THE
EMBO
JOURNAL

SCA8 RAN polySer protein preferentially accumulates in white matter regions and is regulated by eIF3F

Fatma Ayhan, Barbara A. Perez, Hannah K Shorrock, Tao Zu, Monica Banez-Coronel, Tammy Reid, Hirokazu Furuya, H. Brent Clark, Juan C. Troncoso, Christopher A. Ross, S.H. Subramony, Tetsuo Ashizawa, Eric T. Wang, Anthony T. Yachnis, Laura P.W. Ranum

Review timeline:

Submission date:	16th Jan 18
Editorial Decision:	19th Feb 18
Revision received:	13th Jun 18
Editorial Decision:	18th Jul 18
Revision received:	31st Jul 18
Accepted:	2nd Aug 18

Editor: Karin Dumstrei

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 19th Feb 18

Thank you for submitting your manuscript to The EMBO Journal. Your study has now been seen by three referees and their comments are provided below.

As you can see from the comments, the referees find the analysis interesting and insightful. They raise a number of different comments that I would like to ask you to address in a revised version. As you can see, some of the concerns overlap and in general they are very constructive. I anticipate that you should be able to address them, but let me know if we need to discuss anything further.

I should add that it is EMBO Journal policy to allow only a single major round of revision and it is therefore important to address the concerns raised at this stage.

When preparing your letter of response to the referees' comments, please bear in mind that this will form part of the Review Process File, and will therefore be available online to the community. For more details on our Transparent Editorial Process, please visit our website:
http://emboj.embopress.org/about#Transparent_Process

We generally allow three months as standard revision time. As a matter of policy, competing manuscripts published during this period will not negatively impact on our assessment of the conceptual advance presented by your study. However, please contact me as possible upon publication of any related work, to discuss how to proceed. Should you foresee a problem in meeting this three-month deadline, please let me know and I can extend the deadline.

Thank you for the opportunity to consider your work for publication. I look forward to your revision.

REFEREE COMMENTS

Referee #1:

Ayhan et al. analyze neuropathology and disease mechanisms in a mouse model of SCA8 and patient tissue. These mice express at least polyGln (ATG initiated), polySer (RAN-product in another reading frame, focus of this study) and polyAla (RAN-product from third reading frame, reported previously in the same mice, Zu et al., PNAS 2011). The most exciting conclusion is a link of polySer accumulating in white matter and demyelination, because it identifies potentially explains data from human SCA8 patients. However, more data is needed to strengthened this claim. In addition, they propose that enrichment of polySer in white matter is due to differential expression of eIF3F, which seems to modulate RAN-translation. This part is also exciting, what seems premature at the moment. My recommendation would be to focus on the polySer mechanisms for this manuscript.

Major points:
-Fig EV1: Two polySer antibodies detecting the c-terminal tail are shown in EV1, but it is not clear, which one is used in the main figures. Presence of polySer in human tissue should be confirmed with both antibodies
-The authors' claim that RAN-mediated polySer and ATG-mediated polyGln have radically distinct patterns of accumulation, in the white matter and neuronal nuclei respectively, is not currently well-supported. In Figure 1, the authors show abundant polySer staining in the hippocampal CA region, which is not particularly axon-rich. Also, while polyGln and polySer do not colocalize in Fig. 2B, they do appear frequently near each other in the same region and often apparently the same cell. In the areas which are dominated by white matter, it remains unclear whether the signals are originating in neuronal axons or oligodendrocytes. The NeuN staining in Fig 2C looks very different in both panels and is thus not that informative. Is polySer mainly aggregating in nuclei in Fig 2D? The authors should determine whether the white matter signal of polySer originates from neurons or oligodendrocytes, for example by co-staining with neuronal and oligodendrocytes markers (e.g. NFH and MOG) on the RNA or protein level. FACS might also be helpful. Is the ATXN8 transcript expressed in oligodendrocytes? This is important for the understanding of the disease mechanism and for the interpretation of the eIF3F data.
-The claim that white matter abnormalities occur at sites of polySer accumulation is reasonably supported. However, to argue that white matter loss is tied to polySer expression, the authors should also show the white matter staining in areas where polySer is not abundant, for both mouse and human samples. Moreover, the glial expression and regional distribution of polyGln, polyAla and ideally RNA foci should be analyzed.
-Modulation of RAN-translation (for polySer or potentially across diseases) by eIF3F is interesting, but needs further validation. Does higher expression of eIF3 (as reported for white matter) indeed boost RAN-translation? All three reading frames should be addressed for the different constructs - I assume 3T means triple-tag like in Zu et al (PNAS 2011). Most importantly, the authors need to verify the differential expression of eIF3F in glia. The cited reference reports exclusive expression of white but not in grey matter. In contrast, the RNAseq data from the Barres lab shows only modest difference between OPC and neurons (http://web.stanford.edu/group/barres_lab/cgi-bin/geneSearch.py?geneNameIn=eif3f) and the expression pattern in Allen brain atlas looks rather neuronal (http://mouse.brain-map.org/experiment/show/69671953).

Minor points:
-Fig 2C: the figure legend for this section also appears incorrect and should be revised.
-SMI-32 is labeled a phospho-NFH antibody in Fig. 4, but in fact labels total NFH.
-Fig 5 The complicated nomenclature of the different constructs makes Fig 5 a bit confusing.
-Table 2: Effect size should be given for all parameters.

Referee #2:

The manuscript by Ayhan et al describes polySer and polyGln protein accumulation in SCA8 brains from both mice models and human patients. They show regional specificity and that the RAN products accumulate to a greater degree with disease progression. They also perform cell studies to show a correlation between RAN products and eIF3F levels, although at this point causality can only be inferred.
Overall this is an important body of work and there are several suggestions and concerns.

1. In figure 1 it's not clear what alpha-SerCT stands for. It took some digging to find that it may mean the unique C-terminus downstream of the polySer sequence - a predicted RAN product. Some definition in the legend and text is needed.

2. The drop in gait and open field can also reflect muscle pathology - are there RAN products in the muscle or any muscle pathology?

3. The intriguing finding is the cell specificity, which the authors then address by looking for RAN-TAFs, but then they pick a very common target in eIF3F. This is the part of the study that is most problematic - with several concerns.

a. How come the same repeat sequence with an ATG start cannot form a hairpin? It would seem that everything downstream of the methionine would have the same propensity to fold, particularly with that long of a repeat.
b. What cells were used in the KD experiments? This was not explained in the text or legend. How relevant is the cell type? I don't understand how this gets at the specificity point the authors raise - did the authors see in tissues that RAN products and eIF3F expression levels correlate? Does eIF3F interact with the RAN transcripts in RIP studies?
c. The data show a correlation between eIF3F and RAN translation products, but not causality. More studies are needed to conclude that one directly impacts the other. The authors state in the abstract, the results and the discussion that eIF3F may be a therapeutic target but these conclusions should be softened in the absence of additional studies. One could also argue that the data do not show that eIF3F selectively regulates RAN translation - only that reducing it impacts RAN translation products, but there could be many things happening in a pathway linking those two processes as well as other effects of the KD beyond what was examined here.

Methods -
Describe what cells were used and how many times studies were repeated (biological as well as technical replicates).
Was the open field all dark? There was not a light and dark chamber? The assay setup seems unusual. Also for the gait analysis, what parameters were measured? This should be stated as there are numerous types of output for this equipment.

Describe the statistical analysis for the RNA-Seq expression comparison studies.

Referee #3:

Spinocerebellar ataxia type 8 (SCA8) is a dominantly inherited neurodegenerative disease caused by a CTG•CAG microsatellite expansion mutation. The repeats can lead to accumulation of multiple potentially toxic products: RNA foci, polyGln (by ATG translation), polyAla and polySer (both by repeat associated non-ATG (RAN) translation). This manuscript by Ayhan et al. showed the SCA8 RAN polySer protein preferentially accumulates in white matter regions and increases with age and disease severity, accompanied with demyelination and axonal degeneration. The production of RAN proteins polySer and polyAla can be reduced by eIF3F knockdown, a potential therapeutic target. This study is important for our understanding of the pathogenic contribution of RAN translation products to neurodegeneration diseases. However, several evidence need to be further investigated to provide more convincing support.

Major concerns:
1. Only one of the RAN translation product, polySer, is examined in the whole work, except eIF3F is shown to regulate both in the last figure. PolyAla probably has similar expression patterns in the different brain regions and during aging as polySer, and also contribute to the toxicity. Since the lab has the antibody for polyAla (shown in previous publication Zu, Gibbens et al., 2011), the polyAla expression should also be systematically analyzed and compared with polySer.
2. Fig.2 showed different expression/distribution of polySer and polyGln proteins in different brain regions in both SCA8 BAC mouse and human autopsy samples using immunohistochemistry. This method only allows the detection of aggregated forms of the polypeptide. Elisa or dot blot (such as in Fig.5C) of total lysates should also be examined to exclude the possibility that different polypeptides have different solubility or aggregation properties in different regions or cell types.
3. What's the expression level of ATXN8 in different brain regions? Does it correlate with the polyGln aggregates? Why does polyGln have lower expression in white matter? The RNA and protein levels of ATXN8 should be measured at these different brain regions.
4. Fig.2: it is described that polyGln staining is primarily nuclear, but polySer shows perinuclear localization or localized in neuropil (shown in 2C). However, in Fig.2C, most of the polySer is not co-stained with NeuN and looks more like it is expressed in other cell types or extracellularly rather than in neuropil. Co-staining with glia cell markers (astrocyte, oligodendrocyte, microglia) should be included.
5. There is inconsistency of polySer staining between human patient autopsy tissue and SCA8 mouse tissues. It is shown in mouse samples that polySer has perinuclear localization. But in Fig.2D and 4B, it seems the SerCT actually shows staining mostly in the nucleus in human patient autopsy tissue. Does the mouse model really reproduce the human phenotypes?
6. It is nicely shown that the polySer accumulation increases with age and severity of disease in Fig.3. But the brain region is not consistent with ones examined in Fig.2. For example, cerebellar white matter and molecular layer were not examined during aging. This should be included. And quantifications are needed for 3A. In addition, how polyGln accumulation changes during aging should also be examined.
7. In Fig.4E, the RAN translation construct should also be used to compare the toxicity with ATG alt Gln. The relative polySer expression level from the repeat construct (RAN translation) vs ATG Alt Ser construct should be compared and included. Although the current data showed polySer is slightly more toxic than polyGln, but the expression from RAN repeats is probably much lower than AUG translation, therefore leads to lower toxicity in the real context. Furthermore, there is no data on polyAla (neither RAN construct nor ATG version). But in the last sentence of the second paragraph on page 11, it said "polySer expansion proteins are toxic to glial cells independent of the CAG expansion transcripts or polyGln or polyAla proteins".
8. Fig.5A: In order to show the correlation of eif3f levels with polySer, the Eif3f transcript levels in different brain regions should be examined.
9. What's the timing of white matter abnormalities, compared to neuronal defects? This should be evaluated to understand how much the white matter abnormalities could contribute to the disease progression.
Minor points:
1. There are two antibodies targeting SerCT produced, shown in Fig.EV1A,B. But in all the rest of figures in the manuscript, it is not mentioned which antibody was used and whether both antibodies showed similar results.
2. Fig.EV2A: Protein levels from the different constructs should also be compared besides RNA levels.
3. In the abstract and introduction (last sentence of the first paragraph), the authors only mentioned RAN polyAla from the SCA8 CAG expansion as a known fact. But polySer has been shown in previous report (Zu, Gibbens et al., 2011) as well, using in vitro assays. The text should be modified to describe previously known findings more precisely.
4. On page 15, last sentence of the second paragraph, "will be" is duplicated.

1st Revision - authors' response 13th Jun 18

We are grateful to you and the reviewers for the careful and detailed evaluations of our manuscript and for the helpful suggestions. We have now revised the manuscript to address the reviewer concerns. Specifically, we now include additional data demonstrating eIF3F expression is increased in oligodendrocyte enriched white matter brain regions of SCA8 BAC and have performed

additional requested control experiments. We have also softened our language regarding the potential clinical impact of eIF3F as requested. A detailed summary of our responses is below.

Referee #1:

Ayhan et al. analyze neuropathology and disease mechanisms in a mouse model of SCA8 and patient tissue. These mice express at least polyGln (ATG initiated), polySer (RAN-product in another reading frame, focus of this study) and polyAla (RAN-product from third reading frame, reported previously in the same mice, Zu et al., PNAS 2011). The most exciting conclusion is a link of polySer accumulating in white matter and demyelination, because it identifies potentially explains data from human SCA8 patients. However, more data is needed to strengthen this claim. In addition, they propose that enrichment of polySer in white matter is due to differential expression of eIF3F, which seems to modulate RAN-translation. This part is also exciting, what seems premature at the moment. My recommendation would be to focus on the polySer mechanisms for this manuscript.

Major points:
-Fig EV1: Two polySer antibodies detecting the c-terminal tail are shown in EV1, but it is not clear, which one is used in the main figures. Presence of polySer in human tissue should be confirmed with both antibodies.

Response: The antibodies used are now specified for each figure and IHC staining of human tissue with both antibodies is now shown in Figure EV1 D and E. We have clarified this in the results section of the manuscript.

Page 6, line 10-14: "Although both α-SerCT antibodies showed similar punctate straining, α-SerCT was used for IHC analyses of SCA8 BAC mouse tissue as it showed less background reactivity. In SCA8 BAC mice, we detected widespread punctate aggregates of variable size in brain regions primarily affected in the disease, including the cerebellum and brainstem (Figure 1D)."

Page 6, line 20-22: "Both antibodies were also able to detect polySer aggregates in patient autopsy tissue. However, because α-SerCT2 showed less non-specific reactivity in human tissue it was used for subsequent IHC on human tissue (Figure EV1E)."

-The authors' claim that RAN-mediated polySer and ATG-mediated polyGln have radically distinct patterns of accumulation, in the white matter and neuronal nuclei respectively, is not currently well-supported. In Figure 1, the authors show abundant polySer staining in the hippocampal CA region, which is not particularly axon-rich.
-Also, while polyGln and polySer do not colocalize in Fig. 2B, they do appear frequently near each other in the same region and often apparently the same cell.

Response: Because of the unique organization of the cerebellum where Purkinje cell body, axons, and dendrites are located in PC, molecular, and white matter layers respectively; the layer specific distribution patterns of polyGln and polySer in the cerebellum are seen very clearly. In the cortex and hippocampus the polyGln and polySer aggregates are found in adjacent regions, possibly because cell bodies and processes in these brain regions are not as discretely separated within different layers. While the dentate gyrus is not a particularly axon rich area, it receives projections from the entorhinal cortex. It is possible that polySer accumulates within processes in these regions as well. Future studies including different regions of the brain and electron microscopy to more specifically define where the polySer aggregates localize in regions that have less clear separation of cellular compartments would be interesting as a follow-up study, which we feel is beyond the scope of this manuscript. We have now revised the text.

Page 7, lines 10-20: "Although both proteins are expressed from ATXN8 transcripts, their distribution patterns in the cerebellum are strikingly different. IHC performed on serial cerebellar sections shows polyGln, but not RAN polySer aggregates accumulate in Purkinje cell nuclei. In contrast, we detected polySer but not polyGln aggregates in the molecular layer and deep cerebellar white matter, either within the neuropil or as perinuclear aggregates (Figure 2A). In the cortex, hippocampus and brainstem, polyGln and polySer aggregates are often found in close proximity, however they are detected as discrete aggregates. Double-labeling of polyGln and RAN polySer by IF shows no overlap in signal in the frontal cortex, brainstem (pons) or hippocampus. These data

suggest that polySer and polyGln aggregates are found in different cells or different cellular compartments in these brain regions."

- In the areas which are dominated by white matter, it remains unclear whether the signals are originating in neuronal axons or oligodendrocytes.

Response: We agree with the reviewer that the polySer aggregates localized in white matter regions may be located within axonal processes or oligodendrocytes. In order to further clarify this, we performed CNPase IF staining and show that polySer staining appears as perinuclear aggregates or as aggregates located close to or within oligodendrocytes/myelinated axons (Figure EV2). We have included this as an additional supplemental figure and now clarify the localization of polySer aggregates in the text.

Page 8, lines 1-4: "To further clarify the cellular localization of RAN polySer, we performed double-labeling IF of polySer and CNPase, a marker of oligodendrocytes, and show that polySer aggregates located in the neuropil are close to or within oligodendrocytes/myelinated axons (Figure EV2)."

-The NeuN staining in Fig 2C looks very different in both panels and is thus not that informative. Is polySer mainly aggregating in nuclei in Fig 2D?

Response: We agree with the reviewer that the NeuN staining looks different between the 1C2 and polySer double-labelling IFs. This is because the polySerCT antibody was raised in rabbit and the 1C2 antibody was raised in mouse, the NeuN antibodies used for co-staining were raised in different species to allow for secondary antibody detection. This has now been clarified in the methods and figure legend. We also now discuss the different appearance of staining in the human vs. mouse which may be affected by post-mortem interval and more severe end-stage disease in humans.

Page 8, Lines 7-11: "Despite this, there are some differences in the appearance of polySer in human postmortem tissue and SCA8 BAC mouse brain. These differences may be due to the end stage of disease in the human cases and to differences in tissue handling and processing including postmortem delay and the severity of disease at endpoint in human cases."

Methods – page 25, lines 20-23: "Mouse anti-NeuN antibody (ABN78A4, Millipore Sigma, Burlington, MA) was used for co-staining with custom rabbit polyclonal Ser-CT antibody and rabbit anti-NeuN antibody (ab104225, Abcam, Cambridge, MA) was used for co-staining with mouse 1C2 antibody (MAB1574, Millipore Sigma, Burlington, MA)."

Figure 2C legend – page 36, lines 43-47: "IF double staining of end-stage SCA8 BAC frontal cortex shows exclusive localization of polyGln (mouse α-Gln, red, bottom panel) in neurons (rabbit α-NeuN, green, bottom panel). In contrast, polySer (Rabbit α-SerCT, red, top panel) shows widespread accumulation in the frontal cortex including within neurons (mouse α-NeuN, green, top panel)."

- The authors should determine whether the white matter signal of polySer originates from neurons or oligodendrocytes, for example by co-staining with neuronal and oligodendrocytes markers (e.g. NFH and MOG) on the RNA or protein level. FACS might also be helpful. Is the ATXN8 transcript expressed in oligodendrocytes? This is important for the understanding of the disease mechanism and for the interpretation of the eIF3F data.

Response: We have now performed CNPase IF staining and show that polySer appears as perinuclear aggregates or as aggregates located close to or within oligodendrocytes/myelinated axons (Figure EV2). This distribution is distinct from the polyGln protein which appears as nuclear aggregates (Figure2). While the reviewer raises an important point, the current system for dissociating neuronal cell types are efficient in neonatal tissue but challenging and confounded by abundant myelin debris in adult neuronal tissue. While this will be an important part of future experiments and could be explored by generating a cell type reporter SCA8 mouse line, this is beyond the scope of the initial characterization of white matter abnormalities and polySer in SCA8

mice. To address this concern we now show that ATXN8 transcripts are expressed in both white and grey matter. We have now clarify these points in the text.

Page 8, lines 1-4: "To further clarify the cellular localization of RAN polySer, we performed double-labeling IF of polySer and CNPase, a marker of oligodendrocytes, and show that polySer aggregates located in the neuropil are close to or within oligodendrocytes/myelinated axons (Figure EV2)."

Page 14, lines 5-6: "We also evaluated levels of ATXN8 RNA in cerebellar grey and white matter and found no significant difference in the expression of ATXN8 transcripts (Figure EV5B),…"

-The claim that white matter abnormalities occur at sites of polySer accumulation is reasonably supported. However, to argue that white matter loss is tied to polySer expression, the authors should also show the white matter staining in areas where polySer is not abundant, for both mouse and human samples.

Response: We have now included images of unaffected white matter regions in mouse and human neural tissue and demonstrate that regions where polySer is not abundant do not show demyelination (Figure EV3). We have clarified this in the text.

Page 11, lines 9-11: "In contrast, brain regions that did not show polySer aggregates in SCA8 BAC mouse cerebellum (Figure EV3A) and human cortical white matter (Figure EV3B), did not show evidence of demyelination."

-Moreover, the glial expression and regional distribution of polyGln, polyAla and ideally RNA foci should be analyzed.

Response: Previous reports have shown that polyGln is found in Purkinje cell and pontine neuronal nuclei (Moseley et al 2006). This paper has extended these results and now show the regional distribution of polyGln with aggregates present in the frontal cortex, pons and hippocampus (Figure 2B). Furthermore, co-staining with NeuN indicates the neuronal localization of polyGln aggregates (Figure 2C). The regional distribution of RNA foci has been previously reported for the cerebellum, with foci in molecular layer interneurons, Purkinje cells and Bergmann glia (Daughters et al 2009); aside from this, RNA foci are not abundant and are therefore difficult to detect. Zu et al 2011 showed that polyAla is found in Purkinje cells, however, due to background issues in other brain regions, the polyAla antibody was not useful in this study for determining the overall brain distribution of polyAla and requires further optimization. A paragraph discussing this has been added to the discussion. Therefore, while we agree in principle with the reviewer, we suggest these additional detailed studies would be appropriate for a follow-up study in a more specialized journal.

Page 16, lines 10-20: "In addition to our current report of polySer, we previously reported the accumulation of RNA foci, polyAla and polyGln proteins in SCA8 BAC mice and patient autopsy tissue. Initially, it was demonstrated that polyGln accumulates in Purkinje and pontine neuronal nuclei (Moseley et al., 2006). The present work extends the molecular characterization of SCA8 by showing polyGln aggregates are also found in the frontal cortex, pons and hippocampus (Figure 2B), and are consistently located neuronal nuclei (Figure 2C). RAN PolyAla staining was previously reported in Purkinje cell soma and dendritic processes (Zu et al., 2011), and CUG RNA foci were detected in Purkinje cells, molecular layer interneurons, and Bergmann glia (Daughters et al., 2009). Additional studies are needed to investigate the brain distribution of polyAla and possible antisense RAN proteins."

-Modulation of RAN-translation (for polySer or potentially across diseases) by eIF3F is interesting, but needs further validation. Does higher expression of eIF3F (as reported for white matter) indeed boost RAN-translation? All three reading frames should be addressed for the different constructs - I assume 3T means triple-tag like in Zu et al (PNAS 2011). Most importantly, the authors need to verify the differential expression of eIF3F in glia. The cited reference reports exclusive expression of white but not in grey matter. In contrast, the RNAseq data from the Barres lab shows only modest difference between OPC and neurons (http://web.stanford.edu/group/barres_lab/cgi-bin/geneSearch.py?geneNameIn=eif3f) and the expression pattern in Allen brain atlas looks rather

neuronal (http://mouse.brain-map.org/experiment/show/69671953).

(Part 1) All three reading frames should be addressed for the different constructs - I assume 3T means triple-tag like in Zu et al (PNAS 2011). Modulation of RAN-translation (for polySer or potentially across diseases) by eIF3F is interesting, but needs further validation.
Response: Firstly, we clarify that the A8-3T construct is the same construct used in Zu et al 2011. We addressed the polySer and the polyAla frames because the endogenous transcripts do not have an AUG initiation codon. We did not look in detail at the polyGln frame because the endogenous transcripts have an AUG initiation codon. This makes the experiment using the endogenous transcript technically complicated as the products have similar molecular weights, and biologically less relevant.

(Part 2) Does higher expression of eIF3F (as reported for white matter) indeed boost RAN-translation? Most importantly, the authors need to verify the differential expression of eIF3F in glia. The cited reference reports exclusive expression of white but not in grey matter. In contrast, the RNAseq data from the Barres lab shows only modest difference between OPC and neurons (http://web.stanford.edu/group/barres_lab/cgi-bin/geneSearch.py?geneNameIn=eif3f) and the expression pattern in Allen brain atlas looks rather neuronal (http://mouse.brain-map.org/experiment/show/69671953)
Response: We have also now included quantitative RT-PCRs comparing eIF3F expression in tease separated, oligodendrocyte rich white matter versus grey matter of SCA8 BAC mouse neural tissue. These experiments show that eIF3F is more highly expressed in white matter compared to grey matter (Figure 5A), which is consistent with the possibility that eIF3F increases RAN translation.

Page 13, lines 21-23 and page 14 lines 1-11: "To investigate if Eif3f expression is higher in white matter than in grey matter, we performed qRT-PCR comparing tease-separated cerebellar white and grey matter from SCA8 BAC mice. To confirm an enrichment of oligodendrocytes in white matter samples, myelin basic protein (MBP) was measured by qRT-PCR. After confirmation of MBP white matter enrichment (Figure EV5A) we performed qRT-PCR and show Eif3f transcripts have a two-fold increase in SCA8 BAC cerebellar white matter compared to grey matter (Figure 5A). We also evaluated levels of ATXN8 RNA in cerebellar grey and white matter and found no significant difference in the expression of ATXN8 transcripts (Figure EV5B), indicating that changes in ATXN8 RNA expression are unlikely to contribute to the increased polySer aggregate accumulation in white matter regions. The elevated Eif3f levels observed in cerebellar white matter, which shows robust polySer aggregation, are consistent with the possibility that eIF3F increases RAN translation leading to the accumulation of RAN polySer protein in white matter."

Minor points:
-Fig 2C: the figure legend for this section also appears incorrect and should be revised.

Response: We have now corrected the figure legend for Figure 2C.

Page 36, lines 43-47: "IF double staining of end-stage SCA8 BAC frontal cortex shows exclusive localization of polyGln (mouse α-Gln, red, bottom panel) in neurons (rabbit α-NeuN, green, bottom panel). In contrast, polySer (Rabbit α-SerCT, red, top panel) shows widespread accumulation in the frontal cortex including within neurons (mouse α-NeuN, green, top panel)."

-SMI-32 is labeled a phospho-NFH antibody in Fig. 4, but in fact labels total NFH.

Response: "phospho-NFH" now corrected to "NFH" in figure 4, its associated figure legend and the main body of text.

-Fig 5 The complicated nomenclature of the different constructs makes Fig 5 a bit confusing.

Response: We have simplified the names of the constructs in Figure 5B. The labels of the protein blots in the figure are now consistent with the nomenclature used in 5B.

-Table 2: Effect size should be given for all parameters.

Response: We have included a row titled "NT vs SCA8 BAC Difference" to Table 2 that shows the

effect size for each parameter. We have also included the average and SEM for both NT and SCA8 BAC mice for each parameter.

Referee #2:

The manuscript by Ayhan et al describes polySer and polyGln protein accumulation in SCA8 brains from both mice models and human patients. They show regional specificity and that the RAN products accumulate to a greater degree with disease progression. They also perform cell studies to show a correlation between RAN products and eIF3F levels, although at this point causality can only be inferred.
Overall this is an important body of work and there are several suggestions and concerns.

1. In figure 1 it's not clear what alpha-SerCT stands for. It took some digging to find that it may mean the unique C-terminus downstream of the polySer sequence - a predicted RAN product. Some definition in the legend and text is needed.

Response: We have clarified in the legend of Figure 1 that the polySer aggregates in panel 1D and E were detected by unique antibodies to the polySer protein C-terminus. We have also clarified this in the text.

Page 6, lines 2-5: "We generated two rabbit polyclonal antibodies, α-SerCT and α-SerCT2, directed at different non-overlapping peptide sequences within the unique C-terminal region downstream of the predicted SCA8 polySer protein (Figure EV1A)."

2. The drop in gait and open field can also reflect muscle pathology - are there RAN products in the muscle or any muscle pathology?

Response: We thank the reviewer for raising this issue and we have now addressed this in the text. We have previously reported that the SCA8 mouse model does not express ATXN8 or ATXN8OS genes in muscle tissue. Because of the lack of expression and the need to optimize staining in this distinct tissue type we did not pursue this analysis as it seems unlikely that RAN proteins could be expressed at detectable levels (Moseley et al., 2006).

Page 10, lines 3-6: "Although motor abnormalities can be associated with muscle pathology, previous reports did not detect expression of ATXN8 or ATXN8OS transcripts in muscle of SCA8 BAC mice (Moseley et al., 2006), suggesting that these phenotypes are due to CNS features of the disease."

3. The intriguing finding is the cell specificity, which the authors then address by looking for RAN-TAFs, but then they pick a very common target in eIF3F. This is the part of the study that is most problematic - with several concerns.

a. How come the same repeat sequence with an ATG start cannot form a hairpin? It would seem that everything downstream of the methionine would have the same propensity to fold, particularly with that long of a repeat.

Response: We were a little confused about the reviewers question so address this in both the context of the tissue specific expression and also we clarify any possible issues with the ATG-initiated alternative codon constructs.

First, we agree with the reviewer that the same repeat sequence with an ATG start codon can form a hairpin. The cell specificity seen with polyGln accumulation in Purkinje cells and polySer aggregates largely in white matter regions may involve both hairpin formation and also reading frame specific sequence differences that favor RAN translation under specific conditions. In cell culture, we show eIF3F knockdown decreases steady state levels of polySer and polyAla proteins but not ATG-initiated polyAla or ATG-initiated polySer. These data and the cell specificity seen with polyGln in Purkinje cells and polySer in white matter suggests that there may be different requirements for initiation of ATG-initiation and RAN translation which may also be affected by

sequence differences between reading frames, including eIF3F. This is now addressed in the discussion.

Page 18, lines 19-23 and page 19, lines 1-4: "Here we demonstrate that reducing the levels of the mammalian eukaryotic translation initiation factor eIF3F, decrease steady state levels of several RAN but not AUG-initiated expansion proteins in cell culture.
Taken together, these data support a model in which protein translation from reading frames containing upstream AUG start codons initiate using met-tRNAimet and the canonical protein translational machinery. In contrast, RAN translation which does not require an AUG start codon is sensitive to eIF3F levels. The AUG codon present in the polyGln reading frame could therefore explain the cell-type distribution difference between AUG-initiated polyGln and RAN proteins in both SCA8 and HD."

Additionally, we would like to clarify that the ATG-initiated non-hairpin forming constructs used in Figure 4E for the toxicity studies were engineered to encode the same peptide sequences but with different non-hairpin forming nucleotides. For example, "ATG Alt Gln" is ATG initiated and the polyGln protein is encoded by "GAA" repeats which encode glutamine but do not form a hairpin. These constructs were generated specifically to test the toxic effect of the polySer or polyGln proteins independent of RAN translation or possible CAG RNA gain of function effects. This has been clarified in the text.

Page 12, lines 2-9: "We generated alternative codon constructs by selecting different nucleotide sequences that encode the same polyGln or polySer expansion proteins but are unable to form the RAN-permissive hairpin structure (Zu et al, 2011) and instead, only undergo ATG-initiation translation of polyGln or polySer expansion proteins. These alternative codon, non-hairpin forming transcripts allow us to study the toxic effects of individual homopolymeric proteins independent of CAG RNA gain-of-function effects and effects of other RAN proteins (Banez-Coronel 2015)."

b. What cells were used in the KD experiments? This was not explained in the text or legend. How relevant is the cell type? I don't understand how this gets at the specificity point the authors raise - did the authors see in tissues that RAN products and eIF3F expression levels correlate? Does eIF3F interact with the RAN transcripts in RIP studies?

Part 1. Response: HEK293T cells were selected as an experimental model for these knockdown experiments because they undergo high levels of RAN translation (Zu et al, 2011) enabling clear detection of changes in RAN protein levels. This is now more clearly explained in the figure legend and methods.

Page 22, lines 7-8: "HEK293T cells, which express high levels of RAN proteins (Zu et al., 2011), were cultured in DMEM supplemented with 10% fetal bovine serum (FBS)…"

Part 2. We have now included data to show that white matter regions with robust polySer aggregates have increased levels of Eif3f transcripts compared to grey matter regions with fewer polySer aggregates (Figure 5A). These new data are consistent with the hypothesis that elevated eIF3F levels in white matter increase RAN translation.

Page 13, lines 21-23 and page 14, lines 1-5: "To investigate if Eif3f expression is higher in white matter than in grey matter, we performed qRT-PCR comparing tease-separated cerebellar white and grey matter from SCA8 BAC mice. To confirm an enrichment of oligodendrocytes in white matter samples, myelin basic protein (MBP) was measured by qRT-PCR. After confirmation of MBP white matter enrichment (Figure EV5A) we performed qRT-PCR and show Eif3f transcripts have a two-fold increase in SCA8 BAC cerebellar white matter compared to grey matter (Figure 5A)."

Part 3. We agree with the reviewer that it will be important to understand whether eIF3F affects RAN translation via direct or indirect interactions. However, the elucidation of the mechanism of action is beyond the scope of this initial characterization of RAN polySer and white matter abnormalities in SCA8. We have now address this future direction in the discussion.

Page 20, lines 2-11: "In our cell culture experiments, eIF3F knockdown was sufficient to decrease steady state levels of RAN polySer and RAN polyAla, but did not reduce levels of ATG-initiated polySer or ATG-initiated polyAla proteins suggesting eIF3F may preferentially affect RAN proteins and could provide an opportunity to target RAN translation without interfering with canonical translation. Future studies in animal models and the generation and characterization of eIf3f knockout mice will be important to determine if eIF3F/eIf3f knockdown can selectively reduce RAN protein accumulation without deleterious effects on protein homeostasis. Additional studies are also needed to understand how eIF3F affects RAN translation and whether these effects are direct or indirect."

c. The data show a correlation between eIF3F and RAN translation products, but not causality. More studies are needed to conclude that one directly impacts the other. The authors state in the abstract, the results and the discussion that eIF3F may be a therapeutic target but these conclusions should be softened in the absence of additional studies. One could also argue that the data do not show that eIF3F selectively regulates RAN translation - only that reducing it impacts RAN translation products, but there could be many things happening in a pathway linking those two processes as well as other effects of the KD beyond what was examined here.

Response: We agree with the reviewer that the therapeutic potential of eIF3F knockdown needs to be investigated further. We have softened this language in the text. We also agree that we do not show that eIF3F directly regulates RAN translation and that further studies will be required to understand the mechanistic and preclinical relevance of eIF3F. This will be critical to determine the therapeutic potential of eIF3F knockdown. We have edited the abstract to soften the language regarding this point and have expanded the discussion accordingly.

Abstract – page 2, lines 12-13: "Taken together, these data show polySer and WM abnormalities contribute to SCA8 and identify eIF3F as a novel therapeutic target to reduce RAN protein accumulation novel modulator of RAN protein accumulation."

Discussion – page 19, lines 7-11: "While our results suggest that eIF3F plays a role in RAN translation and may be a therapeutic target for multiple microsatellite expansion disorders, it will be important to determine the potential risks and benefits of reduced eIF3F activity in brain. Additionally, more studies are required to understand the mechanism by which eiF3F regulates RAN translation."

Discussion – page 20, lines 2-11: "In our cell culture experiments, eIF3F knockdown was sufficient to decrease steady state levels of RAN polySer and RAN polyAla, but did not reduce levels of ATG-initiated polySer or ATG-initiated polyAla proteins suggesting eIF3F may preferentially affect RAN proteins and could provide an opportunity to target RAN translation without interfering with canonical translation. Future studies in animal models and the generation and characterization of eIf3f knockout mice will be important to determine if eIF3F/eIf3f knockdown can selectively reduce RAN protein accumulation without deleterious effects on protein homeostasis. Additional studies are also needed to understand how eIF3F affects RAN translation and whether these effects are direct or indirect."

Methods -
Describe what cells were used and how many times studies were repeated (biological as well as technical replicates).

Response: HEK293T were used for siRNA-mediated eIF3F knockdown experiments and glial T98 cells were used for RAN protein toxicity experiments. All experiments were performed with a minimum of 3 biological replicates and 5 technical replicates. This has been added to the methods and clarified in the text and figure legends.

Page 22, lines 7-8: "HEK293T cells, which express high levels of RAN proteins (Zu et al., 2011), were cultured in DMEM supplemented with 10% fetal bovine serum (FBS)…"

Page 22, lines 17-23: "Cell toxicity was measured in T98 cells and assessed by measuring lactate dehydrogenase (LDH) released into the medium by dying cells (CytoTox 96 Nonradioactive Cytotoxicity Assay, Promega, Madison, WI) following the manufacturer's protocol. Absorbance was

recorded at 490 nm and total LDH release was measured by lysing the cells with 1% Triton X-100. LDH determinants were measured in three independent experiments and each experiment performed in quintuplicates for each condition (Korzeniewski and Callewaert, 1983).”

Was the open field all dark? There was not a light and dark chamber? The assay setup seems unusual. Also for the gait analysis, what parameters were measured? This should be stated as there are numerous types of output for this equipment.

Response: We have clarified our open field protocol in the methods section and included a reference for this methodology.

Page 24, lines 22-23 and page 25, lines 1-4: “Open field analysis was performed by testing mouse behavior during a 30 min session in a completely dark 17”x17” open chamber (Med Associates, Inc., Fairfax, VT). Mice were allowed to acclimate to the testing room for approximately two hours before the start of analysis. Mice were then placed in the center of the darkened activity-monitoring chamber. Data was analyzed with Activity Monitor software (MED associates, Inc., Fairfax, VT) (Bolivar et al., 2000).”

The 45 parameters measured by DigiGait are standard outputs of Mouse Specifics software and are now listed in Table 2 with the effect size, mean for each genotype, and significance. This has also been clarified in the Methods.

Page 24, lines 18-20: “Data from each paw was analyzed with DigiGait automated gait analysis software and the 45 parameters measured by the software are outlined in Table 2.”

Describe the statistical analysis for the RNA-Seq expression comparison studies.

Response: We have now included this in the methods section.

Page 30, lines 22-24 and page 31, lines 1-3: “Fragment per kilobase million (FPKM) of ATXN8 and ATXN8OS was calculated by normalizing the read count mapped to gene-specific regions of ATXN8 and ATXN8OS to the length. The difference was calculated by dividing both FPKM values by ATXN8OS FPKM value. This analysis was done on three biological replicates and statistical significance was calculated using an unpaired t-test.”

Referee #3:

Spinocerebellar ataxia type 8 (SCA8) is a dominantly inherited neurodegenerative disease caused by a CTG•CAG microsatellite expansion mutation. The repeats can lead to accumulation of multiple potentially toxic products: RNA foci, polyGln (by ATG translation), polyAla and polySer (both by repeat associated non-ATG (RAN) translation). This manuscript by Ayhan et al. showed the SCA8 RAN polySer protein preferentially accumulates in white matter regions and increases with age and disease severity, accompanied with demyelination and axonal degeneration. The production of RAN proteins polySer and polyAla can be reduced by eIF3F knockdown, a potential therapeutic target. This study is important for our understanding of the pathogenic contribution of RAN translation products to neurodegeneration diseases. However, several evidence need to be further investigated to provide more convincing support.

Major concerns:
1. Only one of the RAN translation product, polySer, is examined in the whole work, except eIF3F is shown to regulate both in the last figure. PolyAla probably has similar expression patterns in the different brain regions and during aging as polySer, and also contribute to the toxicity. Since the lab has the antibody for polyAla (shown in previous publication Zu et al., 2011), the polyAla expression should also be systematically analyzed and compared with polySer.

Response: Zu et al 2011 showed that polyAla is found in Purkinje cells, however, due to background issues in other brain regions, the polyAla antibody was not useful in this study for determining the overall brain distribution of polyAla. A paragraph discussing this has been added to the discussion. Therefore, while we agree in principle with the reviewer, we suggest these additional detailed studies would be appropriate for a follow-up study in a more specialized journal. Despite this, we

have been able to examine polyAla in HEK293 cells and show that eIF3F knockdown reduces levels of RAN polyAla.

Page 16, lines 10-20: "In addition to our current report of polySer, we previously reported the accumulation of RNA foci, polyAla and polyGln proteins in SCA8 BAC mice and patient autopsy tissue. Initially, it was demonstrated that polyGln accumulates in Purkinje and pontine neuronal nuclei (Moseley et al., 2006). The present work extends the molecular characterization of SCA8 by showing polyGln aggregates are also found in the frontal cortex, pons and hippocampus (Figure 2B), and are consistently located neuronal nuclei (Figure 2C). RAN PolyAla staining was previously reported in Purkinje cell soma and dendritic processes (Zu et al., 2011), and CUG RNA foci were detected in Purkinje cells, molecular layer interneurons, and Bergmann glia (Daughters et al., 2009). Additional studies are needed to investigate the brain distribution of polyAla and possible antisense RAN proteins."

2. Fig.2 showed different expression/distribution of polySer and polyGln proteins in different brain regions in both SCA8 BAC mouse and human autopsy samples using immunohistochemistry. This method only allows the detection of aggregated forms of the polypeptide. Elisa or dot blot (such as in Fig.5C) of total lysates should also be examined to exclude the possibility that different polypeptides have different solubility or aggregation properties in different regions or cell types.

Response: While we agree in principle, and have now performed this experiment we were unable to obtain a clean signal in protein blots or ELISAs using mouse brain lysates and the polySerCT polyclonal antibodies. The dot blots referred to in Figure 5C were performed on cultured cell lines using monoclonal antibodies against the C-terminal epitope tags on these overexpressed constructs, and so this experiment showed less background reactivity.

3. What's the expression level of ATXN8 in different brain regions? Does it correlate with the polyGln aggregates? Why does polyGln have lower expression in white matter? The RNA and protein levels of ATXN8 should be measured at these different brain regions.

Response: We have now performed quantitative RT-PCR and show ATXN8 RNA expression in white matter and grey matter from cortex and cerebellum (Figure EV5B). These data show no significant difference in expression between these regions. ATXN8 transcripts are is expressed at very low levels throughout the CNS, making it difficult to perform in situ hybridization and to correlate RNA transcript levels with the polyGln and polySer proteins. We would like to clarify that the ATG-initiated ATXN8 protein is encoded by a small open reading frame containing an AUG initiation codon followed by a CAG expansion and stop codons. In other words, this polyGln protein has no other flanking sequence and the ATXN8 protein is a methionine followed by a polyGln expansion so there is no larger ATXN8 *protein to measure. Finally, we believe that the ATG-initiated polyGln is expressed at lower levels in the white matter due to increased levels of eIF3F in white matter (now shown in Figure 5A), a condition which favors RAN translation over ATG-initiated translation.*

Page 13, lines 21-23 and page 14, lines 1-11: "To investigate if Eif3f expression is higher in white matter than in grey matter, we performed qRT-PCR comparing tease-separated cerebellar white and grey matter from SCA8 BAC mice. To confirm an enrichment of oligodendrocytes in white matter samples, myelin basic protein (MBP) was measured by qRT-PCR. After confirmation of MBP white matter enrichment (Figure EV5A) we performed qRT-PCR and show Eif3f transcripts have a two-fold increase in SCA8 BAC cerebellar white matter compared to grey matter (Figure 5A). We also evaluated levels of ATXN8 RNA in cerebellar grey and white matter and found no significant difference in the expression of ATXN8 transcripts (Figure EV5B), indicating that changes in ATXN8 RNA expression are unlikely to contribute to the increased polySer aggregate accumulation in white matter regions. The elevated Eif3f levels observed in cerebellar white matter, which shows robust polySer aggregation, are consistent with the possibility that eIF3F increases RAN translation leading to the accumulation of RAN polySer protein in white matter."

4. Fig.2: it is described that polyGln staining is primarily nuclear, but polySer shows perinuclear localization or localized in neuropil (shown in 2C). However, in Fig.2C, most of the polySer is not co-stained with NeuN and looks more like it is expressed in other cell types or extracellularly rather

than in neuropil. Co-staining with glia cell markers (astrocyte, oligodendrocyte, microglia) should be included.

Response: We have now performed CNPase IF staining and show that polySer appears as perinuclear aggregates or as aggregates located close to or within oligodendrocytes/myelinated axons (Figure EV2). This distribution is distinct from the polyGln protein which appears as nuclear aggregates (Figure 2).

Page 8, lines 1-4: "To further clarify the cellular localization of RAN polySer, we performed double-labeling IF of polySer and CNPase, a marker of oligodendrocytes, and show that polySer aggregates located in the neuropil are close to or within oligodendrocytes/myelinated axons (Figure EV2)."

5. There is inconsistency of polySer staining between human patient autopsy tissue and SCA8 mouse tissues. It is shown in mouse samples that polySer has perinuclear localization. But in Fig.2D and 4B, it seems the SerCT actually shows staining mostly in the nucleus in human patient autopsy tissue. Does the mouse model really reproduce the human phenotypes?

Response: We agree that the staining between human and mouse look somewhat different. These differences may be due to the end stage of disease in the human cases and to differences in tissue handling and processing including postmortem delay for the human autopsy samples. Additionally, due to animal care services and IACUC regulations we are not allowed to age our animals to true disease end point as occurs in humans. We have now addressed these points in the text.

Page 8, lines 7-11: "Despite this, there are some differences in the appearance of polySer in human postmortem tissue and SCA8 BAC mouse brain. These differences may be due to the end stage of disease in the human cases and to differences in tissue handling and processing including postmortem delay and the severity of disease at endpoint in human cases."

6. It is nicely shown that the polySer accumulation increases with age and severity of disease in Fig.3. But the brain region is not consistent with ones examined in Fig.2. For example, cerebellar white matter and molecular layer were not examined during aging. This should be included. And quantifications are needed for 3A. In addition, how polyGln accumulation changes during aging should also be examined.

Response: For Figure 2, we focused on brain regions that include both polyGln and polySer staining simultaneously so that we could compare their regional distribution at animal care services specified end-stage for the mice. The polySer aggregates shown in the brain regions in Figure 2 are small and appear at later stages of disease. Therefore, it was not possible to quantify polySer aggregates in these brain regions throughout disease progression. Because the brainstem and motor cortex are some of the first regions to show polySer aggregates in SCA8 BAC mice, while cerebellum does not show aggregates until later in disease, we chose brainstem and motor cortex for the time course experiment. We have now included quantifications of polySer aggregates in brainstem and motor cortex at the three time points included in Figure 3A. The choice of selected brain regions and references to the quantification are now included in the text. Because the polyGln aggregates are not present until later stages of disease, we could not include them in this time course study.

Page 9, lines 3-10: "We performed IHC at different ages to address how polySer RAN protein aggregates change over time and disease progression (Figure 3A). To do this, we identified regions that showed polySer aggregates at 2 months of age and quantified the number of polySer aggregates in these regions at 2, 6 and 10 months of age. At 2 months, when SCA8 animals show no overt cage behavior phenotypes, IHC showed very small, pin-like polySer aggregates that were found infrequently in brainstem regions including the vestibular nuclei and cuneate nuclei but not detectable in the cerebellum or hippocampus."

Page 9, lines 15-17: "Quantification of polySer aggregates in the motor cortex and brainstem (vestibular nucleus and cuneate nucleus) showed an increase in aggregates from 2-10 and 6-10 months of age (Figure 3A)."

7. In Fig.4E, the RAN translation construct should also be used to compare the toxicity with ATG alt Gln. The relative polySer expression level from the repeat construct (RAN translation) vs ATG Alt Ser construct should be compared and included. Although the current data showed polySer is slightly more toxic than polyGln, but the expression from RAN repeats is probably much lower than AUG translation, therefore leads to lower toxicity in the real context. Furthermore, there is no data on polyAla (neither RAN construct nor ATG version). But in the last sentence of the second paragraph on page 11, it said "polySer expansion proteins are toxic to glial cells independent of the CAG expansion transcripts or polyGln or polyAla proteins".

Response: We appreciate the reviewer's interest in RAN protein toxicity and how this relates to ATG vs RAN levels. We have previously performed similar studies (Banez-Coronel et al., 2015). For the experiment in figure 4E however, our focus is to address the question of whether polySer and polyGln proteins are toxic when overexpressed independent of RNA gain-of-function effects and other proteins in SCA8. It is difficult to compare toxicity from ATG-initiated Alt-codon constructs and individual RAN proteins from our RAN constructs, as the latter express CAG expansion transcripts and polyGln, polySer and polyAla RAN proteins, all of which could contribute to toxicity. Additionally, toxicity in patient or SCA8 BAC mouse brain tissues is further complicated by differences in translational efficiency, protein expression and protein clearance pathways between various cell types. For these reasons, we prefer to keep our message simple and not to distract the readers with these additional experiments.
We thank the reviewer for catching our mistake of including polyAla in our description of toxicity. The reviewer is correct in that there is no data on polyAla as non-hairpin forming codons are not available for polyAla. We have modified the sentence accordingly.

Page 13, line 5-6: "Furthermore, in vitro studies show polySer expansion proteins are toxic to glial cells independent of the CAG expansion transcripts."

8. Fig.5A: In order to show the correlation of eif3f levels with polySer, the Eif3f transcript levels in different brain regions should be examined.

Response: We have now included quantitative RT-PCRs comparing eIF3F expression in tease separated, oligodendrocyte rich white matter versus grey matter of SCA8 BAC mouse neural tissue. These experiments show that eIF3F transcripts are higher in white matter compared to grey matter (Figure 5A).

Page 13, lines 21-23 and page 14, lines 1-11: "To investigate if Eif3f expression is higher in white matter than in grey matter, we performed qRT-PCR comparing tease-separated cerebellar white and grey matter from SCA8 BAC mice. To confirm an enrichment of oligodendrocytes in white matter samples, myelin basic protein (MBP) was measured by qRT-PCR. After confirmation of MBP white matter enrichment (Figure EV5A) we performed qRT-PCR and show Eif3f transcripts have a two-fold increase in SCA8 BAC cerebellar white matter compared to grey matter (Figure 5A). We also evaluated levels of ATXN8 RNA in cerebellar grey and white matter and found no significant difference in the expression of ATXN8 transcripts (Figure EV5B), indicating that changes in ATXN8 RNA expression are unlikely to contribute to the increased polySer aggregate accumulation in white matter regions. The elevated Eif3f levels observed in cerebellar white matter, which shows robust polySer aggregation, are consistent with the possibility that eIF3F increases RAN translation leading to the accumulation of RAN polySer protein in white matter."

9. What's the timing of white matter abnormalities, compared to neuronal defects? This should be evaluated to understand how much the white matter abnormalities could contribute to the disease progression.

Response: While the neuronal defects in the SCA8 BAC mice have not yet been fully characterized, a loss of GABAergic inhibition has previously been reported in the cerebellar molecular layer (Moseley et al., 2006). While we agree that it will be interesting to study how neuronal dysfunction

correlates with white matter abnormalities, these functional studies are beyond the scope of this molecularly focused paper.

Minor points:
1. There are two antibodies targeting SerCT produced, shown in Fig.EV1A,B. But in all the rest of figures in the manuscript, it is not mentioned which antibody was used and whether both antibodies showed similar results.

Response: We have now clarified throughout the manuscript the specific polySer antibody used. We selected the antibody with the least background reactivity for the figures in the manuscript: α-polySerCT was used for mouse staining and anti-polySerCT2 was used for human staining. Examples of mouse and human staining with both antibodies are now included in Figure EV1 D and E.

Page 6, lines 2-5: "We generated two rabbit polyclonal antibodies, α-SerCT and α-SerCT2, directed at different non-overlapping peptide sequences within the unique C-terminal region downstream of the predicted SCA8 polySer protein (Figure EV1A)."

Page 6, lines 10-12: "Although both α-SerCT antibodies showed similar punctate straining, α-SerCT was used for IHC analyses of SCA8 BAC mouse tissue as it showed less background reactivity."

Page 6, lines 20-22: "Both antibodies were also able to detect polySer aggregates in patient autopsy tissue. However, because α-SerCT2 showed less non-specific reactivity in human tissue it was used for subsequent IHC on human tissue (Figure EV1E)."

2. Fig.EV2A: Protein levels from the different constructs should also be compared besides RNA levels.

Response: In Banez-Coronel et al 2015, we showed that similar alternative codons used to encode polyGln and polySer produce predominantly polyGln and polySer respectively, we cite this reference.

Page 12, lines 13-16: "It has previously been shown that alternative codon constructs used to encode polyGln and polySer produce polyGln and polySer proteins respectively (Banez-Coronel et al., 2015)."

3. In the abstract and introduction (last sentence of the first paragraph), the authors only mentioned RAN polyAla from the SCA8 CAG expansion as a known fact. But polySer has been shown in previous report (Zu, Gibbens et al., 2011) as well, using in vitro assays. The text should be modified to describe previously known findings more precisely.

Response: While Zu et al 2011 showed that CAG repeats can make polySer in vitro, this is the first in vivo evidence of polySer in mouse and human brain. The first sentence of the abstract and the last sentence of the first paragraph of the introduction have now been edited accordingly.

Page 2, lines 2-5: "Spinocerebellar ataxia type 8 (SCA8) is caused by a bidirectionally transcribed CAG•CTG expansion that results in the in vivo accumulation of CUG RNA foci, an ATG-initiated polyGln and a polyAla protein expressed by repeat associated non-ATG (RAN) translation."

Page 3, lines 12-16: "In addition, SCA8 CAG expansion transcripts have been shown to result in the expression and accumulation of an ATG-initiated polyglutamine expansion protein and a repeat associated non-ATG (RAN) polyalanine expansion protein in brains of SCA8 BAC mice and SCA8 autopsy tissue (Moseley et al., 2006; Zu et al., 2011)."

4. On page 15, last sentence of the second paragraph, "will be" is duplicated.

Response: the duplicated "will be" has been removed.

2nd Editorial Decision 18th Jul 18

Thank you for submitting your revised manuscript to The EMBO Journal. Your study has now been re-reviewed by the three original referees. As you can see from the comments below, the referees appreciate the introduced changes and support publication here.

Referee #3 has a few remaining comments. Do you have any data on hand to address the major points? The ELISA experiment with polyGln antibody seems like a good suggestion, but lets discuss this further.

When you resubmit the revised version please also take care of the following points:

- Please upload individual figure files

- There is a callout to Figure S3B on p.15 but no such figure.

- Also, we encourage the publication of source data, particularly for electrophoretic gels and blots. It would be great if you could provide me with a PDF file per figure that contains the original, uncropped and unprocessed scans of all or key gels used in the figure? The PDF files should be labeled with the appropriate figure/panel number, and should have molecular weight markers; further annotation could be useful but is not essential. The PDF files will be published online with the article as supplementary "Source Data" files.

- We include a synopsis of the paper (see http://emboj.embopress.org/). Please provide me with a general summary statement and 3-5 bullet points that capture the key findings of the paper.

Happy to discuss further the remaining few points raised by referee #3

REFEREE COMMENTS

Referee #1:

The manuscript greatly improved, although the authors did not do provide all data requested by the reviewers, e.g. the effect of eIF3F overexpression on RAN translation.

Referee #2:

I am satisfied with the revised manuscript and have no further concerns

Referee #3:

This is an improved manuscript that reports on the RAN translation product, SCA8 polySer, that accumulates in white matter and contribute to disease. This study is important for our understanding of the pathogenic contribution of RAN translation products to neurodegeneration diseases. Some of the evidence are still soft for the main conclusion and need further improvement.

Major points:
1. The localization of polySer in white matter is still not clear. The evidence is not sufficient to support "polySer aggregates located in the neuropil" which are "close to or within oligodendrocytes/myelinated axons". Co-staining with axon and dendrite markers might be helpful.

2. Although it is difficult to do ELISA with the SerCT antibody, the authors should still perform

ELISA with polyGln antibody to confirm the total polyGln levels are indeed lower in the white matter. Otherwise the toxicity in white matter could still be contributed from polyGln. And in the response to point 3, the authors said "the ATG-initiated polyGln is expressed at lower levels in the white matter due to increased levels of eIF3F in white matter, a condition which favors RAN translation over ATG-initiated translation". Where is the evidence for that? It should be easily tested in cell lines. Will 2-fold difference of eif3f levels in white matter versus grey matter make such a huge difference?

Minor point:
1. Page 6, line 22: "it was used for subsequent IHC..." should be changed to "α-SerCT was used for subsequent IHC...".

2nd Revision - authors' response 31st Jul 18

Thank you very much for the thoughtful reviews and your careful consideration of our manuscript. A detailed summary of our responses to the additional comments from reviewer 3 is below. We hope these additional changes now make our manuscript suitable for publication in *EMBO*.

Referee #3:

1. The localization of polySer in white matter is still not clear. The evidence is not sufficient to support "polySer aggregates located in the neuropil" which are "close to or within oligodendrocytes/myelinated axons". Co-staining with axon and dendrite markers might be helpful.

Response: We thank the reviewer for raising this issue and now more clearly understand that we have not localized the aggregates specifically to the synaptically dense regions of the neuropil. We now more clearly describe the localization of polySer within the grey and white matter regions of the brain.

2a. Although it is difficult to do ELISA with the SerCT antibody, the authors should still perform ELISA with polyGln antibody to confirm the total polyGln levels are indeed lower in the white matter. Otherwise the toxicity in white matter could still be contributed from polyGln.

Response: The SCA8 polyGln protein is expressed at relatively low levels throughout the brain making western blots and ELISAs difficult assays for SCA8 polyGln detection and quantification. Further complicating this type of experiment is the fact that the 1C2 antibody also recognizes the soluble TATA binding protein which has ~40 glutamines. Therefore, while we can reliably detect aggregated SCA8 polyGln by IHC, accurate measurement of total SCA8 polyGln levels is not currently feasible by ELISA or dot blot (which we tried in the previous round of review).

Although we don't detect polyGln aggregates in the white matter regions, we agree with the reviewer that it is possible that soluble SCA8 polyGln may also contribute to the white matter abnormalities and now mention this possibility in the Discussion.

Page 18, lines 4-7: "The high toxicity of these homopolymeric polySer proteins suggests that they contribute to the white matter abnormalities found in the polySer positive regions in both SCA8 and HD, although additional RAN or polyGln proteins may also contribute.*"*

2b. And in the response to point 3, the authors said "the ATG-initiated polyGln is expressed at lower levels in the white matter due to increased levels of eIF3F in white matter, a condition which favors RAN translation over ATG-initiated translation". Where is the evidence for that? It should be easily tested in cell lines. Will 2-fold difference of eif3f levels in white matter versus grey matter make such a huge difference?

Response: We agree that this question will be interesting as a follow-up experiment, but the most suitable context for this experiment would be to test this question in vivo using cell type specific overexpression. These experiments are beyond the scope of this current manuscript.

Minor points
1.Page 6, line 22: "it was used for subsequent IHC..." should be changed to "α-SerCT was used for subsequent IHC...".

Response: Thank you for raising this, it has now been corrected to:

> *"…α-SerCT2 was used for subsequent IHC…"*

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Laura P. W. Ranum
Journal Submitted to: The EMBO Journal
Manuscript Number: EMBOJ-2018-99023

Reporting Checklist For Life Sciences Articles (Rev. June 2017)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- ➔ the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- ➔ figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- ➔ graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- ➔ if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- ➔ Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- ➔ a specification of the experimental system investigated (eg cell line, species name).
- ➔ the assay(s) and method(s) used to carry out the reported observations and measurements
- ➔ an explicit mention of the biological and chemical entity(ies) that are being measured.
- ➔ an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- ➔ the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- ➔ a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- ➔ a statement of how many times the experiment shown was independently replicated in the laboratory.
- ➔ definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com	Antibodypedia
http://1degreebio.org	1DegreeBio
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repo	ARRIVE Guidelines
http://grants.nih.gov/grants/olaw/olaw.htm	NIH Guidelines in animal use
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm	MRC Guidelines on animal use
http://ClinicalTrials.gov	Clinical Trial registration
http://www.consort-statement.org	CONSORT Flow Diagram
http://www.consort-statement.org/checklists/view/32-consort/66-title	CONSORT Check List
http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tun	REMARK Reporting Guidelines (marker prognostic studies)
http://datadryad.org	Dryad
http://figshare.com	Figshare
http://www.ncbi.nlm.nih.gov/gap	dbGAP
http://www.ebi.ac.uk/ega	EGA
http://biomodels.net/	Biomodels Database
http://biomodels.net/miriam/	MIRIAM Guidelines
http://jjj.biochem.sun.ac.za	JWS Online
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html	Biosecurity Documents from NIH
http://www.selectagents.gov/	List of Select Agents

In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	Because polySer is a novel finding, we have examined many SCA8 animals with littermates controls from independent cohorts (n>15) ussing brain tissue prepared with different method (i.e. fresh frozen, formalin fixed). Once we established the reproducibility and consistency of polySer detection with both IHC and IF; we decided to include n=3 for the following studies detailed in this paper.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	A standard sample size of n=3 was used for quantification of mouse pathology and eIF3F transcript expression, which is sufficient for the statistical analyses performed.
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	No samples and animals were excluded.
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	These experiments were performed on multiple animal cohorts including both sexes.
For animal studies, include a statement about randomization even if no randomization was used.	Affected animals used in the study were chosen at random from a larger pool of available affected animals.
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	Histology and histopathology experiments including quantification were performed blinded. The reproducibility of these results were confirmed by a second investigator independently.
4.b. For animal studies, include a statement about blinding even if no blinding was done	All samples were de-identified before quantification of polySer aggregate number.
5. For every figure, are statistical tests justified as appropriate?	Standard statistical analyses were used in the manuscript.
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Yes the data meets the assumptions of the statistical tests used for analysis.
Is there an estimate of variation within each group of data?	SCA8 polySer staining is consistently negative in gene-negative animals and human controls. White matter degeneration phenotypes are consistent in late stage animals.
Is the variance similar between the groups that are being statistically compared?	Yes.

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	New England Peptide custom made rabbit polyclonal antibodies. Rabbit antisera was raised against synthetic peptides Ac-CRVNLSVEAGSQKRQSE-amide and Ac-CSSSKARFSNMKD-amide and for α-polySer (NEP, Project #1306, Rabbit #F3672) and α-polySer2 (NEP, Project #2953, Rabbit #I2108).

7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	Previously established and characterized cell lines (HEK293T and T98) are used in this study, these were originally obtained from the American Type Culture Collection.

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	Previously described SCA8 BAC transgenic lines on the FVB background (BAC EXP2, 2878) were used (Moseley et al., 2006). FVB mice used for breeding purposes were obtained from The Jackson Laboratory. Hemizygous mice with the SCA8 BAC transgene were by genotyped by PCR as previously described (Moseley et al., 2006). Due to severe motor dysfunction that SCA8 BAC expansion mice exhibit after 5 months of age, additional food (GelDiet, Clear H2O, Westbrook, MA) was provided in the bottom of the cage for animals >5 months. Both genders of animal were used throughout the study with ages ranging from 2 months to endstage (10 months) which are reported in the figure legends.
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	Mice used in this study were housed and treated in according to the NIH Guide for the Care and Use of Laboratory Animals. All animal studies were approved by Animal Care and Use Committee at the University of Florida.
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	We confirm compliance.

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	The Institutional Review Board approved the study protocols under which human tissue was collected, processed and analysed.
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	Control and SCA8 autopsy tissue was collected at the University of Minnesota and the University of Florida with informed consent of patients or their relatives and approval of respective institutional review boards.
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	N/A
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	Human autopsy samples are availble on request providing the requesting party meets the guidelines required by the IRB protocol. The autopsy samples used in this study are rare and so a limited amount of tissue will be available on request.
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	N/A
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	N/A
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	N/A

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	N/A
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	N/A
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	There is no human clinical or genomic dataset associated with this study.
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	N/A

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	N/A