THE
EMBO
JOURNAL

C9ORF72 haploinsufficiency synergizes the toxicity of DPR proteins, a double hit mechanism that can be prevented by drugs activating autophagy

Manon Boivin, Véronique Pfister, Angeline Gaucherot, Frank Ruffenach, Luc Negroni, Chantal Sellier, Nicolas Charlet-Berguerand

Review timeline:

Submission date:	27th Aug 2018
Editorial Decision:	29th Oct 2018
Revision received:	3rd Sep 2019
Editorial Decision:	10th Oct 2019
Revision received:	28th Nov 2019
Accepted:	3rd Dec 2019

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 27th Aug 2018

Thank you for submitting your manuscript to The EMBO Journal. I am sorry for the delay in getting back to you, but I have now received the two referee reports on the manuscript.

As you can see from the comments below, both referees find the analysis interesting and that it provides important insight. However, they also raise major concerns with the analysis that would have to be resolved in order to consider publication here. Should you be able to address the major concerns raised then I would like to invite you to submit a revised manuscript. I should add that it is EMBO Journal policy to allow only a single major round of revision and that it is therefore important to address the raised concerns at this stage.

REFEREE REPORTS:

Referee #1:

GGGGCC repeat expansion in the first intron of C9ORF72 is the most common genetic cause of both ALS and FTD. Recent breathtaking progress have greatly advanced our understanding of pathogenic mechanisms of C9ORF72-ALS/FTD. However, many important questions remain to be addressed. In this short and relatively descriptive study, the authors report several findings that may be interesting for the C9ORF72 field, some of them contribute to current debates on a couple of controversial issues. On the other hand, to further improve the manuscript, additional experiments seem to be warranted, and the writing in some areas needs attention.

1. Instead of transfecting cells with individual DPR proteins, the authors expressed expanded G4C2 repeats to investigate the synergy between partial loss of C9ORF72 protein and DPR proteins synthesized through their native translation initiation sites. This strategy allows several interesting

questions to be addressed. The authors first identified CUG as the translation start site for poly(GA), a finding that is consistent with several recent reports (Green et al., Nat. Commun. 2017; Cheng et al., Nat. Commun. 2018; Tabet et al., Nat. Commun. 2018). Although this result is not novel, the authors went on to show that poly(GA) synthesis is INDEPENDENT of G4C2 repeat length and that short poly(GA) with a small but not a big tag is efficiently degraded. These experiments should be presented in a main figure because they make an important point-that expanded repeats do not affect translation initiation per se but instead increase the size of the open reading frame, resulting in a larger, more stable DPR protein. Although not everyone in the field will agree with this point, the evidence presented here seems convincing. Thus, these findings will contribute to the current debate regarding the mechanisms of DPR protein synthesis in C9ORF72-ALS/FTD.

2. The authors did not provide experimental evidence that, in human cells, the spliced intron or unspliced pre-mRNA serves as the template for DPR synthesis. Therefore, they should not predict the length of the C-terminus of poly(GA) and poly(GR) based on sequences of unspliced pre-mRNA, as shown in Figures 1D and 1F, respectively. In fact, Cheng et al. (Nat. Commun. 2018) argued that the intron is likely translated into DPR proteins. In that case, the predicted C-terminus would be much shorter.

3. The identification of the translation initiation site for poly(PG) on the antisense repeat RNA is novel. Since the start codon for poly(PG) synthesis is AUG, the phrase "repeat-associated non-AUG translation" is inappropriate, and should be deleted from the abstract because the study shows that poly(PG) synthesis is not mediated by "RAN translation". The authors conclude that very little poly(GP) is translated from the sense repeat RNA-a finding that contradicts several published studies in which use of ASOs to degrade sense repeat RNA significantly decreased poly(GP) level. Thus, the authors should confirm whether most poly(GP) is translated from the antisense repeat RNA in human C9ORF72 cells.

4. The toxicity of DPR proteins have been well documented in multiple experimental systems. The idea that decreased C9ORF72 expression also plays role is not totally novel. In iPSC-derived human neurons, DPR proteins are produced and the C9ORF72 level is reduced. Shi et al. (Nat. Med. 2018) reported that ectopic expression of C9ORF72 rescues neuronal cell death in iPSC-derived patient motor neurons, demonstrating the importance of C9ORF72 haploinsufficiency. Moreover, lit has been well documented that loss of C9ORF72 compromises the autophagy pathway, which is essential for aggregate clearance. Thus, the finding that decreased C9ORF72 expression increases DPR aggregate formation and toxicity in cells that express expanded repeats is expected. The lack of significant conceptual novelty in this set of experiments (Figure 3) is a major weakness of the paper. In these experiments, the authors concluded that the toxicity was caused by poly(GA) or poly(GP). However, in principle, poly(GR), which is highly toxic even when expressed at a low level, should be also expressed from the GGGGCC repeats. Tabet et al. (Nat. Commun. 2018) reported that mutations in the near-cognate start codon CUG greatly decreased the levels of both poly(GA) and poly(GR). Therefore, the observed toxicity cannot be attributed to poly(GA) alone. The authors should determine whether mutating AGG to decrease poly(GR) expression rescues the cell death phenotype. Similarly, it is surprising that a low level of poly(GP) induced toxicity. The authors should also consider whether a low level of poly(PR), which is also highly toxic, is produced from their antisense repeat RNA.

5. The authors tested different drugs to boost autophagy activity. It is useful to show that promethazine reduces DPR accumulation and toxicity when repeat plasmids are transiently overexpressed, since this is a clinically approved drug. However, it would be more convincing to show that this drug has therapeutic benefits in iPSC-derived patient neurons or in C9ORF72 BAC transgenic mice.

6. Overall, the writing seems too repetitive. For instance, in several places in the Results section, conclusions are restated again: "to summarize this part, ..." In some other parts, almost identical paragraphs are used to describe a different construct.

Referee #2:

Boivin et al manuscript contributes to the understanding of dipeptides-repeat proteins (DRP) production from C9ORF72 hexanucleotide repeat expansion causing ALS/FTD by profiling the relative abundance of the different DRPs and identifying upstream AUG and near-cognate codons driving protein synthesis in both sense (G4C2) and antisense (C4G2) frames. The study is very elegant by showing that levels of DRP produced from translation of the natural human C9ORF72 are subtoxic and only impact on neuronal viability in vitro when combined with C9ORF72 deficiency by knock-down. Decreased expression of C9ORF72 protein had been reported in ALS/FTD, but its relevance was not clear before the work presented here. The study by Boivin et al also links the synergistic effects of DPR and C9ORF72 loss-of-function to autophagy deficiency, being able to rescue DPR toxicity with autophagy inducers. The manuscript is excellent, very well written with a clear message. I recommend publication in EMBO J after the major points have been addressed.

1- The authors must show that C9ORF72 knock-down actually leads to increased levels of DRP. Please provide analysis by Western-blot.
2- The authors have to quantify DRP aggregation upon C9ORF72 knock-down. Ideally, quantification of GFP-positive aggregates is complemented with biochemical methods.
3- What is the role of C9ORF72 in DRP degradation by autophagy. Does it play a specific role in DRP and repeat proteins clearance? Please provide evidence of the actual role of C9ORF72 in the autophagy pathway?

1st Revision - authors' response 3rd Sep 2019

REFEREE #1:

GGGGCC repeat expansion in the first intron of C9ORF72 is the most common genetic cause of both ALS and FTD. Recent breathtaking progress have greatly advanced our understanding of pathogenic mechanisms of C9ORF72-ALS/FTD. However, many important questions remain to be addressed. In this short and relatively descriptive study, the authors report several findings that may be interesting for the C9ORF72 field, some of them contribute to current debates on a couple of controversial issues. On the other hand, to further improve the manuscript, additional experiments seem to be warranted, and the writing in some areas needs attention.

1. Instead of transfecting cells with individual DPR proteins, the authors expressed expanded G4C2 repeats to investigate the synergy between partial loss of C9ORF72 protein and DPR proteins synthesized through their native translation initiation sites. This strategy allows several interesting questions to be addressed. The authors first identified CUG as the translation start site for poly(GA), a finding that is consistent with several recent reports (Green et al., Nat. Commun. 2017; Cheng et al., Nat. Commun. 2018; Tabet et al., Nat. Commun. 2018). Although this result is not novel, the authors went on to show that poly(GA) synthesis is INDEPENDENT of G4C2 repeat length and that short poly(GA) with a small but not a big tag is efficiently degraded. These experiments should be presented in a main figure because they make an important point-that expanded repeats do not affect translation initiation per se but instead increase the size of the open reading frame, resulting in a larger, more stable DPR protein. Although not everyone in the field will agree with this point, the evidence presented here seems convincing. Thus, these findings will contribute to the current debate regarding the mechanisms of DPR protein synthesis in C9ORF72-ALS/FTD.
As suggested by Reviewer #1, we moved the GFP (large) versus HA (small) tag-repeat experiments as main figures 1 and 2, and their relevance is now better discussed in the text. Overall, these data confirm that translation initiation occurs upstream of the repeat expansion in both normal and pathological conditions. However, we show in this revised draft that, in normal conditions, the resulting DPR proteins are small, unstable and undetectable, while in pathogenic conditions, G4C2 repeats expansion leads to longer proteins that are stable and prone to accumulate into toxic aggregates.

2. The authors did not provide experimental evidence that, in human cells, the spliced intron or unspliced pre-mRNA serves as the template for DPR synthesis. Therefore, they

should not predict the length of the C-terminus of poly(GA) and poly(GR) based on sequences of unspliced pre-mRNA, as shown in Figures 1D and 1F, respectively. In fact, Cheng et al. (Nat. Commun. 2018) argued that the intron is likely translated into DPR proteins. In that case, the predicted C-terminus would be much shorter.
Our scheme was indeed incorrect, potential skipping or retention of intron 2 are now presented in a novel figure 1F.

3. The identification of the translation initiation site for poly(PG) on the antisense repeat RNA is novel. Since the start codon for poly(PG) synthesis is AUG, the phrase "repeat-associated non-AUG translation" is inappropriate, and should be deleted from the abstract because the study shows that poly(PG) synthesis is not mediated by "RAN translation". The authors conclude that very little poly(GP) is translated from the sense repeat RNA-a finding that contradicts several published studies in which use of ASOs to degrade sense repeat RNA significantly decreased poly(GP) level. Thus, the authors should confirm whether most poly(GP) is translated from the antisense repeat RNA in human C9ORF72 cells.
We would like to thank the reviewer for this helpful comment. The abstract and the text were indeed inconsistent with our data and this is now corrected.
Concerning expression of the polyPG DPR encoded by the G4C2 sense repeat expansion, a recent study (Tabet et al., 2018) reports a frameshifting from the GA frame (initiated from the upstream CTG near cognate codon) to the GP frame, explaining some expression of polyGP. However, despite many attempts we detect only trace amount of this protein by immunoblot against a tag (HA or GFP) fused in the GP frame, or using commercial antibodies against GP dipeptides. This is most likely due to the lower sensitivity of our assay (immunoblotting) compared to the more sensitive in vitro translation assays used in the Tabet work. In that aspect, note that using a sensitive nanoluciferase assay, the group of Peter Todd was able to detect some polyGP expression, however at 10 to 20 folds lesser levels than the polyGA expressed from the CUG near cognate codon, suggesting indeed a weak frameshifting from the GA to the GP frame, or alternatively some weak RAN translation initiating within the G4C2 repeats (Green et al., 2017). Note that both studies (Green et al., 2017; Tabet et al., 2018) do not investigate the antisense C9ORF72 transcript and thus do not compare polyGP (sense) to polyPG (antisense) levels. In contrast, our work indicates that the polyPG DPR protein expressed from the antisense C4G2 repeats is easily observed, while we detect only trace amount of polyGP DPR translated from the sense repeats. These results as well as the Tabet and Green studies are now cited and discussed in the text.

4. The toxicity of DPR proteins have been well documented in multiple experimental systems. The idea that decreased C9ORF72 expression also plays role is not totally novel. In iPSC-derived human neurons, DPR proteins are produced and the C9ORF72 level is reduced. Shi et al. (Nat. Med. 2018) reported that ectopic expression of C9ORF72 rescues neuronal cell death in iPSC-derived patient motor neurons, demonstrating the importance of C9ORF72 haploinsufficiency. Moreover, lit has been well documented that loss of C9ORF72 compromises the autophagy pathway, which is essential for aggregate clearance. Thus, the finding that decreased C9ORF72 expression increases DPR aggregate formation and toxicity in cells that express expanded repeats is expected. The lack of significant conceptual novelty in this set of experiments (Figure 3) is a major weakness of the paper. In these experiments, the authors concluded that the toxicity was caused by poly(GA) or poly(GP). However, in principle, poly(GR), which is highly toxic even when expressed at a low level, should be also expressed from the GGGGCC repeats. Tabet et al. (Nat. Commun. 2018) reported that mutations in the near-cognate start codon CUG greatly decreased the levels of both poly(GA) and poly(GR). Therefore, the observed toxicity cannot be attributed to poly(GA) alone. The authors should determine whether mutating AGG to decrease poly(GR) expression rescues the cell death phenotype. Similarly, it is surprising that a low level of poly(GP) induced toxicity. The authors should also consider whether a low level of poly(PR), which is also highly toxic, is produced from their antisense repeat RNA.
We agree with Referee #1 that our results could have been somehow “expected”. Indeed, as DPR proteins are forming aggregates, they were likely to be degraded by autophagy, and as C9ORF72 regulates this mechanism, a synergy of toxicity between DPR expression

and C9ORF72 loss of expression is not an entire surprise. However, this was not demonstrated formally, yet, and we believe that our work has some relevance and is worth to be reported as it reveals for the very first time the mechanisms of DPR proteins translation and their toxicity in ALS through a two-hit mechanism. Concerning the polyPR protein, we detect only trace amount of this DPR. However, concerning the polyGR protein, we have added new data in Figure 3 demonstrating that expressed under its natural AGG near-cognate initiation codon, its expression and toxicity are rather low, but increases upon siRNA-mediated silencing of the C9ORF72 protein. As a control, expression of polyGR under an artificial ATG start codon is highly toxic, as reported in recent studies (Choi et al., 2019; Lopez-Gonzales et al., 2019, Zhang et al., 2018). As a further control, deletion of the polYGR AGG near-cognate start codon abolishes neuronal cell death. These novel results highlight the importance of G4C2 translation into polyGR as potentially the main toxic event in ALS.

5. The authors tested different drugs to boost autophagy activity. It is useful to show that promethazine reduces DPR accumulation and toxicity when repeat plasmids are transiently overexpressed, since this is a clinically approved drug. However, it would be more convincing to show that this drug has therapeutic benefits in iPSC-derived patient neurons or in C9ORF72 BAC transgenic mice.
We agree that finding a drug that correct pathogenicity in iPS patient cells or animal models of ALS would be a true achievement and a breakthrough for the field. However, as we do not have these state-of-the-art cell and animal models in our lab, we feel that establishing these models and performing rescue experiments would be much too long, too expensive and too complex for the present work.

6. Overall, the writing seems too repetitive. For instance, in several places in the Results section, conclusions are restated again: "to summarize this part, ..." In some other parts, almost identical paragraphs are used to describe a different construct.
Indeed, writing and figure presentation of our draft required some improvement.

REFEREE #2:

Boivin et al manuscript contributes to the understanding of dipeptides-repeat proteins (DRP) production from C9ORF72 hexanucleotide repeat expansion causing ALS/FTD by profiling the relative abundance of the different DRPs and identifying upstream AUG and near-cognate codons driving protein synthesis in both sense (G4C2) and antisense (C4G2) frames. The study is very elegant by showing that levels of DRP produced from translation of the natural human C9ORF72 are subtoxic and only impact on neuronal viability in vitro when combined with C9ORF72 deficiency by knock-down. Decreased expression of C9ORF72 protein had been reported in ALS/FTD, but its relevance was not clear before the work presented here. The study by Boivin et al also links the synergistic effects of DPR and C9ORF72 loss-of-function to autophagy deficiency, being able to rescue DPR toxicity with autophagy inducers. The manuscript is excellent, very well written with a clear message. I recommend publication in EMBO J after the major points have been addressed.

1 - The authors must show that C9ORF72 knock-down actually leads to increased levels of DRP. Please provide analysis by Western-blot.
This important control was indeed missing from our previous draft. Immunoblots showing increase levels of polyGA, polyPG and polyGR DPR proteins upon siRNA-mediated silencing of C9ORF72 expression are now presented as figures 3B, 3C and 3D. As positive controls, inhibition of the autophagy pathway by overnight treatment with Bafilomycin A1 also enhances DPR expression.

2- The authors have to quantify DRP aggregation upon C9ORF72 knock-down. Ideally, quantification of GFP-positive aggregates is complemented with biochemical methods.

As suggested by Referee #2, we investigated the formation of DPR aggregates in control versus C9ORF72-reduced conditions. Upon C9ORF72 knockdown, brighter and more polyGA or polyPG aggregates were observed, which is rather expected as expression of these DPR proteins increases (immunoblots presented in the novel figure 3). However and to our surprise, we also observed that the P62 (SQSTM1) autophagic adaptor protein, normally surrounding aggregates to initiate their degradation by autophagy, is less recruited to the polyGA inclusions upon C9ORF72 knockdown. As a control, ubiquitin is present at DPR aggregates in both conditions, suggesting that ubiquitinylation of the proteins to degrade occurs normally. Identical observations were made for the polyPG aggregates. These results suggest that C9ORF72 plays a role in the initiation of autophagy at the step of aggregate recognition and/ or autophagosome formation. These novel data are presented in supplemental figure 3.

3- What is the role of C9ORF72 in DRP degradation by autophagy. Does it play a specific role in DRP and repeat proteins clearance? Please provide evidence of the actual role of C9ORF72 in the autophagy pathway?
As suggested by Referee #2, we investigated the autophagy of DPR proteins and observed a lesser recruitment of P62 to polyGA and polyPG aggregates upon C9ORF72 knockdown (cf. comment #2 above). Furthermore, we also investigated the autophagosome-lysosome traffic and observed that siRNA-mediated depletion of C9ORF72 leads to an increase lysosomal labeling (tested with LAMP1, LAMP2 and M6PR immunofluorescence, shown in the novel supplementary figure 3). These data are similar to various independent results reported recently (Sullivan et al., 2016; Amick et al., 2016; Corrionero et al., 2018, Lan et al., 2019), that all indicate an important role of C9ORF72 at lysosomes. Overall, our data suggest a complex function of C9ORF72 in the regulation of autophagy both at an early step (aggregate recognition?) and at late step (autophagosome-lysosome fusion?). Note that a similar complex regulatory function at early and late autophagic steps has been recently reported for SMCR8, the protein partner of C9ORF72 (Jung et al., 2017).

Overall, we would like to thank both Referees for their very helpful comments and suggestions.

2nd Editorial Decision 10th Oct 2019

While referee #2 is happy with the revised version, referee #1 still raises some issues with the manuscript. I have looked carefully at the raised issues and have also discussed them further with referee # 2. I appreciate the issues raised by referee #1, however I also note that some of them were not raised during initial round of review and that some are beyond the scope of this study.

I would like you to respond to the comments 1, 8 and 12. Should be easy to sort out

Regarding the requested controls in comments 4, 6: do you have any data on hand to address these comments. I don't think they are essential but also good suggestions. Would you get back to me to discuss this issue

We don't need any more data to address the remaining points, please respond to them in the point-by-point response.

REFEREE REPORTS:

Referee #1:

GGGGCC repeat expansion in the first intron of C9ORF72 is the most common genetic cause of both ALS and FTD. Expanded G4C2 repeats can be translated, from both sense and antisense directions, into dipeptide repeat (DPR) proteins that accumulate in brain cells of C9ORF72 patients and are thought to contribute to disease pathogenesis. However, how DPR proteins are synthesized

in patient cells remains largely unknown. In this study, the authors use reporter constructs to identify start codons for poly(GA), poly(GR), and poly(GP) production. While it has been reported by others in 2018 that CUG serves as the start codon for poly(GA) production, the identification of other start codons for poly(GR) and poly(GP) in these reporter constructs, if confirmed with additional control experiments, would be new. Then the authors demonstrate that reduced expression of the C9ORF72 protein itself results in the accumulation of DPR proteins in transfected cells, consistent with several recent papers reporting that C9ORF72 is involved in the autophagy pathway. Moreover, this result confirms what Shi et al. (Nat. Med. 2018) reported in their Fig. 5 that "C9ORF72 levels determine dipeptide repeat turnover". As one would expect, pharmacological activation of autophagy decreases the levels of DPR proteins, which correlates with increased cell survival.

1. In Figure 1, it would be helpful if the authors could please include a cartoon describing their reporter constructs in more detail. For instance, on Page 6, they state that "we cloned 80 G4C2 repeats embedded within the natural human C9ORF72 sequence". However, it is unclear exactly what this C9ORF72 sequence is. One has to check the Methods section to find out. Each construct should be sequenced, and the actual nucleotide sequence between the promotor and the G4C2 repeats should be provided in a supplementary file.

2. One limitation of the current study is that the authors expressed G4C2 repeats in the context of a 5' capped, polyA+ mRNA. Three papers published last year in Nature Communications performed many molecular experiments that are much more sophisticated than the experiments reported here to investigate whether repeats RNA translation is cap-dependent or not, and whether the template for translation is spliced intron or unspliced pre-mRNA. These important issues remain controversial. At least in the case of poly(GA), the current study is somewhat rudimentary and does not provide novel mechanistic insight beyond what have been published.

3. Another limitation of the current study is the cell type used. Laura Ranum already reported several years ago that the efficiency of repeats-associated translation in different frames depends on cell type. The authors here use HEK 293 cells but not postmitotic neurons. Thus, whether some conclusions are relevant to human disease or not is unclear.

4. On Page 6, the authors state that repeats "are predominantly translated into the GA frame and with a much lower efficiency in the GR frame". However, two critical control experiments are missing. The authors should quantify the relative abundance of G4C2 repeats-containing GFP mRNAs. Without this piece of information, one cannot compare relative translation efficiency of different constructs. Similarly, in all other experiments where the authors intend to compare relative translation efficiency, the mRNA levels of different constructs should be measured. For instance, in Figure 1D, different copies of G4C2 repeats may affect mRNA stability as well, but the authors only speculate the difference in protein levels is solely due to GA protein stability.

5. Moreover, what Figure 1A shows is the relative steady-state levels of poly(GA)-, poly(GP)-, or poly(GR)-containing GFP proteins, which is determined by both production and degradation. Without knowing the half-life of GA-GFP, GP-GFP and GR-GFP proteins, one cannot claim the relative translation efficiency of different frames only based on results in Figure 1A.

6. It is important that the relative abundance of GFP mRNAs in Figure 1E should be presented and western blot must be done using cells from the same transfection experiment.

7. Tabet et al. (Nat. Commun. 2018) reported that poly(GR) production also depends on the CUG codon, which is not in-frame with poly(GR). Thus, they proposed that ribosome frameshifting occurs near the first GGGGCC repeat. The result in Figure 1E could be interpreted as being consistent with the frameshifting hypothesis, since AGG>CGG mutation could potentially affect frameshifting. It is curious that the authors did not do mass-spec experiment to identify the first amino acid of the GR-GFP protein as they did for GA-GFP. This important experiment should be done. Secondly, the authors should mutate CUG in their GR-GFP construct and see what happens. This could be a very informative experiment too. Without these additional control experiments (including the one suggested in point #6), the conclusion that AGG is the start codon for poly(GR) is not quite convincing yet.

8. In Figure 1F, the carton is misleading because V2 mRNA does not contain repeats.

9. The identification of AUG as the start codon for poly(GP) translation from the antisense repeat RNA, if confirmed in patient cells, would be novel. First, similar control experiments as described above need to be done. Second, we do not know, in patient cells, how antisense repeat RNA is transcribed (where is the promoter) and whether it has a polyA tail or not. If the authors could provide these molecular information, that would be a useful contribution to the field. Otherwise, Figure 2 seems to be too rudimentary for such a high-profile molecular biology journal. Here the authors made an artificial construct to place the AUG-containing sequence and C4G2 repeats in the context of a polyA+ mRNA. Then it is totally expected that AUG in this construct can serve as a start codon. The authors should also use CRISPR to delete this AUG codon in C9ORF72 patient cells (lymphoblast cell lines, fibroblast cell lines, or iPSC lines) to investigate whether indeed it serves as the start codon for poly(PG).

10. Figure 3 also seems to be rudimentary, especially considering that Shi et al. (Nat. Med. 2018) already reported that C9orf72 regulates DPR protein turnover. The use of GT1-3 cell line but not mammalian primary neurons is also a concern.

11. In Figure 4, it is simply a correlation between decreased DPR proteins and increased cell survival after autophagy activation by drugs. The accumulation of other misfolded proteins and damaged organelles caused by C9orf72 KD should be decreased as well. Without additional experimental evidence, one could argue that the drugs simply rescued the detrimental effects of C9orf72 KD on autophagy in addition to DPR proteins.

12. The whole manuscript needs to be carefully edited. For instance, some wrong references are cited, especially in the Introduction section. Throughout the text, some statements are inaccurate, for instance, on Page 10, "...fluphenazine, decreases expression of poly(GA) and poly(PG)....." but these compounds should only affect the accumulation of DPR proteins.

Referee #2:

The authors addressed all points raised and improved their manuscript. This study is a substantial contribution to understand the role of C9 hexanucleotide expansion in the pathology of ALS/FTD. Publication is recommended after minor points are corrected.

Minor points:

1- In page 5, 'In conclusion, these results provide a molecular mechanism for the RAN translation of G4C2 sense and C4G2 antisense repeats, and support a double hit mechanism in ALS/FTD, where the reduce expression of C9ORF72 synergizes the accumulation and toxicity of DPR proteins.' The authors should rephrase the conclusion by saying that they found an alternative molecular mechanism to RAN translation that explains sense and antisense production of specific DPR.
2- In page 9, 'and suggest a complex function of C9ORF72 both at early (protein aggregates recognition?) and late steps (lysosomal fusion?) of the autophagic process.' Sounds too speculative and should say only 'and suggest a complex function of C9ORF72 both at early and late steps of the autophagic process.

2nd Revision - authors' response 28th Nov 2019

REFEREE #1

(Report for Author)
GGGGCC repeat expansion in the first intron of C9ORF72 is the most common genetic cause of both ALS and FTD. Expanded G4C2 repeats can be translated, from both sense and antisense directions, into dipeptide repeat (DPR) proteins that accumulate in brain cells of C9ORF72 patients and are thought to contribute to disease pathogenesis. However, how DPR proteins are synthesized in patient cells remains largely unknown. In this study, the authors use reporter constructs to identify start codons for poly(GA), poly(GR), and poly(GP) production. While it has been reported by others in 2018 that CUG

serves as the start codon for poly(GA) production, the identification of other start codons for poly(GR) and poly(GP) in these reporter constructs, if confirmed with additional control experiments, would be new. Then the authors demonstrate that reduced expression of the C9ORF72 protein itself results in the accumulation of DPR proteins in transfected cells, consistent with several recent papers reporting that C9ORF72 is involved in the autophagy pathway. Moreover, this result confirms what Shi et al. (Nat. Med. 2018) reported in their Fig. 5 that "C9ORF72 levels determine dipeptide repeat turnover". As one would expect, pharmacological activation of autophagy decreases the levels of DPR proteins, which correlates with increased cell survival.

1. In Figure 1, it would be helpful if the authors could please include a cartoon describing their reporter constructs in more detail. For instance, on Page 6, they state that "we cloned 80 G4C2 repeats embedded within the natural human C9ORF72 sequence". However, it is unclear exactly what this C9ORF72 sequence is. One has to check the Methods section to find out. Each construct should be sequenced, and the actual nucleotide sequence between the promotor and the G4C2 repeats should be provided in a supplementary file.
Sequence and scheme of the G4C2 repeats cloned into pcDNA3 is now presented as supplementary figure 1A.

2. One limitation of the current study is that the authors expressed G4C2 repeats in the context of a 5' capped, polyA+ mRNA. Three papers published last year in Nature Communications performed many molecular experiments that are much more sophisticated than the experiments reported here to investigate whether repeats RNA translation is cap-dependent or not, and whether the template for translation is spliced intron or unspliced pre-mRNA. These important issues remain controversial. At least in the case of poly(GA), the current study is somewhat rudimentary and does not provide novel mechanistic insight beyond what have been published.
Indeed, three recent and independent studies found that the polyGA DPR protein is produced through translation initiation at a CUG near-cognate codon located upstream of the G4C2 repeats (Green et al., 2017; Tabet et al., 2018; Sonobe et al., 2018). Furthermore, deletion of the genomic sequence containing this CUG codon abolishes polyGA expression in neurons differentiated from C9-iPS cells (Almeida et al., 2019). Our work fully confirms these studies, but also uncovers the mechanism of translation initiation for other DPR proteins (polyGR and polyPG), as well as a novel mechanism of toxicity where C9ORF72 reduced expression synergizes DPR toxicity. Thus, while our data on polyGA translation initiation are not novel, results about polyGR and polyPG expression and our proposed double hit mechanism are novel finding that we believe are worth reporting. Green and Tabet studies were previously cited in the manuscript, and we have now added references for Sonobe et Almeida work.

3. Another limitation of the current study is the cell type used. Laura Ranum already reported several years ago that the efficiency of repeats-associated translation in different frames depends on cell type. The authors here use HEK293 cells but not post mitotic neurons. Thus, whether some conclusions are relevant to human disease or not is unclear.
We performed additional transfections into neuronal N2A mouse cells and observed identical results, suggesting that data obtained in HEK293 cells are valid. However, we do agree that a definitive confirmation of these results would require CRISPR/Cas9 deletion of the genomic sequences initiating the translation of the polyGR and polyPG proteins in C9-iPS cells. However, genome editing in iPS cells would be much too long, too expensive and too complex for the present work.

4. On Page 6, the authors state that repeats "are predominantly translated into the GA frame and with a much lower efficiency in the GR frame". However, two critical control experiments are missing. The authors should quantify the relative abundance of G4C2 repeats-containing GFP mRNAs. Without this piece of information, one cannot compare relative translation efficiency of different constructs. Similarly, in all other experiments where the authors intend to compare relative translation efficiency, the mRNA levels of different constructs should be measured. For instance, in Figure 1D, different copies of G4C2 repeats may affect mRNA stability as well, but the authors only speculate the difference in protein levels is solely due to GA protein stability.

We added RT-qPCR quantification data showing that polyGA, polyGP and polyGR constructs are expressed at similar RNA levels (novel supplementary figure 1C). Thus, absence of polyGP protein is not caused by reduced RNA expression and/or a bias in RNA stability.

5. Moreover, what Figure 1A shows is the relative steady-state levels of poly(GA)-, poly(GP)- or poly(GR)-containing GFP proteins, which is determined by both production and degradation. Without knowing the half-life of GA-GFP, GP-GFP and GR-GFP proteins, one cannot claim the relative translation efficiency of different frames only based on results in Figure 1A.
Transfection of GFP-tagged constructs with polyGA, polyGP and polyGR cloned downstream of artificial ATG start codons show similar levels of protein expression in fluorescence and immunoblot. These controls, presented in supplementary figures 1D and 1E, suggest that the absence of detection of the polyGP protein expressed under its native C9ORF72 sequence is not caused by a bias of protein stability.

6. It is important that the relative abundance of GFP mRNAs in Figure 1E should be presented and western blot must be done using cells from the same transfection experiment.
As for polyGA vs polyGP and polyGR expression, RT-qPCR quantification shows that mutation of either the CTG or the AGG near-cognate codon does not affect RNA expression. These controls are presented as supplementary figures 1H and 1J.

7. Tabet et al. (Nat. Commun. 2018) reported that poly(GR) production also depends on the CUG codon, which is not in-frame with poly(GR). Thus, they proposed that ribosome frameshifting occurs near the first GGGGCC repeat. The result in Figure 1E could be interpreted as being consistent with the frameshifting hypothesis, since AGG>CGG mutation could potentially affect frameshifting. It is curious that the authors did not do mass-spec experiment to identify the first amino acid of the GR-GFP protein as they did for GA-GFP. This important experiment should be done. Secondly, the authors should mutate CUG in their GR-GFP construct and see what happens. This could be a very informative experiment too. Without these additional control experiments (including the one suggested in point #6), the conclusion that AGG is the start codon for poly(GR) is not quite convincing yet.
We failed to identify by mass spectrometry the N-terminal part of the polyGR protein. This might be due to its very short N-terminus as its potential methionine is right adjacent to the GR repeats.
We believe that expression of the polyGR protein is independent of the CTG codon as we tested G4C2 constructs deleted of the C9ORF72 sequence preceding the AGG codon, and these plasmids still express the polyGR protein. Furthermore, a recent work (Almeida et al., 2019) demonstrates that deletion of the genomic C9ORF72 sequence containing the CUG near-cognate codon abolishes polyGA expression, but not polyGR expression, in neurons differentiated from C9-iPS cells with a thousand G4C2 repeats. Hence, polyGR might not originate from a CUG-initiated frameshifting event, but is more likely produced through RAN translation initiating within the G4C2 repeats and/or from classical translation initiation to an upstream near-cognate codon located in the GR frame, as suggested by the present work.

8. In Figure 1F, the carton is misleading because V2 mRNA does not contain repeats.
This scheme is now corrected.

9. The identification of AUG as the start codon for poly(GP) translation from the antisense repeat RNA, if confirmed in patient cells, would be novel. First, similar control experiments as described above need to be done. Second, we do not know, in patient cells, how antisense repeat RNA is transcribed (where is the promoter) and whether it has a polyA tail or not. If the authors could provide these molecular informations, that would be a useful contribution to the field. Otherwise, Figure 2 seems to be too rudimentary for such a high-profile molecular biology journal. Here the authors made an artificial construct to place the AUG-containing sequence and C4G2 repeats in the context of a polyA+ mRNA. Then it is

totally expected that AUG in this construct can serve as a start codon. The authors should also use CRISPR to delete this AUG codon in C9ORF72 patient cells (lymphoblast cell lines, fibroblast cell lines, or iPSC lines) to investigate whether indeed it serves as the start codon for poly(PG).
Indeed, CRISPR/Cas9 editing of the C9ORF72 genomic sequence initiating polyPG translation will be required to prove its importance. However, such experiments are much too long, too expensive and too complex for the present work.

10. Figure 3 also seems to be rudimentary, especially considering that Shi et al. (Nat. Med. 2018) already reported that C9orf72 regulates DPR protein turnover. The use of GT1-7 cell line but not mammalian primary neurons is also a concern.
The work of Shi and collaborator uses DPR proteins expressed under artificial ATG start codons, resulting in higher expression and consequently higher toxicity. The present study uses DPR proteins expressed under their native initiating sequences, resulting in lower expression and toxicity, which we believe is important to report, especially for the polyGR protein.

11. In Figure 4, it is simply a correlation between decreased DPR proteins and increased cell survival after autophagy activation by drugs. The accumulation of other misfolded proteins and damaged organelles caused by C9orf72 KD should be decreased as well. Without additional experimental evidence, one could argue that the drugs simply rescued the detrimental effects of C9orf72 KD on autophagy in addition to DPR proteins.
Indeed, our data do not demonstrate a drug specificity to C9ORF72 haploinsufficiency and/or DPR accumulation. On the contrary, drug boosting autophagy could be of therapeutic interest in various diseases caused by accumulation of proteins prone to aggregation and thus resistant to proteasome clearance. This is now adequately discussed.

12. The whole manuscript needs to be carefully edited. For instance, some wrong references are cited, especially in the Introduction section. Throughout the text, some statements are inaccurate, for instance, on Page 10, "...fluphenazine, decreases expression of poly(GA) and poly(PG)....." but these compounds should only affect the accumulation of DPR proteins.
This is now corrected, and we would like to thank the Referee for his/her comments.

REFEREE #2

(Report for Author)
The authors addressed all points raised and improved their manuscript. This study is a substantial contribution to understand the role of C9 hexanucleotide expansion in the pathology of ALS/FTD. Publication is recommended after minor points are corrected.

Minor points:

1- In page 5, 'In conclusion, these results provide a molecular mechanism for the RAN translation of G4C2 sense and C4G2 antisense repeats, and support a double hit mechanism in ALS/FTD, where the reduce expression of C9ORF72 synergizes the accumulation and toxicity of DPR proteins.' The authors should rephrase the conclusion by saying that they found an alternative molecular mechanism to RAN translation that explains sense and antisense production of specific DPR.

2- In page 9, 'and suggest a complex function of C9ORF72 both at early (protein aggregates recognition?) and late steps (lysosomal fusion?) of the autophagic process.' Sounds too speculative and should say only 'and suggest a complex function of C9ORF72 both at early and late steps of the autophagic process.'

Thanks to Referee comments, these two points are now corrected.

3rd Editorial Decision 3rd Dec 2019

Thank you for submitting your revised manuscript to The EMBO Journal. I have taken a careful look at everything and all looks OK.

I am therefore very pleased to accept the manuscript for publication in The EMBO Journal.

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: Nicolas Charlet-Berguerand
Journal Submitted to: EMBO Journal
Manuscript Number: EMBOJ-2018-100574R

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 $\chi 2$ 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.

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?	Samples size were chosen so as detect a 10% variation
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	Not applicable, no animals were used in that study.
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	No samples 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.	Samples were randomized
For animal studies, include a statement about randomization even if no randomization was used.	Not applicable, no animals were used in that study.
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.	Samples analyzed were randomized
4.b. For animal studies, include a statement about blinding even if no blinding was done	Not applicable, no animals were used in that study.
5. For every figure, are statistical tests justified as appropriate?	Yes,statistical test are indicated and appropriate
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	yes the data meet a normal distribution as assessed by statistical test (Kolmogorov-Smirnov)

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repo

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title

http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tun

http://datadryad.org

http://figshare.com

http://www.ncbi.nlm.nih.gov/gap

http://www.ebi.ac.uk/ega

http://biomodels.net/

http://biomodels.net/miriam/
http://jjj.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

Is there an estimate of variation within each group of data?	Yes, there is an estimation of the variation between groups
Is the variance similar between the groups that are being statistically compared?	Yes, the variance are similar between analyzed groups

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).	All reference and/or clone number of the antibodies used in that study are provided in the material and method section.
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	Cell linesused in that study were long established and are referenced as so in the litterature

* 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.	No animal model in that study
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	no animals were studied here
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.	Not applicable

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	No human subjects in this study.
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.	Not applicable, no human samples in this study.
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	No human data or samples in this study.
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	Not applicable.
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	Not applicable.
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.	No applicable
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.	Not applicable

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	Not applicable
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).	Not applicable
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).	Not applicable, no human data in that 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.	Not applicable,

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.	Not applicable, all experiments were standart laboratory procedure that were performed under the French and European legislation.