Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Is There Any Influence of *TBP* CAG / CAA Repeats on Huntington's Disease Age at Onset?

Iane dos Santos da Silva
Universidade Federal do Estado do Rio de Janeiro (UNIRIO)
Thays Andrade Apolinário
Programa de Pós-Graduação em Neurologia, (UNIRIO)
Luciana de Andrade Agostinho (polucita@yhaoo.com.br)
Programa de Pós-Graduação em Neurologia, (UNIRIO) https://orcid.org/0000-0001-5604-8386
Carmen Lúcia Antão Paiva
Universidade Federal do Estado do Rio de Janeiro (UNIRIO)

Research Article

Keywords: Huntington's Disease, HTT, TBP, modifier gene, Brazilian sample

Posted Date: July 9th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-680274/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Huntington's disease (HD) is a genetic neurodegenerative progressive and fatal disease characterized by motor disorder, cognitive impairment and behavioral problems, caused by expanded repeats of CAG trinucleotides in the *HTT* gene. The aim of this study was to investigate the influence of *TBP* gene CAG/CAA repeats in conjunction with *HTT* gene CAG repeats, on the age of HD onset in Brazilian individuals. Individuals diagnosed as molecularly negative for HD, presented 29-39 *TBP* CAG/CAA (mean = 36 ± 2; median = 36). The most frequent allele had 36 repeats. The heterozygosity was 84%. In individuals diagnosed as molecularly positive for HD, a range of 25-40 *TBP* CAG/ CAA was found (mean = 36 ± 2; median = 36). The most frequent *TBP* allele had 38 repeats and the heterozygosity was 81%. We also conducted *TBP* direct Sanger sequencing of some samples which demonstrated other *TBP* structures different from the wild-type. The *HTT* expanded CAG and *TBP* CAG/CAA repeat sizes jointly explained 66% of the age at onset (AO) in our HD patients. The strongest variable in the model associated to AO was the number of expanded *HTT* CAG repeats. The difference between the association of HD AO with *HTT* expanded CAG together with *TBP* CAG / CAA and the association of HD AO with *HTT* expanded CAG was 0.001 (ΔR^2). Therefore, we found a weak association (0.1%) of *TBP* CAG/CAA repeats on HD AO, if any.

1. Introduction

Huntington's disease (HD) (OMIM: 143100) is a genetic, neurodegenerative, progressive, and fatal disease, characterized by motor disorder, cognitive impairment and behavioral problems. It is caused by a CAG trinucleotide repeat expansion in the first exon of the *HTT* gene (GENE ID: 3064) located on chromosome 4p16.3 (Group 1993).

It is known that the most critical determinant of HD age at onset (AO) is the number of CAG repeats in the *HTT* gene which accounts for about 70% of the AO(Djousse et al. 2003; Rubinsztein et al. 1997; Wexler et al. 2004). The other 30% are assigned to the modifier genes and/or environmental factors (Rubinsztein et al. 1997).

A disease modifier gene is that gene whose structure, or expression, alters the expression of phenotypes associated with the primary mutation that causes the disease. The main strategy used to search modifier genes has been the investigation of genes linked to metabolic processes or to molecular pathways allegedly involved in HD (Gusella and MacDonald 2009).

The importance of this article resides in the fact that there are two main reasons for searching modifier genes in association with Huntington’s disease. Knowing the genetic modifiers and how they act would provide a better understanding of the disease, as well as a better genetic counseling. The knowledge about genetic modifier genes is also important for selecting a more homogeneous population, aiming to allow better clinical trial for testing a candidate therapeutic drug (Gusella et al. 2014).

The *TBP* gene (Gene ID: 6908) is a supposed modifier gene which encodes for a TATA box binding protein. This protein is a transcriptional factor required for transcription initiation. Its N-terminal domain is formed by a glutamine track encoded by CAG / CAA repeats. Normal individuals have up to 42 CAG / CAA repeats and individuals with more than 42 repetitions develop spinocerebellarataxia 17, also known as HDL4 (Huntington's disease like 4) (Nakamura et al. 2001) (OMIM: 607136). In 2001, an inherited ataxy was first described which was associated with the elongation of an unstable (CAG)n element combined with the loss of CAA interruptions within the complex repetitive sequence of *TBP* gene (Zuhlke et al. 2001).

TBP protein has been suggested as a participant in the pathogenesis of HD (van Roon-Mom et al. 2002) as well as in Alzheimer's disease (Reid et al. 2009) and to be associated with the risk of Schizophrenia (Ohi et al. 2009). Furthermore, *TBP* CAG/CAA repeats were also associated with the AO of Spinocerebellar Ataxia 7 (Tezenas du Montcel et al. 2014).

The aim of this study was to investigate the influence of a candidate modifier gene (*TBP*) in conjunction with the *HTT* gene on the age at onset of the HD symptoms in Brazilian individuals.

2. Material And Methods

2.1 Individuals

The subjects were selected from the Clinical Genetics Service held at the University Hospital Gaffrée and Guinle (HUGG/UNIRIO) and from families enrolled in the Brazilian Huntington's Association - Brazil. They were all Brazilian by birth. One hundred and four individuals from 19 unrelated families were investigated: 51 women and 53 men. All subjects were from the following Brazilian states: 48% from Rio de Janeiro, 38% from Minas Gerais, 9% from Espírito Santo, 2% from São Paulo, 1% from Bahia, 1% from Pará and 1% from Maranhão.

The AO of the disease (when the manifestation of motor symptoms began) was self-reported by the affected individual, by his/her caregiver, or by the family. All the participants in this investigation signed the "Informed Consent" form. This study was approved by the HUGG Research Ethics Committee, Rio de Janeiro, Brazil, under the number CAAE 26387113.1.0000.5258.

2.2 DNA extraction

DNA samples were obtained from 1-3 mL peripheral blood (in EDTA tube), swabs or scrapings from oral mucosa.

DNA extraction was performed according to the extraction kit protocol (Illustra Blood Genomic Prep Mini Spin, GE Healthcare, Buckinghamshire, UK).

2.3 Analysis of *HTT* CAG region

The number of CAG repeats in the *HTT* gene was determined according to the protocol suggested by Agostinho et al (2012) (Agostinho Lde et al. 2012). The following primers were used: HD1 (forward, 6 FAM 5'-TGGCGACCCTGGAAAAGCTGAT-3') and HD3 (reverse, 5'-GCGGTGGCGGCTGTTGCTGCT-3') at the concentration of 10 ρmoles/uL. The mixture for PCR was prepared with 1uL of each primer, plus 6.25 uL of GoTaq® Green Master Mix (containing 1.5 mM $MgCl_2$ and 200 uM of each dNTP) (Promega Wisconsin, USA), and 4.25 uL of DNA (20-100 ηg/uL); at a final volume of 12.5 uL. The conditions for PCR were: 1 cycle at 94°C for 5 min; and 35 cycles of 94°C for 1 min, 59.1ºC for 1 min and 72°C for 2 min; followed by a final cycle of 72ºC for 50 min.

The amplicons were detected by automatic capillary electrophoresis for fragment analysis (ABI 3500 or ABI 3070 Life Technologies, Applied Biosystems, Foster City, CA, USA) and the results were analyzed with the Gene Mapper V4.1 software.

It is important to mention that some samples had the CAG region sequenced, as suggested by Andrew et al. (1994) (Andrew et al. 1994) and were used as size standards for the fragment analysis.

2.4 Analysis of *TBP* CAG / CAA region

For the PCR reaction, the primers suggested by Koide et al. (1999) (Koide et al. 1999) were used: TBP-F (forward, 6 FAM 5'-GACCCCACAGCCTATTCAGA-3') and TBP-R (reverse, 5'-TTGACTGCTGAACGGCTGCA-3'). The temperature cycles were those described by Ohi et al. (2009) (Ohi et al. 2009): 1 cycle at 94° C for 10 min; 30 cycles of 94°C for 1 min, 60°C for 1 min and 72°C for 1 min; and a final cycle of 72ºC for 10 min. The primers were used at a concentration of 5 ρmoles / uL. The reaction mixture for PCR contained 0.5 uL of each primer, plus 6.25 uL of GoTaq® Green Master Mix (containing 1.5 mM $MgCl_2$ and 200 uM of each dNTP) (Promega, Wisconsin, USA); and 2 uL of 10-50 ηg/uL of DNA, and 3.25 uL of DNAase free water to a final volume of 12.5 uL.

The amplicons were detected by automatic capillary electrophoresis for fragment analysis (ABI 3500 or ABI 3070 Life Technologies, Applied Biosystems, Foster City, CA, USA) and the results were analyzed with the Gene Mapper V4.1 software.

In order to validate this assay, three samples had their respective CAG / CAA regions sequenced. For that, the following primers were used: forward (5'-AGCCAGCCTAACCTGTTTTTC-3') and reverse (5'-TGCGGTACAATCCCAGAACT-3'). The sequenced fragments were used as size standards for fragment analysis.

3. Results

Out of 104 subjects, 72 showed positive molecular result for HD (≥ 36 CAG in the *HTT* gene) and 32 showed negative molecular results (<36 CAG repeats).

Sixty four of the 72 molecularly positive individuals reported their HD AO, which ranged from 18 to 67 years (mean 42 ± 10).

3.1 Number of *HTT* CAG repeats

Among the 32 subjects diagnosed as molecularly negative for HD, 26 subjects were heterozygous and six homozygous for the CAG region. The number of repeats in this group of patients ranged from 12-30 repetitions (mean: 19 ± 4; median: 17). All homozygotes showed 17 CAG repeats in both alleles. Regarding the classification of alleles, 94% (60 alleles) were classified as normal and 6% (4 alleles) as normal mutable (or intermediate alleles) (Table 1 and Figure 1).

Among the 72 individuals diagnosed as molecularly positive for expanded *HTT* all of them were heterozygous with an expanded allele (≥ 36 CAG repeats) and a non-expanded allele (<36 CAG repeats). The number of CAG repeats of non-expanded alleles ranged from 14-30 (mean:19 ± 4; median: 17) and the most frequent allele had 17 repeats. Repetitions in the expanded alleles ranged from 39-62 (mean:45 ± 4; median:44), and the most frequent expanded allele had 44 CAG repeats. Concerning the non-expanded alleles, 94% (68 alleles) were classified as normal and 6% (4 alleles) as mutable normal (intermediate). Among the HD expanded alleles, 100% (72 alleles) were completely penetrant CAG alleles (Table 1 and Figure 2).

3.2 Number of *TBP* CAG/ CAA repeats

Individuals diagnosed as molecularly negative for HD presented with a range of 29-39 *TBP* CAG / CAA repeats (mean: 36 ± 2; median: 36) and their most frequent allele had 36 repeats (Table 1 and Figure 3). The heterozygosity was 84% (Table 1).

A range of 25-40 *TBP* CAG/CAA repeats (mean: 36 ± 2; median: 36) was found in individuals diagnosed as molecularly positive for HD. The most frequent *TBP* allele had 38 repeats (Table 1 and Figure 4) and the heterozygosity was 81% (Table 1).

There was no statistically significant difference between the number of *TBP* CAG / CAA repeats of the HD affected (expanded *HTT*) and non-affected individuals (normal *HTT*).

The sequence of the *TBP* gene was determined in three samples by direct sequencing (Sanger method). The *TBP* structures were different when compared with the wild-type allele according to reference GCh37.p13 found in the NCBI database, as well as when compared with the basic structure proposed by Gostout (1993) (Gostout et al. 1993), who categorized the *TBP* structures according to five regions: I, II, III, IV and V (Table 2).

We found four *TBP* different structures in our Brazilian sample: two HD patients had structures number 8, 9 and 10 and one individual who bore an intermediate *HTT* allele (with 30 uninterrupted CAG repeats) had the structure number 11 (Table 2).

One HD patient harboring 48 *HTT* CAG repeats was homozygous for the *TBP* structure shown as number 9 (Table 2). Another HD patient with 48 *HTT* CAGs had two different *TBP* structures: represented by number 8 (normal allele) and number 10 (expanded allele).

3.3 Influence of *TBP* CAG / CAA repeats associated with the *HTT* CAG repeats on the HD AO

The *HTT* expanded CAG and *TBP* CAG/CAA repeat sizes jointly explained 66% of the AO in our HD patients (table 3). The strongest variable in the model associated to AO was the number of expanded *HTT* CAG repeats. Table 3 shows the influence of *HTT* expanded CAG repeats together with *TBP* CAG / CAA repeats on the HD AO. The difference between the association of HD AO with *HTT* expanded CAG together with *TBP* CAG / CAA and the association of HD AO with *HTT* expanded CAG was 0.001 (ΔR^2). Therefore, we found a weak association (0.1%) of *TBP* CAG/CAA repeats on HD AO, if any.

4. Discussion

Although the expanded number of *HTT* CAG repeats is considered to be the major determinant of the AO of HD symptoms, other factors act together with the *HTT* CAG expansions. These factors can be stochastic, environmental or genetic. Although all these three factors may be involved to some extent in the determination of the AO, the genetic factors can be more easily studied due to the advances in molecular techniques (Gusella and MacDonald 2009).

Variants in CAA-CAG sequence downstream the *HTT* (CAG)n repeats can influence the AO in the following way: the exchange of one adenine nucleotide in a CAA codon (CAA>CAG), turning the region into an uninterrupted CAG region, is associated with dramatically earlier AO, despite the same polyglutamine length in individuals with the interrupting penultimate CAA codon. Besides that, another variant in this region (where the CAA-CAG sequence is duplicated) was associated with later AO. Identification of these cis-acting modifiers have potentially important implications for genetic counselling in HD-affected families (Wright et al. 2019).

A study shows that increased *FAN1* expression of a nuclease involved in DNA interstrand cross-link repair, is significantly associated with delayed AO and slower progression of HD, suggesting *FAN1* is protective in the context of an expanded *HTT* CAG repeat (Goold et al. 2019).

In this investigation, we found that expanded *HTT* alone is responsible for 65.7% of the AO. Furthermore, other studies have shown that other mechanisms can also be involved such as epigenetic-chromatin deregulation, as well as RNA toxicity and transcription aberrations (Marti 2016; Nalavade et al. 2013). The microtubule associated protein tau (MAPT), which is involved in several neurodegenerative disorders, has also been implicated in HD (Vuono et al. 2015). It is also important to mention that the expanded huntingtin (mHTT) shows a pleiotropic effect, as it is broadly present in different cellular compartments (e.g. cytosol, nucleus, mitochondria) as well as in all cell types of the human body at all developmental stages (Bassi et al. 2017).

In addition, we searched the number of *TBP* CAG / CAA repeats and its influence on the AO of motor symptoms in 72 HD patients. TBP has been suggested as a protein that plays an important role in the HD pathogenesis. The normal form of TBP has been found, along with the mHTT, at higher levels in the brains of HD patients than in control group (van Roon-Mom et al. 2002). We found 84% of *TBP* heterozygosity among individuals who were molecularly negative for HD and 81% for molecularly positive patients. A similar result described 78.5% of heterozygosity for the *TBP* alleles (Tomiuk et al. 2007).

In one of the first studies to investigate the allelic frequency of *TBP* CAG / CAA repeats showed that, in Algeria, these repeats varied from 32-39, and the most frequent alleles had 38 repeats. Black South Africans and Sub-Saharan Africans had alleles with 33-39 repeats, and their most frequent allele had 35 repeats. Indian people had CAG / CAA repeats ranging from 27 to 39 and their most frequent allele had 38 repeats (Rubinsztein et al. 1996).

An investigation of the number of *TBP* CAG / CAA repeats in different neurodegenerative diseases (Alzheimer, Parkinson, Huntington, Schizophrenia and different Ataxias) revealed that the *TBP* repeats, in the group of patients, varied from 27-46 (mean 36 ± 2; median 36). For the control group, the *TBP* repetitions ranged from 30 to 43 (mean 37 ± 1; median of 36) (Wu et al. 2005). On the other hand, in our study, we observed that molecularly tested HD negative individuals had 29-39 *TBP* CAG / CAA repeats and the most frequent *TBP* allele had 36 repeats. While molecularly positive HD individuals had 25-40 *TBP* CAG / CAA repeats and the most frequent allele had 38 CAG / CAA.

De novo expansions of CAG repeats in HD and SCA17 patients have been reported to occur from intermediate alleles, which contain uninterrupted pure CAG repeats (Myers et al. 1993). It is known that the *TBP* CAG/CAA repeat track is highly variable, the number of glutamines ranges from 25 to 42 in the American population. The CAA interruptions in *TBP* contribute to the stability of CAG polymorphic regions (II and IV) as suggested by Gostout et al. (1993) (Gostout et al. 1993), (table 2).

The implication of new modifier genes on the modulation at AO in Tunisian HD patients was investigated, such as the *TBP* (CAG/CAA) polymorphisms (Hmida-Ben Brahim et al. 2014). The authors took into account the genetic polymorphisms such as expanded *HTT* (CAG) together with *TBP* (CAG/CAA), as well as the expanded *HTT* CAG polymorphisms independently. They observed an increase of only 0.2% (ΔR^2) of variation in the AO in Tunisian HD patients when *TBP* was included (Hmida-Ben Brahim et al. 2014). We also investigated the same influence of *TBP* on the HD AO and observed a similar weak association (ΔR^2= 0.1%). It is important to mention that these two searches are comprised of Brazilian or Tunisian individuals who have been living in different environments and exposed to different mechanisms of selection. Furthermore, the two studies also differ in the sample size: a small sample size (n=15) (Hmida-Ben Brahim et al. 2014) and our sample of 144 *TBP* alleles (n=72 HD patients) (Table 3).

The normal number of *TBP* CAG / CAA repeats was described as modulator of the AO of spinocerebellar ataxia 7 (SCA7), with evidence that the largest number of CAG / CAA repeats leads to a

decrease in the age of SCA7 onset. This ataxia is caused by *ATXN7* gene which also codes for a track of glutamine repeats (Tezenas du Montcel et al. 2014). Furthermore, a study indicates that genotypes with more than 35 *TBP* CAG / CAA repeats have been associated with risk of schizophrenia, AO of the disease and prefrontal cortical function in the Japanese population (Ohi et al. 2009).

It is worth mentioning that the first polymorphic region of *TBP* structure (Region II) shows sequences with different CAG numbers from 6 to 47 (table 2). The NCBI reference (GRCh37.p13) only reports the region II as a sequence of eight repeats $(CAG)_8$ and the CAG variations were reported as Inframe deletions; however they do not have clinical significance presented by ClinVar database.

We suggest that subjects bearing *TBP* structure 7.2 could have two SNPs (rs55736770 and rs62430309) in region III which would lead to uninterrupted CAG track with 45 to 47 repeats. Although considered as a region of higher instability (Imbert et al. 1994), variant rs55736770 is reported by ClinVar database as likely-benign. In addition, the clinical significance of rs62430309 is not reported in ClinVar database.

The alternative sequence (*1) localized between regions III and IV (rs112083427), present in 7.2 structure, has a mutation (CAG CA(G>A) CAG_{17}) which is a variant classified as VUS by ClinVar. This *TBP* region has a polulational frequency of 24.9% in Europe, 23.4% in East Asia, 12% in Africa and 34% in the American Continent according to the Genome Aggregation Database (gnomAD).

The second polymorphic *TBP* region IV, according to GRCh37.p13, may vary from 10 to 24 repeats with benign or likely-benign clinical significance. Imbert et al (1994) (Imbert et al. 1994) consider normal *TBP* alleles with a maximum of 21 triplets at region IV. Table 2 shows that the second polymorphic region has a variation of 9 to 31 CAGs but structure 7.2 which has a expansion >40 CAGs. Our patient, who bore the *TBP* structure 8, had 17 uninterrupted CAGs between regions I and V. The variants which have more than 24 CAG repeats are not reported in the NCBI database.

The alternative region (*2), located between regions IV and V, was only described in a single study (structure number 3.2- table 2) where five G>A led to a CAG sequence interrupted by five CAA (synonymous variant). As a consequence of its presence, region IV has a smaller number of CAG repeats (9 repeats). Although, the variant sequence is reported in the NCBI database it does not have any clinical significance in ClinVar database. Its frequency is less than 1% in the world population (1000 Genome Bank).

Considering *TBP* as candidate to be a modifier gene, neither the *TBP* variants observed in the Brazilian HD patients, nor those reported by Goustout et al. (1993) (Gostout et al. 1993), are available in the ClinVar database. Consequently, we could not compare these different structures in relation to their clinical significance.

In conclusion, the rate-limiting mechanisms of AO in HD still remain elusive: many different processes are commonly disrupted in HD cell lines and animal models, as well as in HD patient cells (Bassi et al. 2017).

It would be important if future searches further investigated the association of the size of *TBP* CAG/CAA expansions with or without CAA in the CAG polymorphic region of this gene, once the long CAG expansions have already been associated with loss of genetic stability.

Declarations

Funding

This work was supported by CAPES (scholarships for I.S.S, T.A.A and L.A.A. during their post-graduate studies), FAPERJ (grant n° E-26/010.000956/2016 to C.L.A.P), UNIRIO and FINEP.

Conflicts of interest/Competing interests

The authors have no conflict of interest to disclose

Ethics approval

This study was approved by the HUGG Research Ethics Committee, Rio de Janeiro, Brazil, under the number CAAE 26387113.1.0000.5258.

Consent to participate

All the participants in this investigation signed the "Informed Consent" form.

Consent for publication

Participant authorizes, in consent form, publication of data in journal article form without having their identity revealed.

Availability of data and material (data transparency)

Not applicable

Code availability (software application or custom code)

Not applicable

Authors' contributions

Da Silva and Apolinário were responsible for the experiments. All authors contributed to the result analysis, statistical analysis and write the manuscript. Agostinho and Paiva reviewed the final version of the article.

ACKNOWLEDGEMENTS

Many thanks are due to the DNA Sequencing and Fragment Analysis Platform of IOC / FIOCRUZ / RJ and to the Huntington Disease Research Group, the Centre for Molecular Medicine and Therapeutics (UBC), Vancouver, Canada, for helping sequencing the amplicons. We also thank Dr. Suely Rodrigues dos Santos, Dr. Mariana Spitz, Dr. João Santos Pereira and Dr. Luiz Felipe Vasconcellos for the referral of some patients, as well as the Brazilian Huntington's Disease Association (ABH). Finally we thank all the patients and their families for participating in this investigation

References

1. Agostinho Lde A et al. (2012) Haplotype analysis of the CAG and CCG repeats in 21 Brazilian families with Huntington's disease J Hum Genet 57:796-803 doi:10.1038/jhg.2012.120
2. Andrew SE, Goldberg YP, Theilmann J, Zeisler J, Hayden MR (1994) A CCG repeat polymorphism adjacent to the CAG repeat in the Huntington disease gene: implications for diagnostic accuracy and predictive testing Hum Mol Genet 3:65-67 doi:10.1093/hmg/3.1.65
3. Bassi S, Tripathi T, Monziani A, Di Leva F, Biagioli M (2017) Epigenetics of Huntington's Disease Adv Exp Med Biol 978:277-299 doi:10.1007/978-3-319-53889-1_15
4. Djousse L et al. (2003) Interaction of normal and expanded CAG repeat sizes influences age at onset of Huntington disease Am J Med Genet A 119A:279-282 doi:10.1002/ajmg.a.20190
5. Fujigasaki H et al. (2001) CAG repeat expansion in the TATA box-binding protein gene causes autosomal dominant cerebellar ataxia Brain 124:1939-1947 doi:10.1093/brain/124.10.1939
6. Goold R et al. (2019) FAN1 modifies Huntington's disease progression by stabilizing the expanded HTT CAG repeat Hum Mol Genet 28:650-661 doi:10.1093/hmg/ddy375
7. Gostout B, Liu Q, Sommer SS (1993) "Cryptic" repeating triplets of purines and pyrimidines (cRRY(i)) are frequent and polymorphic: analysis of coding cRRY(i) in the proopiomelanocortin (POMC) and TATA-binding protein (TBP) genes Am J Hum Genet 52:1182-1190
8. Group THsDCR (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes Cell 72:971-983 doi:10.1016/0092-8674(93)90585-e
9. Gusella JF, MacDonald ME (2009) Huntington's disease: the case for genetic modifiers Genome Med 1:80 doi:10.1186/gm80
10. Gusella JF, MacDonald ME, Lee JM (2014) Genetic modifiers of Huntington's disease Mov Disord 29:1359-1365 doi:10.1002/mds.26001
11. Hmida-Ben Brahim D et al. (2014) Modulation at age of onset in tunisian huntington disease patients: implication of new modifier genes Genet Res Int 2014:210418 doi:10.1155/2014/210418
12. Imbert G, Trottier Y, Beckmann J, Mandel JL (1994) The gene for the TATA binding protein (TBP) that contains a highly polymorphic protein coding CAG repeat maps to 6q27 Genomics 21:667-668 doi:10.1006/geno.1994.1335
13. Koide R et al. (1999) A neurological disease caused by an expanded CAG trinucleotide repeat in the TATA-binding protein gene: a new polyglutamine disease? Hum Mol Genet 8:2047-2053

doi:10.1093/hmg/8.11.2047

14. Marti E (2016) RNA toxicity induced by expanded CAG repeats in Huntington's disease Brain Pathol 26:779-786 doi:10.1111/bpa.12427
15. Myers RH et al. (1993) De novo expansion of a (CAG)n repeat in sporadic Huntington's disease Nat Genet 5:168-173 doi:10.1038/ng1093-168
16. Nakamura K et al. (2001) SCA17, a novel autosomal dominant cerebellar ataxia caused by an expanded polyglutamine in TATA-binding protein Hum Mol Genet 10:1441-1448 doi:10.1093/hmg/10.14.1441
17. Nalavade R, Griesche N, Ryan DP, Hildebrand S, Krauss S (2013) Mechanisms of RNA-induced toxicity in CAG repeat disorders Cell Death Dis 4:e752 doi:10.1038/cddis.2013.276
18. Ohi K et al. (2009) TATA box-binding protein gene is associated with risk for schizophrenia, age at onset and prefrontal function Genes Brain Behav 8:473-480 doi:10.1111/j.1601-183X.2009.00497.x
19. Reid SJ, Whittaker DJ, Greenwood D, Snell RG (2009) A splice variant of the TATA-box binding protein encoding the polyglutamine-containing N-terminal domain that accumulates in Alzheimer's disease Brain Res 1268:190-199 doi:10.1016/j.brainres.2009.03.004
20. Rubinsztein DC, Leggo J, Chiano M, Dodge A, Norbury G, Rosser E, Craufurd D (1997) Genotypes at the GluR6 kainate receptor locus are associated with variation in the age of onset of Huntington disease Proc Natl Acad Sci U S A 94:3872-3876 doi:10.1073/pnas.94.8.3872
21. Rubinsztein DC, Leggo J, Crow TJ, DeLisi LE, Walsh C, Jain S, Paykel ES (1996) Analysis of polyglutamine-coding repeats in the TATA-binding protein in different human populations and in patients with schizophrenia and bipolar affective disorder American Journal of Medical Genetics 67:495-498 doi:10.1002/(sici)1096-8628(19960920)67:5<495::Aid-ajmg12>3.0.Co;2-i
22. Silveira I et al. (2002) Trinucleotide repeats in 202 families with ataxia: a small expanded (CAG)n allele at the SCA17 locus Arch Neurol 59:623-629 doi:10.1001/archneur.59.4.623
23. Tezenas du Montcel S et al. (2014) Modulation of the age at onset in spinocerebellar ataxia by CAG tracts in various genes Brain 137:2444-2455 doi:10.1093/brain/awu174
24. Tomiuk J et al. (2007) Repeat expansion in spinocerebellar ataxia type 17 alleles of the TATA-box binding protein gene: an evolutionary approach Eur J Hum Genet 15:81-87 doi:10.1038/sj.ejhg.5201721
25. van Roon-Mom WM, Reid SJ, Jones AL, MacDonald ME, Faull RL, Snell RG (2002) Insoluble TATA-binding protein accumulation in Huntington's disease cortex Brain Res Mol Brain Res 109:1-10 doi:10.1016/s0169-328x(02)00450-3
26. Vuono R et al. (2015) The role of tau in the pathological process and clinical expression of Huntington's disease Brain 138:1907-1918 doi:10.1093/brain/awv107
27. Wexler NS et al. (2004) Venezuelan kindreds reveal that genetic and environmental factors modulate Huntington's disease age of onset Proc Natl Acad Sci U S A 101:3498-3503 doi:10.1073/pnas.0308679101

28. Wright GEB et al. (2019) Length of Uninterrupted CAG, Independent of Polyglutamine Size, Results in Increased Somatic Instability, Hastening Onset of Huntington Disease Am J Hum Genet 104:1116-1126 doi:10.1016/j.ajhg.2019.04.007
29. Wu YR et al. (2005) Analysis of polyglutamine-coding repeats in the TATA-binding protein in different neurodegenerative diseases J Neural Transm (Vienna) 112:539-546 doi:10.1007/s00702-004-0197-9
30. Zuhlke C et al. (2001) Different types of repeat expansion in the TATA-binding protein gene are associated with a new form of inherited ataxia Eur J Hum Genet 9:160-164 doi:10.1038/sj.ejhg.5200617

Tables

Table 1: Number of *HTT* CAG repeats and *TBP* CAG / CAA repeats in our sample (n=208 alleles/ 104 individuals)

	Nº CAG repeats *(HTT)*			Nº CAG / CAA repeats (*TBP*)	
	Molecularly Negative Individuals	Alleles of Molecularly Positive individuals		Molecularly Negative individuals normal non-expanded *HTT*	Molecularly positive individuals expanded *HTT*
		Non-expanded	Expanded		
Nº of repeats (range)	12-30	14-30	39-62	29-39	25-40
Mean	19±4	19±4	45±4	36±2	36±2
Median	17	17	44	36	36
Heterozygosity Rate	81%	100%		84%	81%
Allele Classification	94% (60 normal alleles) 6% (4 intermediate alleles)	94% (68 normal alleles) 6% (4 intermediate alleles)	99% (71 complete penetrant alleles) 1% (1 reduced penetrant allele)		

Table 2: The basic structure of TBP gene showing the polymorphic regions: I, II, III, IV and V.

Different *TBP* structures	REGIONS I	II	III	*	IV	*	V	Reference	Sample observed
1	$(CAG)_3(CAA)_3$	(CAG)x - first polymorphic region	CAACAGCAA		(CAG)x - second polymorphic region		CAACAG	GOSTOUT et al., 1993(Gostout et al. 1993)	Basic structure
2	$(CAG)_3(CAA)_3$	$(CAG)_3$	CAACAGCAA		$(CAG)_{19}$		CAACAG	NCBI database	According to GCh37.p13
3.1	$(CAG)_3(CAA)_3$	$(CAG)_3$	CAACAGCAA		$(CAG)_{16,17}$		CAACAG	KOIDE et al., 1999 (Koide et al. 1999)	Normal alleles for SCA17
3.2	$(CAG)_3(CAA)_3$	$(CAG)_3$	CAACAGCAA		$(CAG)_9$	$(CAA)_3$ $(CAG)_9$ CAA CAG CAA $(CAG)_{13}$	CAACAG	KOIDE et al., 1999(Koide et al. 1999)	Expanded alleles for SCA17
4	$(CAG)_3(CAA)_3$	$(CAG)_3$	CAACAGCAA		$(CAG)_{26}$		CAACAG	FUJIGASAKI et al., 2001 (Fujigasaki et al. 2001)	Autosomal dominant cerebelar ataxia
5	$(CAG)_3(CAA)_3$	$(CAG)_3$	CAACAGCAA		$(CAG)_{23}$		CAACAG	SILVEIRA et al., 2002 (Silveira et al. 2002)	Expanded allele for SCA 17
6	$(CAG)_3(CAA)_3$	$(CAG)_5$	CAACAGCAA		$(CAG)_{20}$		CAACAG	WU et al., 2005 (Wu et al. 2005)	Patient with Parkinson and Alzheimer
7.1	$(CAG)_3(CAA)_3$	$(CAG)_3$	CAACAGCAA		$(CAG)_{30,31}$		CAACAG	ZUHLKE et al., 2001(Zuhlke et al. 2001)	2 patients with ataxia
7.2	$(CAG)_3(CAA)_3$	$(CAG)_{45, 46, 47}$					CAACAG	ZUHLKE et al., 2001(Zuhlke et al. 2001)	3 genotypes in one family with ataxia
8	**$(CAG)_3(CAA)_3$**	**$(CAG)_{17}$**					**CAACAG**	**Our study**	**Brazilian patient with HD (normal allele)**
9	**$(CAG)_3(CAA)_3$**	**$(CAG)_9$**	**CAACAGCAA**		**$(CAG)_{18}$**		**CAACAG**	**Our study**	**Brazilian patient with HD expanded allele**
10	**$(CAG)_3(CAA)_3$**	**$(CAG)_9$**	**CAACAGCAA**	**CAGCAA**	**$(CAG)_{16}$**		**CAACAG**	**Our study**	**Brazilian patient with HD expanded allele**
11	**$(CAG)_3(CAA)_3$**	**$(CAG)_9$**	**CAACAGCAA**		**$(CAG)_{17}$**		**CAACAG**	**Our study**	**Brazilian patient with HD intermediate allele**

Legend: * alternative regions.

Table 3: Influence of *HTT* expanded CAG repeats and *TBP* CAG / CAA repeats on the HD AO (n=72 individuals).

Independent Variables	Adjusted R^2	p-value (regression model)
***HTT* expanded CAG**	0,657	<0,0001
***TBP* CAG / CAA**	Excluded variable	0,31
***HTT* expanded CAG + TBP CAG / CAA**	0,658	<0,0001

*p-value< 0,05

Figures

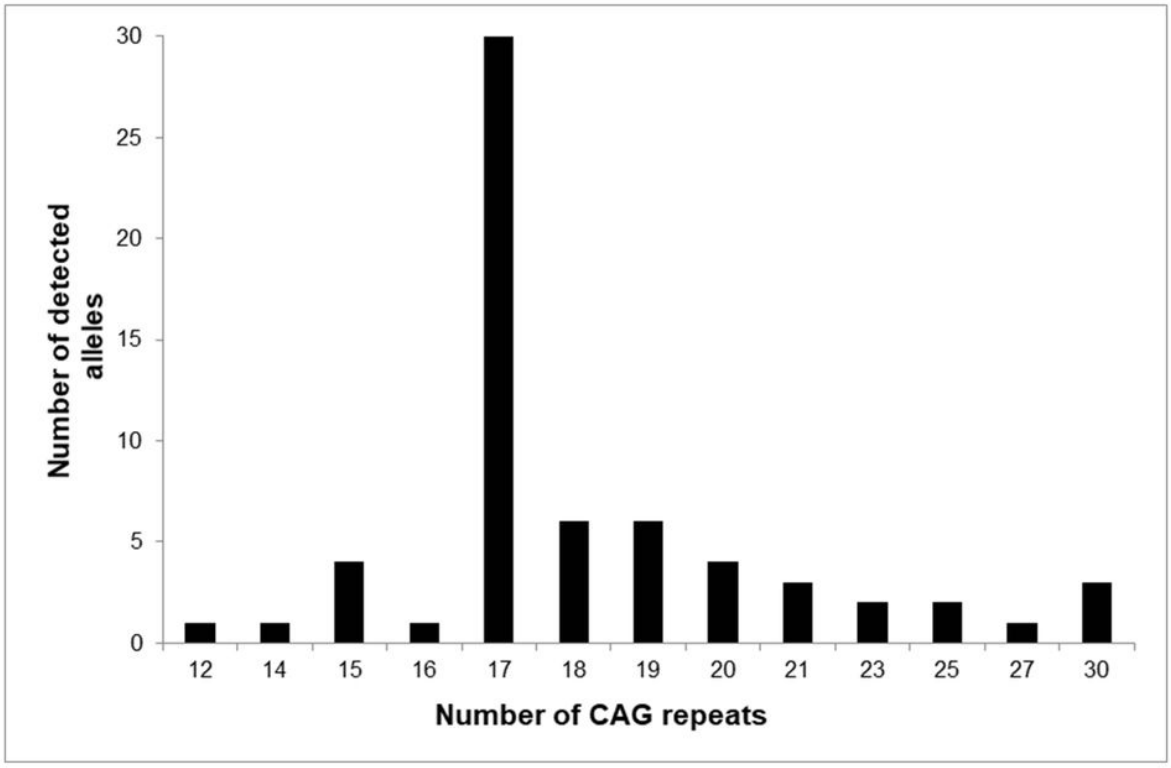

Figure 1

Frequency of CAG repeats in HTT of molecularly negative HD individuals (n = 32 individuals, 64 alleles)

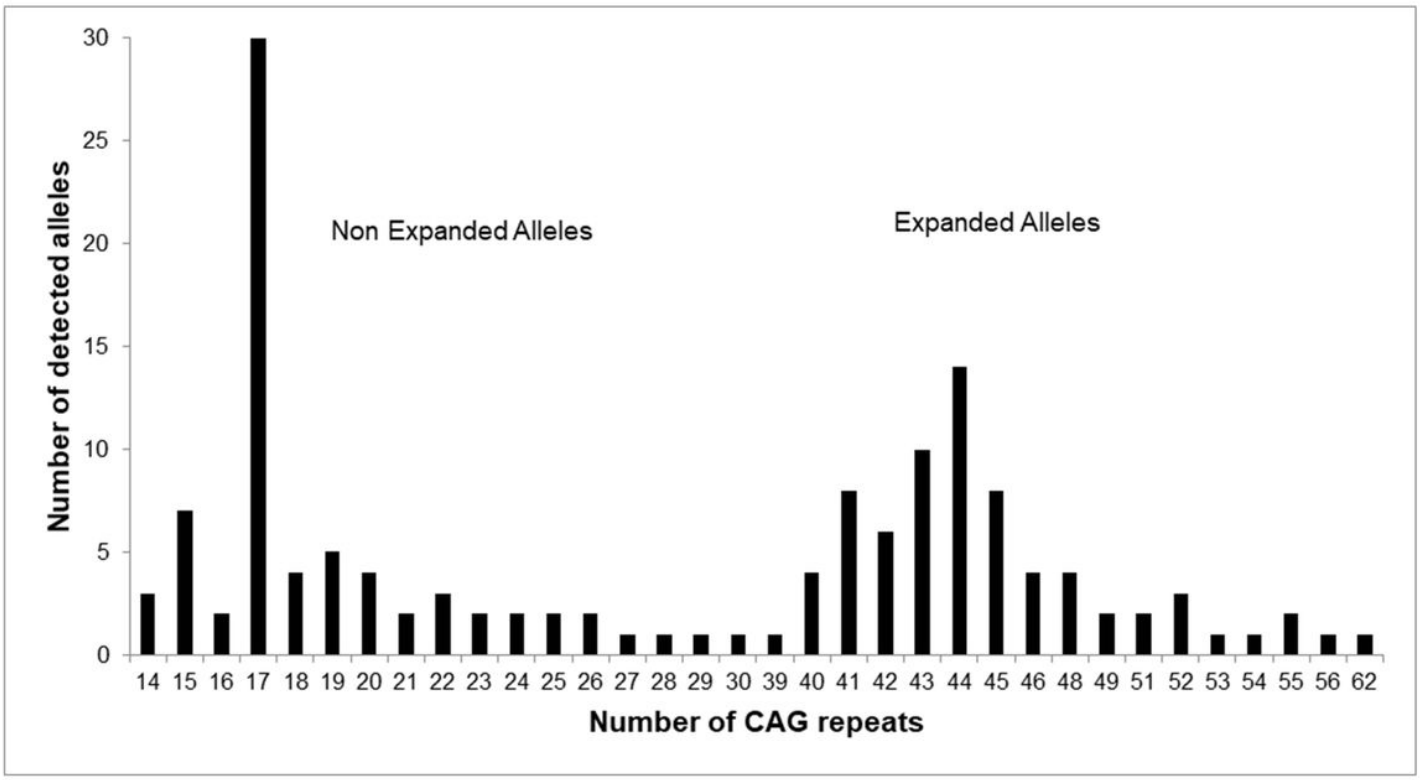

Figure 2

Frequency of CAG repeats in HTT of molecularly positive DH individuals (n = 72 individuals bearing 144 alleles: 72 non-expanded and 72 expanded alleles). Classification of alleles: 14-26 Normal, 27-35 intermediate, 36-39 incompletely penetrant; >40 completely penetrant.

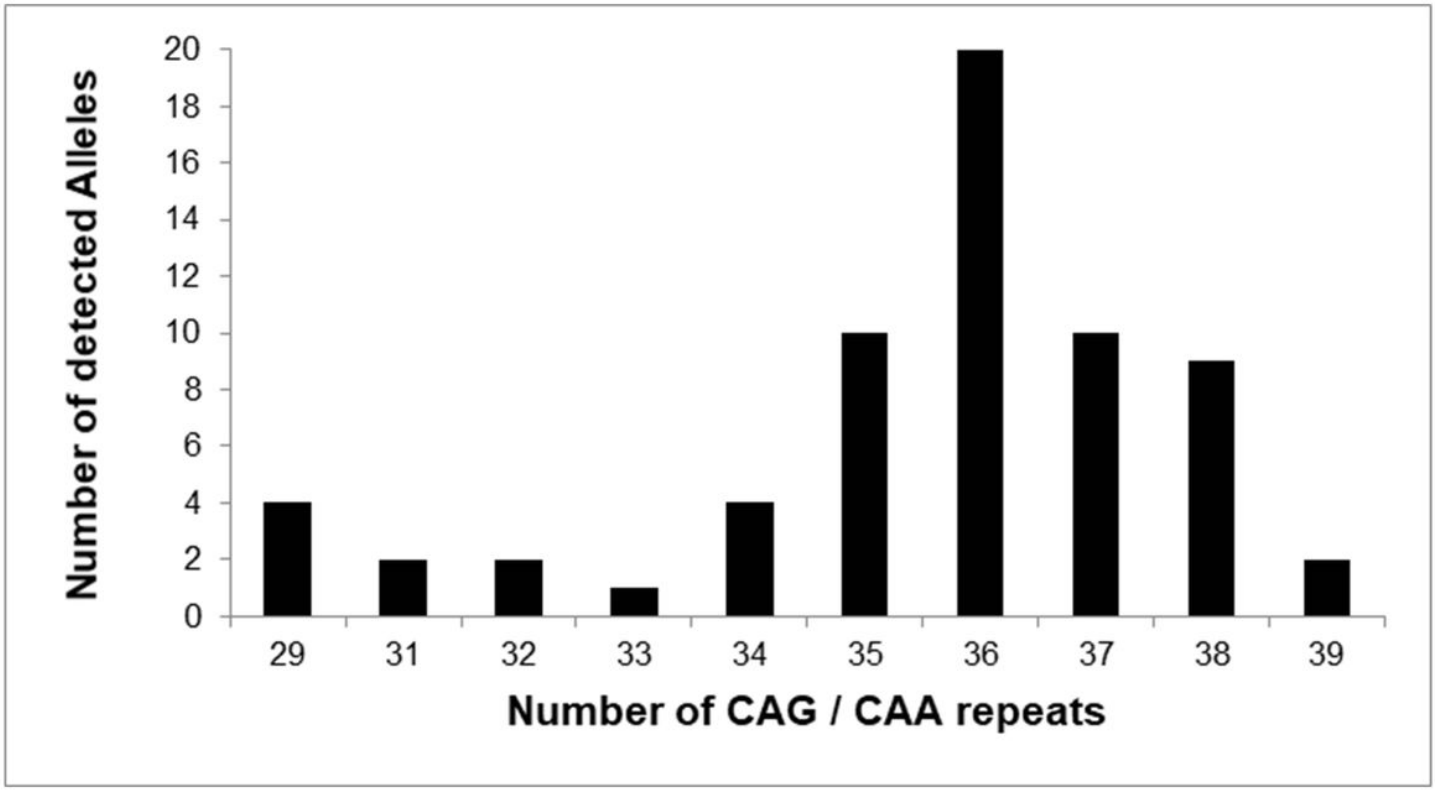

Figure 3

Frequency of CAG / CAA repeats in TBP gene of molecularly tested HD negative individuals (n = 32 individuals or 64 alleles)

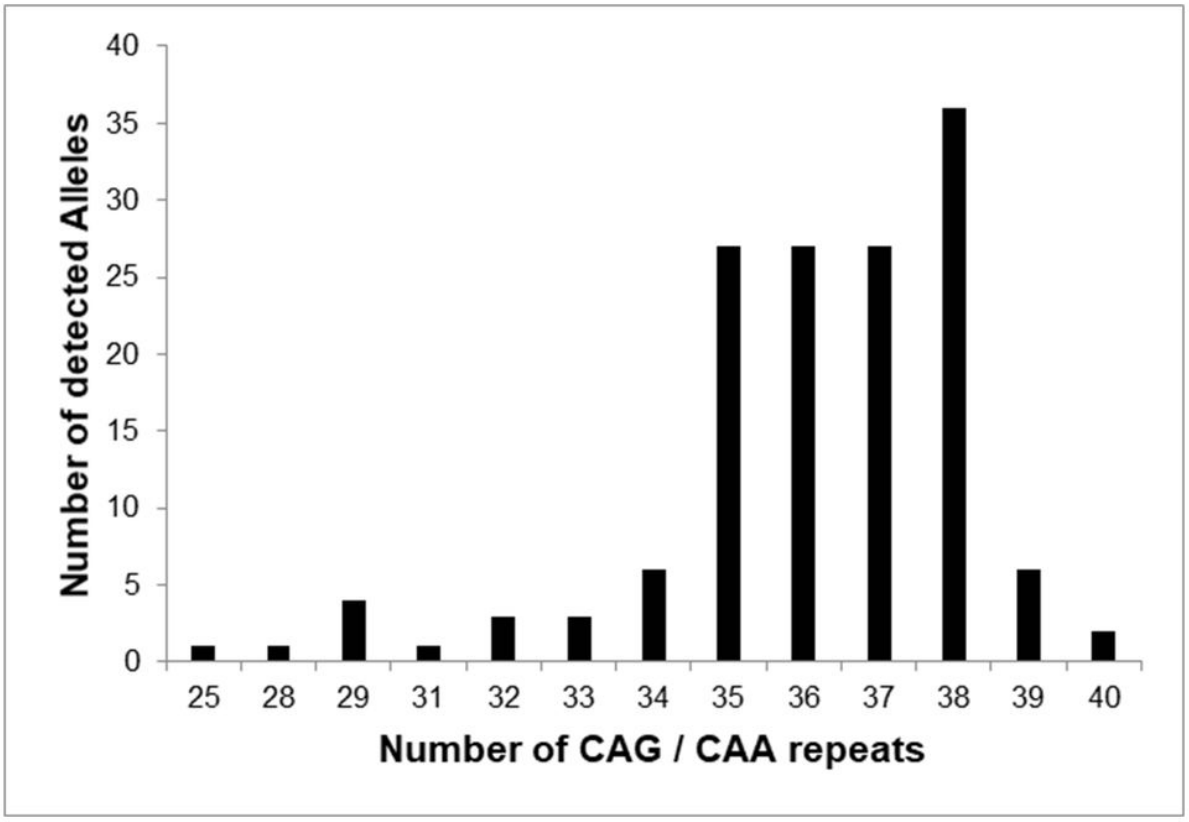

Figure 4

Frequency of CAG / CAA repeats in the TBP gene of molecularly positive HD individuals (n = 72 individuals or 144 alleles)