CDD*press*

Check for updates

EDITORIAL

Extreme conservation of the poly-glutamine tract in huntingtin is related to neurodevelopmental functions: the "better" may become the "enemy of the good" in the course of evolution

The protein huntingtin was identified when its coding gene, now named HTT, was found to be the locus of the genetic defects underlying Huntington's disease (HD) [1]. HD is a rare inherited neurodegenerative disorder affecting young adults characterized by choreiform movements, cognitive defects, and psychiatric changes and is associated with brain atrophy, especially affecting the striatum and white matter tracks [2]. The disease is lethal within ~15 years from disease onset and there is no treatment to delay the onset or slow down the progression of HD. HD is due to a pathological expansion of a CAG repeat tract in the HTT gene, which leads to an expansion of a poly-glutamine (polyQ) tract in the N-terminal part of huntingtin. More than 5000 publications deal with the question of how mutant huntingtin triggers neurotoxicity. Many hypotheses have been suggested, pointing to multiple cellular and molecular culprits such as transcription, cellular transport, autophagy, energy metabolism, and synapses [3]. Anomalies of neurodevelopmental processes associated with brain maturation may also be at stake [4]. In comparison, only a small proportion of studies tried to decipher the physiological function(s) of huntingtin. However, without going into the details, compelling evidence indicates that wild-type huntingtin is expressed in many (if not all) cell types, and plays key roles in development and cellular homeostasis [5]. One of the most striking aspects of huntingtin biology is that the protein interacts with more than 300 molecular partners, emphasizing its probable implication in a variety of cellular processes.

The presence of the trinucleotide CAG repeat in huntingtin is intriguing. PolyQ repeat length is quite small (~5) in animal reign (e.g., 7 in mice, 11 in non-human primates). Only humans show a quite high number of glutamines, which is highly variable with 15–35 in the normal range. Above 35–36, gene carriers develop HD symptoms. The highly conserved polyQ tracts in huntingtin are reminiscent of the polyQ tracts found in development-related transcription factors. Since the knockout of the mouse *htt* gene (*hdh*) is lethal in embryos when the neural crest is forming [6], huntingtin probably has a central role in neurodevelopmental processes.

Why has evolution tightly maintained a stable number of glutamines in huntingtin in chordates? Why did evolution select longer CAG repeats in humans? Could it be that this "physiological expansion" is related to neurodevelopmental advantages associated with amplified complexity of the brain? If so, why does an increase in the CAG repeat above 35–36 precipitate the human brain toward a lethal neurodegenerative disorder? This appears paradoxical. To address these questions, a better

understanding of the wild-type function of the huntingtin polyQ tract is required.

In this issue of *Cell Death and Differentiation*, Iennaco et al. from Elena Cattaneo's laboratory at the Department of Biosciences, University of Milan, and the Istituto Nazionale Genetica Molecolare tackled this very challenging question of the biological function of the CAG repeats in the HTT gene [7]. More specifically, they asked whether the mild increase in CAG repeat length in the course of evolution could be the result of a positive selection linked to a neurodevelopmental function of the protein. They used bioinformatics analyses of ~100 HTT genes or orthologues from already available sequences deposited in public genetic databases and in order to enrich the analysis, they sequenced HTT exon 1 in more than 100 various species. One powerful approach for investigating whether a sub-domain of a protein plays a key functional role is to look into the evolution of the corresponding gene across taxa. High conservation across species, especially the near absence of non-synonymous substitution and a low rate of synonymous substitution of DNA bases indicate a purifying process during evolution. This suggests that a minor modification of the gene domain is strongly suppressed by evolution pressure, underlining that the function of the domain is crucial for the survival of the species. Cattaneo's team used this approach and found that CAG-encoded polyQ tract of huntingtin protein is under purifying selection during evolution. Interestingly, they compared the results from the analysis of the HTT genes with genes of other polyQ-containing proteins whose abnormal expansion also leads to neurodegenerative diseases such as ataxin-3 or ataxin-7. The "purification" process for ataxin genes was not so high as compared to that seen for HTT gene. In addition, the purifying process of the huntingtin polyQ tract is almost as high as that of polyQ-containing transcription factors that play crucial roles in neurodevelopment such as TBP (TATA-box-binding protein). Thus, conservation and purity in the huntingtin polyQ tract could have paved the way for its evolutionary expansion in mammals, possibly leading to an improved biological function.

The authors also provide experimental evidence that, in the mouse genetic background, the size of the polyQ tract is determinant for neurodevelopmental functions of huntingtin. The authors created knockin models of mouse embryonic stem (mES) cells with various repeats (from 0 to 17 Q). Looking at transcriptomic and morphologic developmental markers (lumen size of nestin-positive rosettes), they found that the number of Q correlates with stronger marks of neurodevelopment, zero glutamine showing the poorest marks. Intriguingly, the effect plateaued above 7 Q, which is the normal size of the polyQ tract in mice. Then, focusing on the mouse cell models with 0, 2, 4, and 7 Q, the authors
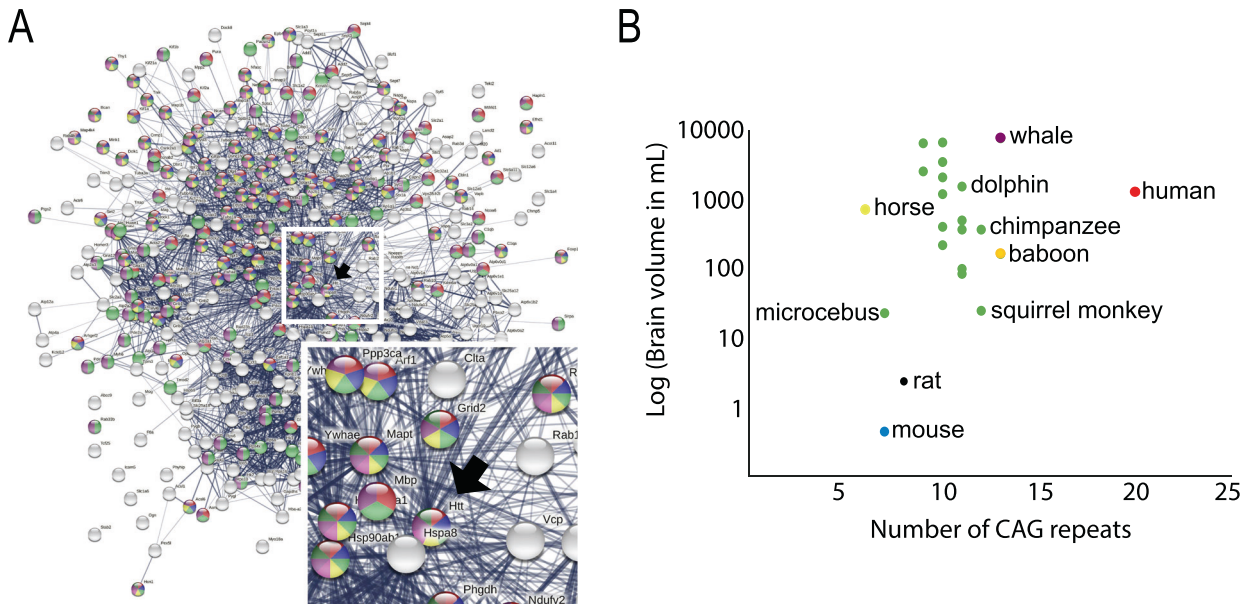
Fig. 1 Huntingtin interactome and the high conservation of its polyQ tract contribute to show a key role of the protein in neurodevelopment. A Representation of the network of proteins interacting with huntingtin using String (https://string-db.org). Each of the 409 proteins shown is represented by a gray ball if unrelated to development or colored when its biological function is related to development (according to Gene Ontology). Interactions between proteins are represented by gray lines. The bottom panel is a zoom of the network showing the enrichment of developmental (colored) proteins. The black arrow indicates HTT. The list of the 409 proteins is from a list of 412 proteins available in Shirazaki et al. [9]. **B** The graph shows the number of CAG repeats in different mammals across taxa (X-axis) and the corresponding logarithm of the volume of the brain of these animals (Y-axis). For clarity, only common names of well-known species are indicated. It is remarkable that, in the course of evolution in mammals, the number of CAG repeats remained higher than 6, for example, the number found in horses (*Equus caballus*). Also note that this number is relatively constant and always below 13, with the exception of the human huntingtin gene that has an average of 20 repeats. The number of CAG for each animal shown has been extracted using NCBI resources. Approximates of the size of the brain (volume) for each animal were from recent publications [13–15].

demonstrate that the increase in the number of Q in huntingtin provides a significant advantage for neurodevelopment: genes differentially regulated as a function of Q at different stages of differentiation are massively linked to early and late neurodevelopmental signaling.

These puzzling results further demonstrate the key role of huntingtin in development. However, the underlying molecular mechanisms await further studies to be understood. Not all cellular functions of wild-type huntingtin are "Q-dependent." Indeed, Iennaco et al. showed that the length of the polyQ tract was not changing cell death in mES cells, indicating that the anti-apoptotic effect of huntingtin is not dependent on the polyQ length. Many cellular components and processes might be involved. For example, an intriguing study points to energy metabolism. Marcy McDonald's team reported that polyQ tract length (<20) is correlated with the mitochondrial capacity to generate ATP [8], which is consistent with the fact that huntingtin interacts with many proteins with ATPase activity and/or with mitochondrial localization (e.g. Atp1a1, Cct1, Cct8, Hsp90ab1, Hspd1, Ndufs3) [9].

What are the implications of this exciting study by Iennaco et al.? It suggests that the size of the polyQ tracts in huntingtin was totally maintained to four Q in bonefishes, crocodiles, and birds and i could influence brain structure and function in mammalian species where the polyQ tract has generaly more than 4 Q. However, search in public databases and key publications that provide a fair estimation of brain volumes in mammalian species show that the relationship between the number of CAG repeats and brain size across different species is not simply linear (Fig. 1). Small mammals like mice and rats have 7 and 8 Q, respectively. In non-human primates, it remains stable around 9–11. In animals whose brains are ~3–4 times bigger than the human brain (i.e., elephants and Cetaceans such as killer

whales and dolphins), the HTT gene codes for only a 9–12 Q tract. Could it be that increasing the number of repeats in the course of evolution could have produced advantage to the human brain? MRI studies of brain anatomy and genetic determination of CAG repeat length in healthy population showed that the higher the CAG repeat number, the bigger the brain and these two readouts also correlate with the quantitative proxy of what could be considered "intelligence" in a human subject [10]. Thus, it is plausible that the polyQ tract in huntingtin in the human genetic background possesses unique developmental properties that are not present in other mammals. How the polyQ tract leads to better brain function in humans is unknown. Experiments in human models (iPSC, organoids) expressing human huntingtin with varying polyQ in the normal range would be extremely interesting in this regard.

If this speculative evolutionary process of progressive CAG repeat expansion occurred in the past when it could provide an advantage against environmental pressure, it is puzzling that more recently, the increase in CAG repeat in the HTT gene, although a very rare event, can be let out of control, leading to a lethal brain disorder. Addressing this question may help discover the covariant genes that allowed this dire event to occur. Recent GWAS data obtained from the analysis of large cohorts of mutant HTT gene carriers may contribute to identify the culprits [11]. Thus, the implications of the present study are important for HD. It emphasizes that the physiological role of wild-type huntingtin is crucial to shape our brain. Evolution may have used huntingtin to this objective. It requires further studies to be fully understood. In the context of rapidly developing HTT lowering therapeutic strategies [12], basic research on the function of wild-type huntingtin has to be pursued to optimize the efficacy and safety of future therapies for HD.

268

Emmanuel Brouillet ⓘ[1]✉

[1]Université Paris-Saclay, Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), Centre National de la Recherche Scientifique (CNRS), Molecular Imaging Research Center (MIRCen), Laboratoire des Maladies Neurodégénératives, Fontenay-aux-Roses, France. ✉email: emmanuel.brouillet@cea.fr

REFERENCES

1. The Huntington's Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell. 1993;72:971–83. https://doi.org/10.1016/0092-8674(93)90585-e
2. Walker FO. Huntington's disease. Lancet Lond Engl. 2007;369:218–28.
3. Bates GP, Dorsey R, Gusella JF, Hayden MR, Kay C, Leavitt BR, et al. Huntington disease. Nat Rev Dis Prim. 2015;1:15005. https://doi.org/10.1038/nrdp.2015.5.
4. Barnat M, Capizzi M, Aparicio E, Boluda S, Wennagel D, Kacher R, et al. Huntington's disease alters human neurodevelopment. Science. 2020;369:787–93. https://doi.org/10.1126/science.aax3338
5. Saudou F, Humbert S. The biology of huntingtin. Neuron. 2016;89:910–26. https://doi.org/10.1016/j.neuron.2016.02.003
6. Duyao MP, Auerbach AB, Ryan A, Persichetti F, Barnes GT, McNeil SM, et al. Inactivation of the mouse Huntington's disease gene homolog Hdh. Science. 1995;269:407–10. https://doi.org/10.1126/science.7618107
7. Iennaco R, Formenti G, Trovesi C, Rossi R, Zuccato C, Lischetti T et al. The evolutionary history of the polyQ tract in huntingtin sheds light on its functional proneural activities. Cell Death Diff. 2022; in press
8. Seong IS, Ivanova E, Lee JM, Choo YS, Fossale E, Anderson M, et al. HD CAG repeat implicates a dominant property of huntingtin in mitochondrial energy metabolism. Hum Mol Genet. 2005;14:2871–80. https://doi.org/10.1093/hmg/ddi319. PMID: 16115812
9. Shirasaki DI, Greiner ER, Al-Ramahi I, Gray M, Boontheung P, Geschwind BH, et al. Network organization of the huntingtin proteomic interactome in mammalian brain. Neuron Neuron. 2012;75:41–57. https://doi.org/10.1016/j.neuron.2012.05.024
10. Lee JK, Conrad A, Epping E, Mathews K, Magnotta V, Dawson JD, et al. Effect of trinucleotide repeats in the Huntington's gene on intelligence. EBioMedicine. 2018;31:47–53. https://doi.org/10.1016/j.ebiom.2018.03.031
11. Genetic Modifiers of Huntington's Disease (GeM-HD) Consortium. CAG repeat not polyglutamine length determines timing of Huntington's disease onset. Cell. 2019;178:887–900.e14. https://doi.org/10.1016/j.cell.2019.06.036
12. Tabrizi SJ, Flower MD, Ross CA, Wild EJ. Huntington disease: new insights into molecular pathogenesis and therapeutic opportunities. Nat Rev Neurol. 2020;16:529–46. https://doi.org/10.1038/s41582-020-0389-4
13. Ridgway SH, Carlin KP, Van Alstyne KR, Hanson AC, Tarpley RJ. Comparison of dolphins' body and brain measurements with four other groups of cetaceans reveals great diversity. Brain Behav Evol. 2016;88:235–57. https://doi.org/10.1159/000454797
14. Isler K, Kirk CE, Miller JM, Albrecht GA, Gelvin BR, Martin RD. Endocranial volumes of primate species: scaling analyses using a comprehensive and reliable data set. J Hum Evol. 2008;55:967–78. https://doi.org/10.1016/j.jhevol.2008.08.004
15. Veitschegger K. The effect of body size evolution and ecology on encephalization in cave bears and extant relatives. BMC Evol Biol. 2017;17:124 https://doi.org/10.1186/s12862-017-0976-1

COMPETING INTERESTS

The author declares no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Emmanuel Brouillet.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.