigger of Alzheimer's disease pathophysiology. Biochim Biophys Acta *1802*,
1253 2-10.

1254 Murray, D.H., Jahnel, M., Lauer, J., Avellaneda, M.J., Brouilly, N., Cezanne, A., Morales-
1255 Navarrete, H., Perini, E.D., Ferguson, C., Lupas, A.N., *et al.* (2016). An endosomal tether
1256 undergoes an entropic collapse to bring vesicles together. Nature *537*, 107-111.

1257 Nguyen, L.T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and
1258 effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol
1259 *32*, 268-274.

1260 Nielsen, E., Christoforidis, S., Uttenweiler-Joseph, S., Miaczynska, M., Dewitte, F., Wilm, M.,
1261 Hoflack, B., and Zerial, M. (2000). Rabenosyn-5, a novel Rab5 effector, is complexed with
1262 hVPS45 and recruited to endosomes through a FYVE finger domain. J Cell Biol *151*, 601-612.

1263 Nothwang, H.G., Kim, H.G., Aoki, J., Geisterfer, M., Kubart, S., Wegner, R.D., van Moers,
1264 A., Ashworth, L.K., Haaf, T., Bell, J., *et al.* (2001). Functional hemizygosity of PAFAH1B3
1265 due to a PAFAH1B3-CLK2 fusion gene in a female with mental retardation, ataxia and atrophy
1266 of the brain. Hum Mol Genet *10*, 797-806.

1267 Ortiz-Gonzalez, X.R., Tintos-Hernandez, J.A., Keller, K., Li, X., Foley, A.R., Bharucha-
1268 Goebel, D.X., Kessler, S.K., Yum, S.W., Crino, P.B., He, M., *et al.* (2018). Homozygous
1269 boricua TBCK mutation causes neurodegeneration and aberrant autophagy. Ann Neurol *83*,
1270 153-165.

1271 Pagel, M. (1994). Detecting Correlated Evolution on Phylogenies: A General Method for the
1272 Comparative Analysis of Discrete Characters. Proceedings of the Royal Society of London
1273 Series B *255*, 37.

1274 Paradis, E., and Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and
1275 evolutionary analyses in R. Bioinformatics *35*, 526-528.

1276 Park, J.S., Davis, R.L., and Sue, C.M. (2018). Mitochondrial Dysfunction in Parkinson's
1277 Disease: New Mechanistic Insights and Therapeutic Perspectives. Curr Neurol Neurosci Rep
1278 *18*, 21.

1279 Pfeffer, S.R. (2013). Rab GTPase regulation of membrane identity. Curr Opin Cell Biol *25*,
1280 414-419.

1281 Philips, A.K., Pinelli, M., de Bie, C.I., Mustonen, A., Maatta, T., Arts, H.H., Wu, K., Roepman,
1282 R., Moilanen, J.S., Raza, S., *et al.* (2017). Identification of C12orf4 as a gene for autosomal
1283 recessive intellectual disability. Clin Genet *91*, 100-105.

1284 Popovic, D., Nijenhuis, W., Kapitein, L.C., and Pelkmans, L. (2020). Co-translational targeting
1285 of transcripts to endosomes. bioRxiv.

1286 Preibisch, S., Saalfeld, S., and Tomancak, P. (2009). Globally optimal stitching of tiled 3D
1287 microscopic image acquisitions. Bioinformatics *25*, 1463-1465.

1288 Quentin, D., Schuhmacher, J.S., Klink, B.U., Lauer, J., Shaikh, T.R., Huis in't Veld, P.J., Welp,
1289 L., Urlaub, H., Zerial, M., and Raunser, S. (2021). Structure of the human FERRY Rab5
1290 effector complex.

1291 Rangaraju, V., Tom Dieck, S., and Schuman, E.M. (2017). Local translation in neuronal
1292 compartments: how local is local? EMBO Rep *18*, 693-711.

1293 Reddy, P.H., Mao, P., and Manczak, M. (2009). Mitochondrial structural and functional
1294 dynamics in Huntington's disease. Brain Res Rev *61*, 33-48.

1295 Rehman, A.U., Najafi, M., Kambouris, M., Al-Gazali, L., Makrythanasis, P., Rad, A.,
1296 Maroofian, R., Rajab, A., Stark, Z., Hunter, J.V., *et al.* (2019). Biallelic loss of function variants
1297 in PPP1R21 cause a neurodevelopmental syndrome with impaired endocytic function. Hum
1298 Mutat *40*, 267-280.

1299 Riechmann, V., and Ephrussi, A. (2001). Axis formation during Drosophila oogenesis. Curr
1300 Opin Genet Dev *11*, 374-383.

1301 Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015).
1302 limma powers differential expression analyses for RNA-sequencing and microarray studies.
1303 Nucleic Acids Res *43*, e47.

1304 Rozewicki, J., Li, S., Amada, K.M., Standley, D.M., and Katoh, K. (2019). MAFFT-DASH:
1305 integrated protein sequence and structural alignment. Nucleic Acids Res *47*, W5-W10.

1306 Schnatwinkel, C., Christoforidis, S., Lindsay, M.R., Uttenweiler-Joseph, S., Wilm, M., Parton,
1307 R.G., and Zerial, M. (2004). The Rab5 effector Rabankyrin-5 regulates and coordinates
1308 different endocytic mechanisms. PLoS Biol *2*, E261.

1309 Sleister, H.M., and Rao, A.G. (2001). Strategies to generate antibodies capable of
1310 distinguishing between proteins with >90% amino acid identity. J Immunol Methods *252*, 121-
1311 129.

1312 Spiegel, A., Bachmann, M., Jurado Jimenez, G., and Sarov, M. (2019). CRISPR/Cas9-based
1313 knockout pipeline for reverse genetics in mammalian cell culture. Methods *164-165*, 49-58.

1314 Suleiman, J., Al Hashem, A.M., Tabarki, B., Al-Thihli, K., Bi, W., and El-Hattab, A.W. (2018).
1315 PPP1R21 homozygous null variants associated with developmental delay, muscle weakness,
1316 distinctive facial features, and brain abnormalities. Clin Genet *94*, 351-355.

1317 Tabari, E., and Su, Z. (2017). PorthoMCL: Parallel orthology prediction using MCL for the
1318 realm of massive genome availability. Big Data Anal *2*.

1319 Team, R.C. (2019). R: A language and environment for statistical computing. R Foundation
1320 for Statistical Computing.

1321 Turner-Bridger, B., Caterino, C., and Cioni, J.M. (2020). Molecular mechanisms behind
1322 mRNA localization in axons. Open Biol *10*, 200177.

1323 UniProt, C. (2019). UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res *47*,
1324 D506-D515.

1325 Wandinger-Ness, A., and Zerial, M. (2014). Rab proteins and the compartmentalization of the
1326 endosomal system. Cold Spring Harb Perspect Biol *6*, a022616.

1327 Whelan, S., and Goldman, N. (2001). A general empirical model of protein evolution derived
1328 from multiple protein families using a maximum-likelihood approach. Mol Biol Evol *18*, 691-
1329 699.

1330 Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag New
1331 York).

1332 Wilson, J.M., de Hoop, M., Zorzi, N., Toh, B.H., Dotti, C.G., and Parton, R.G. (2000). EEA1,
1333 a tethering protein of the early sorting endosome, shows a polarized distribution in
1334 hippocampal neurons, epithelial cells, and fibroblasts. Mol Biol Cell *11*, 2657-2671.

1335 Yates, A.D., Achuthan, P., Akanni, W., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M.R.,
1336 Armean, I.M., Azov, A.G., Bennett, R.*, et al.* (2020). Ensembl 2020. Nucleic Acids Res *48*,
1337 D682-D688.

1338 Young, G., Hundt, N., Cole, D., Fineberg, A., Andrecka, J., Tyler, A., Olerinyova, A., Ansari,
1339 A., Marklund, E.G., Collier, M.P.*, et al.* (2018). Quantitative mass imaging of single biological
1340 macromolecules. Science *360*, 423-427.

1341 Yu, G., Lam, T.T., Zhu, H., and Guan, Y. (2018). Two Methods for Mapping and Visualizing
1342 Associated Data on Phylogeny Using Ggtree. Mol Biol Evol *35*, 3041-3043.

1343 Yu, G., Smith, D.K., Zhu, H., Guan, Y., and Lam, T.T.-Y. (2017). ggtree: an r package for
1344 visualization and annotation of phylogenetic trees with their covariates and other associated
1345 data. Methods in Ecology and Evolution *8*, 28-36.

1346 Zapata-Aldana, E., Kim, D.D., Remtulla, S., Prasad, C., Nguyen, C.T., and Campbell, C.
1347 (2019). Further delineation of TBCK - Infantile hypotonia with psychomotor retardation and
1348 characteristic facies type 3. Eur J Med Genet *62*, 273-277.

1349 Zhang, J., Reiling, C., Reinecke, J.B., Prislan, I., Marky, L.A., Sorgen, P.L., Naslavsky, N.,
1350 and Caplan, S. (2012). Rabankyrin-5 interacts with EHD1 and Vps26 to regulate endocytic
1351 trafficking and retromer function. Traffic *13*, 745-757.

1352 Zhang, X., Smits, A.H., van Tilburg, G.B., Ovaa, H., Huber, W., and Vermeulen, M. (2018).
1353 Proteome-wide identification of ubiquitin interactions using UbIA-MS. Nat Protoc *13*, 530-
1354 550.

1355