Neurobiology of Aging 84 (2019) 236.e17–236.e28

Contents lists available at ScienceDirect

Neurobiology of Aging

journal homepage: www.elsevier.com/locate/neuaging

Repeat variations in polyglutamine disease–associated genes and cognitive function in old age

Sarah L. Gardiner [a,b,*], Stella Trompet [c], Behnam Sabayan [d], Merel W. Boogaard [e], J. Wouter Jukema [f], P. Eline Slagboom [g], Raymund A.C. Roos [a], Jeroen van der Grond [h], N. Ahmad Aziz [i,j]

[a] Department of Neurology, Leiden University Medical Centre, Leiden, the Netherlands
[b] Department of Human Genetics, Leiden University Medical Centre, Leiden, the Netherlands
[c] Department of Internal Medicine, Section of Gerontology and Geriatrics, Leiden University Medical Centre, Leiden, the Netherlands
[d] The Ken and Ruth Davee Department of Neurology, Northwestern University, Chicago, IL, USA
[e] Department of Clinical Genetics, Leiden University Medical Centre, Leiden, the Netherlands
[f] Department of Cardiology, Leiden University Medical Centre, Leiden, the Netherlands
[g] Department of Molecular Epidemiology, Leiden University Medical Centre, Leiden, the Netherlands
[h] Department of Radiology, Leiden University Medical Centre, Leiden, the Netherlands
[i] Population Health Sciences, German Centre for Neurodegenerative Diseases (DZNE), Bonn, Germany
[j] Department of Neurology, University of Bonn, Bonn, Germany

ARTICLE INFO

Article history:
Received 30 April 2019
Received in revised form 3 August 2019
Accepted 4 August 2019
Available online 9 August 2019

Keywords:
Cognitive function
Polyglutamine diseases
Tandem repeats
Missing heritability
Polyglutamine disease–associated genes
Huntington disease

ABSTRACT

Although the heritability of cognitive function in old age is substantial, genome-wide association studies have had limited success in elucidating its genetic basis, leaving a considerable amount of "missing heritability." Aside from single nucleotide polymorphisms, genome-wide association studies are unable to assess other large sources of genetic variation, such as tandem repeat polymorphisms. Therefore, here, we studied the association of cytosine-adenine-guanine (CAG) repeat variations in polyglutamine disease–associated genes (PDAGs) with cognitive function in older adults. In a large cohort consisting of 5786 participants, we found that the CAG repeat number in 3 PDAGs (*TBP*, *HTT*, and *AR*) were significantly associated with the decline in cognitive function, which together accounted for 0.49% of the variation. Furthermore, in an magnetic resonance imaging substudy, we found that CAG repeat polymorphisms in 4 PDAGs (*ATXN2*, *CACNA1A*, *ATXN7*, and *AR*) were associated with different imaging characteristics, including brain stem, putamen, globus pallidus, thalamus, and amygdala volumes. Our findings indicate that tandem repeat polymorphisms are associated with cognitive function in older adults and highlight the importance of PDAGs in elucidating its missing heritability.

1. Introduction

Cognitive function is a key determinant of quality of life and independence in old age (Blazer et al., 2015; Cigolle et al., 2007; Pusswald et al., 2015). The proportion of the world's elderly population is rapidly increasing with the number of people aged 65 years or older estimated to rise from 8.5% of the world's population in 2015 to 16.7% by 2050 (He et al., 2016). Therefore, the necessity to understand and investigate the pathophysiology and risk factors of cognitive decline in the aging population is urgent. The contribution of genetics to brain function in older adults is substantial (Deary et al., 2009; Lee et al., 2010; Plomin et al., 2008; Singer et al., 2006). Numerous genome-wide association studies (GWASs) have identified several single nucleotide polymorphisms (SNPs) associated with age-related cognitive decline (Lin et al., 2017). The largest meta-analysis to date, investigating 57 population-based cohorts comprising a total of 300,486 participants aged between 16 and 102 years, reported 148 SNPs associated with general cognitive function at a genome-wide significance level. The maximum proportion of phenotypic variance in general cognition explained in 3 independent samples by a prediction score derived from these data ranged between 2.63% and 4.31% (Davies et al., 2018), leaving a substantial proportion of unexplained or "missing heritability." Despite their essential role in genetic

* Corresponding author at: Department of Neurology, Leiden University Medical Centre (LUMC), Albinusdreef 2 2333 ZA, Leiden, the Netherlands. Tel.: +31 71 52 69602; fax: +31 71 52 68285.
E-mail address: s.l.gardiner@lumc.nl (S.L. Gardiner).

https://doi.org/10.1016/j.neurobiolaging.2019.08.002

research, GWASs are limited because of the fact that these studies cannot assess polymorphisms in tandem repeat sequences (Lander et al., 2001). Nonetheless, 3% of genomic DNA consists of such repetitive DNA stretches and thus have a considerable impact on genetic variation (Hannan, 2010a, b, 2018).

Nine hereditary neurodegenerative diseases, known as polyglutamine disorders, including Huntington disease (HD), are the most prevalent disorders associated with tandem repeats (Mirkin, 2007; Shao and Diamond, 2007). These diseases are caused by an elongated cytosine-adenine-guanine (CAG) repeat sequence in the protein-coding region of the respective polyglutamine disease–associated gene (PDAG) (Table 1) (Benton et al., 1998; Bettencourt et al., 2016; Chung et al., 1993; David et al., 1997; Del-Favero et al., 1998; Fan et al., 2014; Giunti et al., 1995; Johansson et al., 1998; Koide et al., 1999; La Spada et al., 1991; Maciel et al., 1995; Maruyama et al., 1995; Matilla et al., 1993; Matsuyama et al., 1997; Nagafuchi et al., 1994; Nakamura et al., 2001; Orr et al., 1993; Pulst et al., 1996; Sanpei et al., 1996; Zhuchenko et al., 1997). The elongated repeat sequence is present from birth. However, symptoms usually do not arise until middle age, making age a prominent risk modifier for disease onset. Furthermore, aside from progressive motor defects and psychiatric disturbances, polyglutamine disorder symptoms often affect cognitive function. Increasing evidence implicates the polyglutamine domains of PDAGs as key regulators of transcriptional regulation, synaptic plasticity, calcium homeostasis, and mitochondrial energy production, dysregulation of which have been associated with cognitive aging (Burke and Barnes, 2006; Hands et al., 2008; Paulson et al., 2017). However, to what extent more common CAG repeat length variations in PDAGs are associated with cognitive decline in older adults and changes in relevant brain structures is still unknown. Because polyglutamine diseases are caused by expanded repeat sequences, the most rational hypothesis would be that larger repeat numbers in PDAGs are associated with a worse cognitive function. However, previous studies demonstrated more complex, nonlinear associations between CAG repeat size variations and neurodegenerative and neuropsychiatric disorders (Gardiner et al., 2019b; Gardiner et al., 2017b; Gardiner et al., 2017d; Stuitje et al., 2017). Therefore, in this study, we aimed to assess the association of cognitive function in older adults with the CAG repeat number in the 9 known PDAGs, investigating different linear and nonlinear associations in an unbiased fashion.

2. Subjects and methods

2.1. Subjects

The 9 known PDAGs (including *ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7*, *TBP*, *HTT*, *ATN1*, and *AR*) were genotyped in all participants with sufficient amounts of DNA available from blood samples of the Prospective Study of Pravastatin in the Elderly at Risk study. This well-characterized cohort included 5786 men and women aged between 70 and 83 years with a pre-existing vascular disease or a raised risk for such a disease. The study originally was a prospective multicentre randomized placebo-controlled trial to assess the effect of treatment with pravastatin on the risk of major vascular events in the elderly. Previous research showed that the treatment with pravastatin in this cohort did not affect cognitive decline or brain volume (ten Dam et al., 2005; Trompet et al., 2010). Participants were recruited from Scotland (n = 2517), Ireland (n = 2173), and the Netherlands (n = 1096). All participants were Caucasian, and none were diagnosed with a polyglutamine disease. The study was approved by the institutional ethics review boards of all centers, and informed consent was obtained from all participants (Shepherd et al., 2002).

2.2. Cognitive function assessment

The Mini–Mental State Examination was used to assess global cognitive function. To measure various cognitive domains, 4 neuropsychological tests were administered to all subjects. The Stroop-Color-Word test was applied to assess executive function by evaluating attention span. The outcome parameter for this test was the total number of seconds required to complete the third Stroop card (Stroop III). The Letter-Digit Coding Test was applied to assess processing speed with the total number of correct entries completed in 60 seconds as outcome. Memory was assessed with the 15-Picture Learning Test, which measured immediate and delayed recall. Immediate recall was defined as the accumulated number of pictures recalled over 3 learning trials, and delayed recall was defined as the number of pictures recalled after 20 minutes. Cognitive function was measured at baseline, after 9, 18, and 30 months of follow-up, and at the end of the study, which varied between participants (36–48 months) (Shepherd et al., 1999).

Table 1
Summary of genotyped PDAGs

PDAG	Disease	Protein	CAG repeat ranges			PROSPER			
			Normal	Pathological	Allele	Mean ± SD	Median	N	Range
ATXN1	SCA1	Ataxin-1	6–38	39–83	Short	29.2 ± 1.0	29	5633	17–36
					Long	30.8 ± 1.7	30	5633	26–44
ATXN2	SCA2	Ataxin-2	14–32	33–500	Short	21.9 ± 0.6	22	5548	11–27
					Long	22.4 ± 1.2	22	5548	22–27
ATXN3	SCA3	Ataxin-3	12–44	52–87	Short	19.0 ± 4.4	20	5544	14–34
					Long	24.3 ± 3.8	23	5543	14–29
CACNA1A	SCA6	CACNA1A	4–18	20–33	Short	10.6 ± 2.1	11	5633	4–14
					Long	12.5 ± 1.1	13	5633	7–17
ATXN7	SCA7	Ataxin-7	3–19	37–460	Short	10.1 ± 0.5	10	5285	7–16
					Long	10.8 ± 1.2	10	5285	10–25
TBP	SCA17	TBP	25–43	45–66	Short	36.3 ± 1.8	37	5559	26–40
					Long	37.9 ± 1.0	38	5559	30–47
HTT	HD	HD	6–26	36–121	Short	16.9 ± 2.1	17	5602	9–29
					Long	20.2 ± 3.5	19	5602	10–38
ATN1	DRPLA	Atrophin-1	3–38	48–93	Short	12.3 ± 3.1	14	5633	5–20
					Long	15.6 ± 2.2	16	5633	8–27
AR	SMBA	Androgen receptor	6–34	36–72	Short	21.2 ± 2.7	21	5546	7–35
					Long	22.8 ± 2.9	23	5546	7–39

Key: CACNA1A, calcium channel, voltage-dependent P/Q-type, α1A subunit; DRPLA, dentatorubropallidoluysian atrophy; HD, Huntington disease; PROSPER, Prospective Study of Pravastatin in the Elderly at Risk; PDAG, polyglutamine disease–associated gene; SD, standard deviation; SCA, spinocerebellar ataxia; SBMA, spinal bulbar muscular atrophy; TBP, thymine-adenine-thymine-adenine (TATA) box binding protein.

2.3. Imaging characteristics

Two successive magnetic resonance imaging (MRI) scans of the brain were obtained from a total of 646 Dutch participants with a follow-up of 30 months. MRI was performed on a system operating at 1.5-T field strength (Philips Medical Systems) and proton density-T2/dual fast-spin-echo images were collected (time to echo, 27/120 ms; time to repeat, 3000 ms; echo train length factor, 10; 48 contiguous 3-mm slices; matrix 256x256; field of view, 220) (ten Dam et al., 2005). From these images, several MRI parameters were determined by segmenting parenchyma (whole-brain) volume using semiautomated software developed by the Division of Image Processing Department of Radiology of the Leiden University Medical Center. Percentage of atrophy was defined as the difference between intracranial and parenchymal volume divided by the parenchymal volume. Increase in atrophy percentage per year was determined by dividing the difference in atrophy between the first and the second scan by 2.5 (30 months/12 months). Furthermore, baseline values of gray matter volume, white matter volume, and total brain volume were reported in mm^3 normalized to Montreal Neurological Institute space, whereas the volume of the brain stem and the subcortical structures (i.e., caudate nucleus, putamen, thalamus, globus pallidus, nucleus accumbens, amygdala, and hippocampus) were reported in unnormalized mm^3. The volumes of the right and left subcortical structures were summed to generate one overall volume per subcortical structure.

2.4. Genotyping

To determine the CAG repeat length in the 9 PDAGs for each individual, a polymerase chain reaction (PCR) was performed in a TProfessional thermocycler (Biometra, Westburg) with labeled primers flanking the CAG stretch of the PDAGs (Biolegio) (Supplemental Table 1). The PCR was performed using 10 ng of genomic DNA, 1x OneTaq master mix (New England Biolabs, OneTaq Hot start with GC Buffer master mix), 1 μL of primer Mix A or B (Supplemental Table 1) and Aqua B. Braun water to a final volume of 10 μL. The PCR was run with 27 cycles of 30 seconds, denaturation at 94 °C, one minute of annealing at 60 °C, and 2 minutes of elongation at 69 °C, preceded by 5 minutes of initial denaturation at 94 °C. Final elongation was performed at 69 °C for 5 minutes. Every PCR included a negative control without genomic DNA and a reference sample of CEPH 1347-02 genomic DNA. The PCR products were run on an ABI 3730 automatic DNA sequencer (Applied Biosystems) and analyzed using the GeneMarker software, version 2.4.0. For every analysis, we included 3 controls with known CAG repeat lengths for each PDAG to assure every run was performed reliably. All assessments were performed by randomizing study participants across batches while researchers were blinded with respect to the clinical information.

2.5. Statistical analysis

To analyze the association between the CAG repeat number in PDAGs and cognitive function in older adults, we initially performed a principal component analysis to aggregate the different cognitive test scores into one outcome. However, the first principal component of "cognition at baseline" and "cognition over time" both did not explain more than 52.7%–55.8% of the variance. Therefore, we did not find these components to summarize the actual outcomes reliably. Instead, we calculated Z-scores for all cognitive assessment test outcomes, including Mini–Mental State Examination score, total number of seconds to complete Stroop III, total number of correct entries in the Letter-Digit Coding Test, and total number of pictures recalled in the immediate and delayed recall of the 15-Picture Learning Test. The data of Stroop III had a skewed distribution and thus was log-transformed and inverted as higher scores indicated a worse performance. We summed up the Z-scores of the data derived at baseline to create a baseline summary score of cognitive function. A higher score indicated a better cognitive function. Generalized linear mixed-effect models with this summary score as a dependent variable were applied to assess the association between the CAG repeat numbers in the PDAGs and cognitive function at baseline. Country (i.e., Scotland, Ireland, or the Netherlands) was set as random effect to account for potential population stratification, and the CAG repeat numbers of both alleles were set as fixed effects. Each allele was defined as either "short" or "long" relative to the CAG repeat number in the other allele. The alleles were defined as such because previous research has shown the "shorter" and the "longer" allele to have different associations with various outcomes, such as age of onset in polyglutamine diseases, body mass index, and depression (Aziz et al., 2009; Gardiner et al., 2017a; Gardiner et al., 2017c; Gardiner et al., 2018; Stuitje et al., 2017). The CAG repeat size in the X-linked *AR* gene was examined by (1) analyzing men and women separately and (2) investigating either the shorter or the longer *AR* allele separately in men and women combined. To assess potential interaction or nonlinear effects (Gardiner et al., 2017c; Tezenas du et al., 2014), we also included a product term of both alleles and a quadratic term for each allele as fixed effects.

To assess the association between the CAG repeat number in the PDAGs and the decline in cognitive function over time, we created a summary score of the cognitive function over time by adding the Z-scores of the data derived per visit (the Z-scores were created with respect to the baseline distribution). Consequently, each individual had one cognitive summary score per visit. We then applied similar generalized linear mixed-effects models with this summary score as a dependent variable. The time of assessment, defined as the number of months of follow-up (i.e., 0, 9, 18, 30 or 36–48), and product terms of this follow-up time and the CAG repeat number in both alleles were added as fixed effects, in addition to the variables in the previous model. To model the clustering by country as well as the longitudinal repeated measurements within each individual, we applied a random intercept-and-slope model: Apart from a random intercept for country, a random slope for the time of assessment was added to account for the longitudinal interindividual variation in the rate of cognitive decline (i.e., the random slope for the time effect) and to allow treatment of missing values at each time point under the missing-at-random assumption.

MRI scans were only performed in Dutch participants, and MRI parameters were not assessed at more than one point in time. Therefore, linear regression instead of generalized linear mixed-effect modeling was applied to analyze the association between the different imaging parameters and the CAG repeat number in the PDAGs. The MRI variables were set as dependent variables, and the CAG repeat numbers of both alleles, including the product term and the quadratic terms, were included as independent variables.

To assess whether CAG repeat variation in the PDAGs could explain a significant additional amount of variation in the respective dependent variable, we first performed an omnibus test (i.e., a restricted F-test) per PDAG by including as independent variables all terms associated with the respective locus (i.e., CAG repeat size of both alleles, their interaction, a quadratic term for each allele and the product of follow-up time, and the CAG repeat number in both alleles if applicable). We applied a false discovery rate correction to account for multiple testing, assuming 9 independent tests with the false discovery rate q set at 0.05, to determine whether an omnibus test result was significant (Benjamini and Hochberg, 1995). Only in cases where an omnibus test was statistically significant after multiple testing correction, we performed post hoc tests by

Table 2
Summary of PROSPER characteristics

Variable	Observations (n)	Percentage
Sex		
Men	2798	48.4
Women	2988	51.6
Country		
The Netherlands	1096	18.9
Scotland	2517	43.5
Ireland	2173	37.6
Education		
Elementary school	131	2.3
> Elementary school	5655	97.7

Variable	Observations (n)	Mean	SD	Range
Age at baseline	5786	75.33	3.35	69.37 to 83.39
Cognitive function				
Baseline measurements				
MMSE (score)	5718	28.03	1.54	22 to 30
Stroop III (sec.)	5370	66.59	26.91	20.09 to 346.30
LDT (correct #)	5407	23.07	7.83	5 to 54
PLTi (correct #)	5444	9.32	1.95	4 to 14.33
PLTd (correct #)	5444	10.14	2.59	3 to 15
Measurements from all time points				
MMSE (score)	25,401	28.10	1.81	8 to 30
Stroop III (sec.)	23,803	65.64	27.17	10.09 to 346.30
LDT (correct #)	24,340	22.90	7.81	0 to 55
PLTi (correct #)	24,670	9.38	2.00	1 to 15
PLTd (correct #)	24,670	10.08	2.80	0 to 15
Summary scores[a]				
Baseline cognitive function	5003	0.18	3.57	−12.55 to 9.92
Decline in cognitive function	22,399	0.21	3.63	−16.30 to 10.15
MRI variables (*The Netherlands*)				
Baseline atrophy (%)	536	26.21	3.10	12.74 to 37.04
Atrophy after 30 mo (%)	542	27.23	3.43	7.61 to 38.41
Increase in atrophy per year (%)	527	−0.42	0.84	−5.34 to 4.81
Gray matter volume (mm^3)[b]	494	$5.91*10^5$	$4.45*10^4$	$4.24*10^5$ to $7.16*10^5$
White matter volume (mm^3)[b]	494	$7.68*10^5$	$3.85*10^4$	$6.46*10^5$ to $8.98*10^5$
Total brain volume (mm^3)[b]	494	$1.36*10^6$	$6.57*10^4$	$1.14*10^6$ to $1.56*10^6$
Nucleus caudatus (mm^3)	423	$7.52*10^3$	$9.89*10^2$	$5.17*10^3$ to $1.11*10^4$
Putamen (mm^3)	409	$1.05*10^4$	$1.10*10^3$	$7.68*10^3$ to $1.39*10^4$
Thalamus (mm^3)	428	$1.62*10^4$	$1.29*10^3$	$1.28*10^4$ to $1.92*10^4$
Globus pallidus (mm^3)	387	$3.87*10^3$	$7.23*10^2$	$2.21*10^3$ to $6.26*10^3$
Nucleus accumbens (mm^3)	209	$1.13*10^3$	$2.53*10^2$	$5.84*10^2$ to $1.98*10^3$
Amygdala (mm^3)	423	$4.00*10^3$	$5.80*10^2$	$2.36*10^3$ to $5.38*10^3$
Hippocampus (mm^3)	413	$9.28*10^3$	$1.06*10^3$	$6.71*10^3$ to $12.2*10^4$
Brain stem (mm^3)	429	$2.34*10^4$	$2.60*10^3$	$1.72*10^4$ to $30.0*10^4$

Key: CAG, cytosine adenine guanine; MMSE, Mini–Mental State Examination; MRI, magnetic resonance imaging; Stroop III, interference segment of the Stroop test; LDT, Letter-Digit Test; PLTi, Average Total Word Learning Test Immediate Recall; PLTd, word Learning Test Delayed Recall; SD, standard deviation.
[a] Calculated by adding the Z-scores derived from the results of the cognitive tests.
[b] Normalized to Montreal Neurological Institute (MNI) space.

assessing the effect of the individual PDAG alleles and their associated higher-order terms to gain more insight into the nature of the association. Nonsignificant higher-order terms were removed, and the analyses were repeated to arrive at a final model. All final models were corrected for age, sex, and population structure using principal components generated from genome-wide genotyping data (Shiffman et al., 2012). Although having a pre-existing vascular disease or a raised risk for such a disease probably has an effect on cognition and cognitive decline, this variable is unlikely to be associated with CAG repeat polymorphisms in PDAGs and can thus not be considered a confounder. Therefore, we did not correct for this covariate. To reduce multicollinearity, the CAG repeat numbers were centered around their respective means. To account for potential effects of heteroscedasticity and influential points, all statistical significance tests were based on robust estimators of standard errors and all CAG repeat lengths with a frequency of less than 10 for the analyses of cognitive function or less than 5 for the analyses of imaging characteristics were excluded (Supplemental Tables 2 and 3). We calculated the R^2 per PDAG for each final model to determine the amount of variance explained by each gene (Nakagawa and Schielzeth, 2013).

We determined the amount of variance in cognitive function explained by the significant repeat polymorphisms in total by incorporating the fixed effects of the final significant models with cognitive function over time as outcome into one model and calculating its marginal R^2. To illustrate this combined effect, we (1) calculated the residual cognitive function after regression on age, sex, follow-up time, and principal components generated from genome-wide genotyping data as fixed factors and country and the follow-up time as random factors in a generalized mixed-effects model, (2) performed linear regression with the CAG repeat sizes of the significant PDAGs, including all interaction and nonlinear effects, which were identified as significant in the main analyses as independent variables and the residual cognitive function as the outcome, (3) divided the total cohort into 2 equally sized groups based on the predicted values of this regression model, and (4) plotted the average residual cognitive function of these 2 groups. All data are displayed as means and 95% confidence intervals unless

Table 3
The association between summary scores of cognitive function and the CAG repeat number in PDAGs

Outcome	Variable[a]	β-coefficient[d]	SE	t	*p*-value	95% CI		R^2
Baseline cognitive function								
	AR_s ♀	$7.78*10^{-2}$	$4.51*10^{-2}$	1.72	0.085	$-1.07*10^{-2}$	$1.66*10^{-1}$	$8.95*10^{-3}$
	AR_l ♀	$1.52*10^{-2}$	$5.50*10^{-2}$	0.28	0.782	$-9.27*10^{-2}$	$1.23*10^{-1}$	
	AR_l2 ♀	$\mathbf{1.70*10^{-2}}$	$\mathbf{4.14*10^{-3}}$	**4.12**	$\mathbf{0.383*10^{-4}}$[e]	$\mathbf{8.93*10^{-3}}$	$\mathbf{2.52*10^{-2}}$	
	AR_s ♀[b]	$7.51*10^{-2}$	$3.15*10^{-2}$	2.38	0.017	$1.33*10^{-2}$	$1.37*10^{-1}$	
	AR_l ♀[b]	$-8.20*10^{-4}$	$5.45*10^{-2}$	−0.02	0.988	$-1.08*10^{-1}$	$1.06*10^{-1}$	
	AR_l2 ♀[b]	$\mathbf{2.01*10^{-2}}$	$\mathbf{4.18*10^{-3}}$	**4.82**	$\mathbf{0.144*10^{-5}}$[e]	$\mathbf{1.19*10^{-2}}$	$\mathbf{2.83*10^{-2}}$	
Cognitive function over time								
	TBP_s	$9.67*10^{-3}$	$5.03*10^{-3}$	1.93	0.054	$-1.76*10^{-4}$	$1.95*10^{-2}$	$4.78*10^{-3}$
	TBP_l	$\mathbf{-4.79*10^{-2}}$	$\mathbf{2.29*10^{-2}}$	**−2.10**	**0.036**[e]	$\mathbf{-9.27*10^{-2}}$	$\mathbf{-3.10*10^{-3}}$	
	TBP_s[c]	$2.01*10^{-2}$	$9.13*10^{-3}$	2.20	0.028	$2.22*10^{-3}$	$3.80*10^{-2}$	
	TBP_l[c]	$\mathbf{-6.08*10^{-2}}$	$\mathbf{2.29*10^{-2}}$	**−2.65**	**0.008**[e]	$\mathbf{-1.06*10^{-1}}$	$\mathbf{1.59*10^{-2}}$	
	HTT_s	$-1.90*10^{-2}$	$3.32*10^{-2}$	−0.57	0.566	$-8.40*10^{-2}$	$4.59*10^{-2}$	$3.06*10^{-3}$
	HTT_l	$9.40*10^{-3}$	$1.23*10^{-2}$	0.77	0.443	$-1.46*10^{-2}$	$3.34*10^{-2}$	
	HTT_l2	$\mathbf{-4.74*10^{-3}}$	$\mathbf{1.71*10^{-3}}$	**−2/78**	**0.005**[e]	$\mathbf{8.08*10^{-3}}$	$\mathbf{1.40*10^{-3}}$	
	HTT_s[c]	$-8.76*10^{-3}$	$3.04*10^{-2}$	−0.29	0.773	$-6.84*10^{-2}$	$5.08*10^{-2}$	
	HTT_l[c]	$1.48*10^{-2}$	$1.46*10^{-2}$	1.02	0.310	$-1.38*10^{-2}$	$4.34*10^{-2}$	
	HTT_l2[c]	$\mathbf{-5.20*10^{-3}}$	$\mathbf{1.83*10^{-3}}$	**−2.84**	**0.004**[e]	$\mathbf{-8.78*10^{-3}}$	$\mathbf{-1.62*10^{-3}}$	
	AR_s ♀	$5.87*10^{-2}$	$2.24*10^{-2}$	2.63	0.009	$1.49*10^{-2}$	$1.03*10^{-1}$	$2.37*10^{-2}$
	AR_l ♀	$1.24*10^{-2}$	$1.94*10^{-2}$	0.64	0.523	$-2.57*10^{-2}$	$5.05*10^{-2}$	
	AR_l2 ♀	$\mathbf{1.66*10^{-2}}$	$\mathbf{7.51*10^{-3}}$	**2.20**	**0.028**[e]	$\mathbf{1.83*10^{-3}}$	$\mathbf{3.13*10^{-2}}$	
	AR_s ♀[b]	$5.72*10^{-2}$	$1.84*10^{-2}$	3.12	0.002	$2.12*10^{-2}$	$9.32*10^{-2}$	
	AR_l ♀[b]	$-2.47*10^{-3}$	$1.58*10^{-2}$	−0.16	0.876	$-3.35*10^{-2}$	$2.85*10^{-2}$	
	AR_l2 ♀[b]	$\mathbf{1.91*10^{-2}}$	$\mathbf{1.77*10^{-3}}$	**10.83**	$\mathbf{0.000*10^{-7}}$[e]	$\mathbf{1.57*10^{-2}}$	$\mathbf{2.26*10^{-2}}$	
	AR_s	$2.94*10^{-3}$	$1.54*10^{-2}$	−0.19	0.851	$-3.37*10^{-2}$	$2.78*10^{-2}$	$7.86*10^{-4}$
	AR_s_m	$\mathbf{-5.48*10^{-4}}$	$\mathbf{-5.48*10^{-4}}$	**−4.02**	$\mathbf{0.571*10^{-4}}$[e]	$\mathbf{-8.14*10^{-4}}$	$\mathbf{-2.81*10^{-4}}$	
	AR_s[c]	$-3.24*10^{-3}$	$1.63*10^{-2}$	−0.20	0.842	$-3.52*10^{-2}$	$2.87*10^{-2}$	
	AR_s_m[c]	$\mathbf{-4.00*10^{-4}}$	$\mathbf{1.36*10^{-4}}$	**−2.94**	**0.003**[e]	$\mathbf{-6.66*10^{-4}}$	$\mathbf{-1.33*10^{-4}}$	

Bold indicates the variables that drive the significant association.
Key: CAG, cytosine adenine guanine; CI, confidence interval; l, relatively longer allele; l2, quadratic term relatively longer allele; _m, interaction with month; PDAGs, polyglutamine disease–associated genes; R^2, estimated variance explained by the significant association; s, relatively shorter allele; s2, quadratic term relatively shorter allele; SE, standard error.
[a] Only the variables of PDAGs are reported in the cases where the omnibus test was statistically significant after multiple testing correction.
[b] Corrected for age and population structure using principal components.
[c] Corrected for age, sex, and population structure using principal components.
[d] This column indicates the change in summary score of cognitive function per unit CAG repeat size increase.
[e] *p*-value statistically significant.

otherwise specified. All analyses were performed in STATA/SE, version 14.2 (StataCorp LLC).

3. Results

We were able to determine the CAG repeat length of between 5285 and 5633 individuals for each gene. The lacking samples were due to too little available DNA and were missing completely at random (Table 1). In total, we found 79 participants with CAG repeat numbers within the pathological range of at least one PDAG, including *ATXN1* (n = 4), *ATXN2* (n = 2), *TBP* (n = 66), and *HTT* (n = 7). For a more extensive report on these findings, please see our previous publication (Gardiner et al., 2019a). Between 0.1% and 2.9% of the cases per gene were excluded from the analyses of cognitive function because of CAG repeat numbers with a frequency of less than 10, and between 0.8% and 7.2% of the cases were excluded from the analyses of the imaging parameters because of CAG repeat numbers with a frequency of less than 5 (Supplemental Tables 2 and 3). The cohort characteristics are summarized in Table 2. The summary scores of baseline cognitive function and cognitive function over time ranged from −12.55 to 9.92 and from −16.30 to 10.15, respectively.

3.1. Baseline cognitive function was associated with CAG repeat variations in AR in women

Initial omnibus tests with cognitive function at baseline as the outcome were significant for *ATXN2* and *TBP* ($p \leq 0.008$). In addition, the omnibus tests for the shorter *AR* allele assessed in both sexes and the longer *AR* allele assessed in women only were statistically significant ($p \leq 0.004$). The post hoc tests subsequently indicated a significant nonlinear association between cognitive function at baseline and the CAG repeat number in the longer *AR* allele in women (Table 3). Both relatively small and relatively large CAG repeat numbers in the longer *AR* allele in women were associated with a higher cognitive function at baseline (Fig. 1). A CAG repeat number of about 22–23 in the longer *AR* allele was

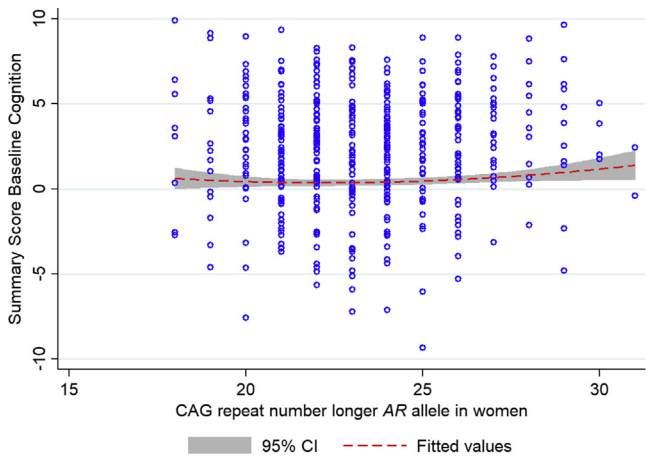

Fig. 1. The association between cognitive function at baseline and the CAG repeat number in the longer *AR* allele in women. The cognitive function assessed at baseline had a nonlinear association with the CAG repeat number in the longer *AR* allele in women. In women, both relatively small and relatively large CAG repeat numbers were associated with a higher cognitive function at baseline.

associated with the lowest cognitive function at baseline. These results did not change materially after correcting for age, sex, and population structure (Table 3). Post hoc tests for *ATXN2*, *TBP*, and the shorter *AR* allele assessed in both sexes did not indicate a significant association with cognitive function at baseline.

3.2. *The CAG repeat variations in* TBP, HTT, *and* AR *were associated with cognitive function over time*

The omnibus tests with cognitive function over time as outcome were significant for *ATXN2*, *TBP*, and *HTT* ($p < 0.005$). Furthermore, the omnibus tests were statistically significant, when the longer and the shorter *AR* allele were assessed separately in both sexes, and when both *AR* alleles were assessed in women only ($p < 0.009$). Subsequent post hoc tests indicated a significant association between cognitive function over time and CAG repeat variations in the longer *TBP* allele, the longer *HTT* allele, the longer *AR* allele assessed in only women, and the shorter *AR* allele assessed in both sexes (Table 3). Larger CAG repeat numbers in the longer *TBP* allele were associated with a lower summary score of cognitive function at each point in time that cognitive function was assessed. This association is visualized by observing that as individuals age, the summary score of cognitive function declines, but the negative association with the CAG repeat number in the longer *TBP* allele remains apparent in each age group (Fig. 2A). Similarly, in each age group, the CAG repeat number in the longer *HTT* allele and the longer *AR* allele in women had a nonlinear associations with cognitive function over time. Both relatively small and relatively large CAG repeat numbers in the longer *HTT* allele were associated with a lower cognitive function (Fig. 2B). Conversely and similar to its association with cognitive function at baseline, smaller and larger CAG repeat numbers in the longer *AR* allele assessed in women were associated with a higher cognitive function (Fig. 2C). The CAG repeat number in the shorter *AR* allele in both sexes was associated with the rate of decline in cognitive function, based on the significant association between the interaction term of the CAG repeat number in the shorter *AR* allele and follow-up time (Table 3). Larger CAG repeat numbers in the shorter *AR* allele assessed in both sexes were associated with a faster decline in cognitive function over time (Fig. 2D). These associations remained significant and largely unaltered after correction for age, sex, and population structure. Post hoc tests for *ATXN2* and the longer *AR* allele assessed in both sexes did not indicate a significant association with cognitive function over time.

In combination, CAG repeat polymorphisms in *TBP*, *HTT*, and the shorter *AR* allele assessed in both sexes explained 0.49% of the total variance in cognitive function (Fig. 3). In this calculation, we did not

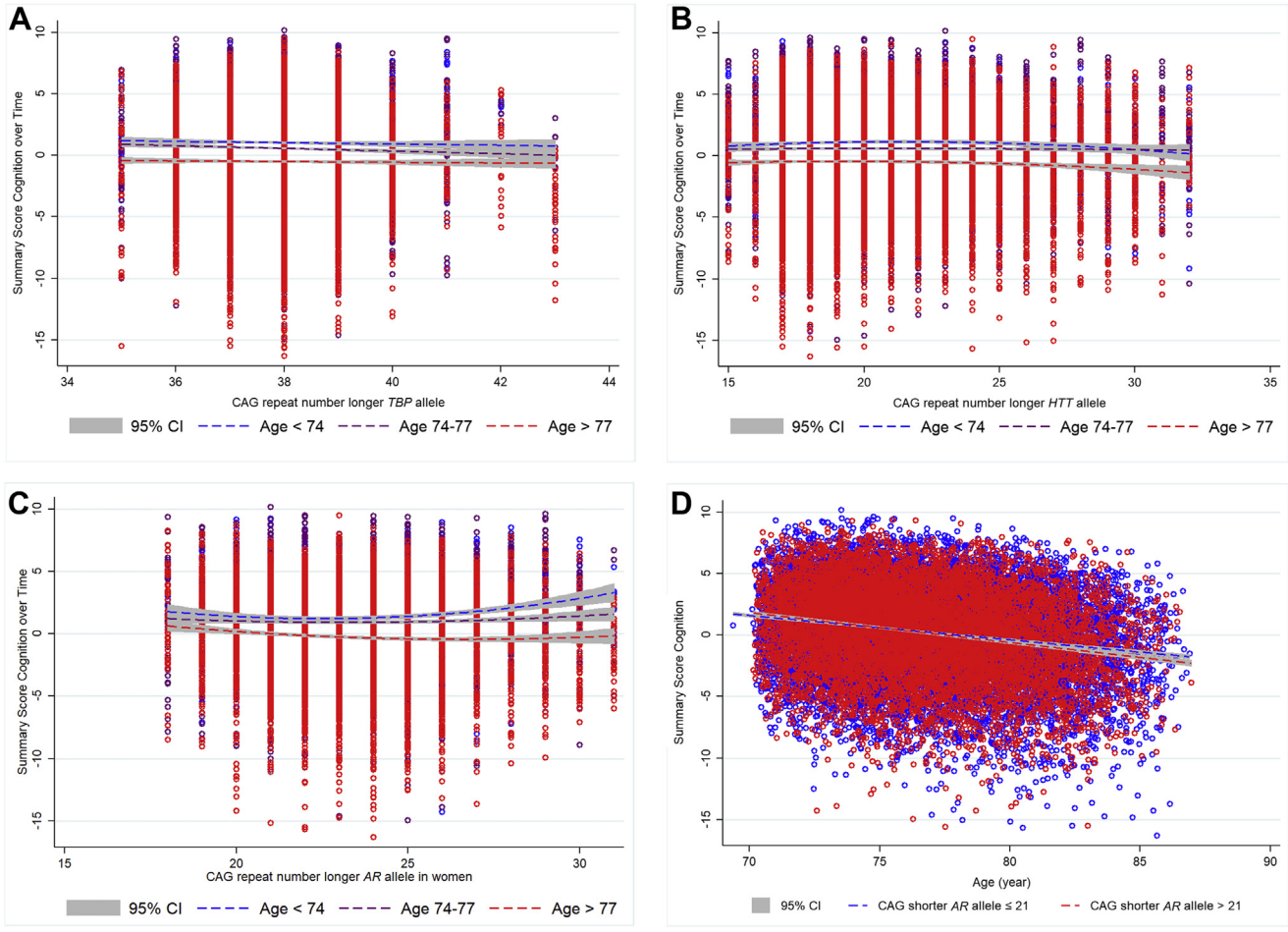

Fig. 2. The association between cognitive function over time and the CAG repeat number in *TBP*, *HTT*, and *AR*. (A) Larger CAG repeat numbers in the longer *TBP* allele were associated with a lower cognitive function. This association remained significant throughout the entire assessment period and is thus visible in all age groups. (B) Both relatively small and relatively large CAG repeat numbers in the longer *HTT* allele were associated with a decreased cognitive function. This association was also present at all assessment time points and can be observed over the 3 different age groups. (C) Relatively small and relatively large CAG repeat numbers in the longer *AR* allele assessed in women were associated with a higher cognitive function. This association was again consistent over the entire assessment period and all age groups. (D) CAG repeat numbers >21 CAG repeats were associated with a faster decline in cognitive function over time compared with CAG repeat numbers ≤21.

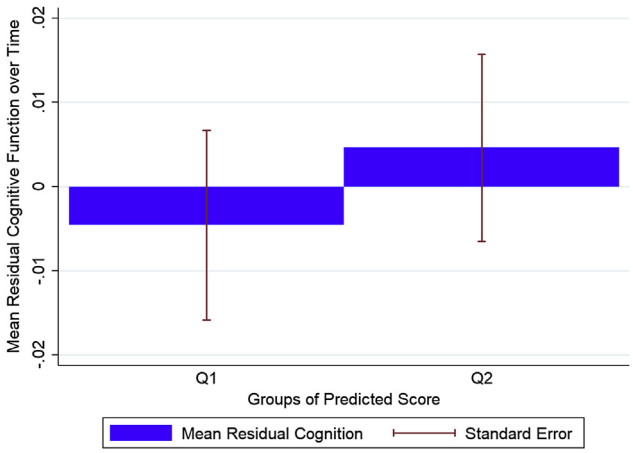

Fig. 3. The effect of CAG repeat size variations in polyglutamine disease–associated genes (PDAGs) on the summary score of cognitive function over time. This plot illustrates that the combination of CAG repeat size variations in the 3 PDAGs can account for a variation of up to 0.02 summary score points.

incorporate the effect of the CAG repeat polymorphisms in the longer *AR* allele because this effect was only significant when assessed in women.

3.3. *CAG repeat variations in* ATXN2, CACNA1A, ATXN7, *and* AR *were significantly associated with brain stem and subcortical structure volumes*

The MRI variables of which omnibus tests and subsequent post hoc tests indicated significant associations with CAG repeat polymorphisms in the PDAGs are presented in Table 4. In women, the CAG repeat number in the longer *AR* allele had a nonlinear association with the volume of the amygdala. Both smaller and larger CAG repeat numbers were associated with a higher volume of the amygdala (Supplemental Fig. 1A). The longer *AR* allele in both sexes had a similar nonlinear association with the volume of the amygdala, the putamen, and the thalamus (Supplemental Fig. 1B, C and D). In contrast, larger CAG repeat numbers in the shorter *AR* allele assessed in both sexes were associated with a smaller volume of the amygdala (Supplemental Fig. 2).

When the shorter *CACNA1A* allele had less than 11 repeats, larger CAG repeat numbers in the longer *CACNA1A* allele were associated with an increased volume of the amygdala (Supplemental Fig. 3A). CAG repeat variations in *CACNA1A* were also associated with the volume of the thalamus. Larger CAG repeat numbers in the shorter *CACNA1A* allele were associated with a larger thalamic volume (Supplemental Fig. 3B). The CAG repeat number in the longer *ATXN2* allele and the shorter *ATXN7* allele had inverse associations with the volume of the brain stem (Supplemental Fig. 4A and B). The CAG repeat number in the shorter *ATXN2* allele had nonlinear associations with the volume of the putamen and the thalamus (Supplemental Fig. 4C and D). Both smaller and larger CAG repeat numbers in the shorter *ATXN2* allele were associated with smaller volumes of both subcortical structures. The association between the CAG repeat number in *ATXN2* and the increase in percentage of atrophy per year, the volume of the amygdala and the volume of the globus pallidus, and the association between the interaction of both *ATXN7* alleles and the volume of the nucleus accumbens were mainly driven by a few influential points and, therefore, unlikely to be robust (Supplemental Fig. 5A, B, C and 6). CAG repeat variations in the other PDAGs were not significantly associated with atrophy at baseline, gray matter volume, white matter volume, total brain volume, the volume of the hippocampus, or the volume of the nucleus caudatus. All results remained materially unaltered after correcting for age, sex, and population structure (Table 4).

4. Discussion

To our knowledge, this study is the first to investigate the association between CAG repeat variations within the "normal" range of PDAGs and cognitive function in older adults. We found that CAG repeat number variations in *TBP*, *HTT*, and *AR* were significantly associated with cognitive function independent of age. In addition, we found that the CAG repeat variations in the shorter *AR* allele were associated with the rate of cognitive decline. Together, the CAG repeat polymorphisms in these genes explained about 0.49% of the variance in cognitive function. Furthermore, we found that CAG repeat polymorphisms in *ATXN2*, *CACNA1A*, *ATXN7*, and *AR* were significantly associated with the volume of the brain stem and several subcortical structures, including the amygdala, the thalamus, and the putamen. Remarkably, the associations of *AR* repeat polymorphisms with cognitive function closely matched the associations between the *AR* repeat polymorphisms and MRI variables.

The maximum proportion of phenotypic variance in general cognition explained in 3 independent samples by a prediction score derived from data of 57 population-based cohorts evaluated in a meta-analysis, ranged between 2.63% and 4.31% (Davies et al., 2018). In this study, we found that CAG repeat polymorphisms within the "normal" range in the 9 known PDAGs explained almost 0.49% of the variance in cognitive function. Although external validation of our results is warranted, this remarkable finding indicates the enormous potential of tandem repeat polymorphisms in the genetic research of cognitive function.

The polyglutamine stretch in *HTT* has been implicated previously as relevant for cognitive function. *Lee* et al. reported that in children aged 6–18 years, larger CAG repeat numbers within the "normal" range in the longer *HTT* allele were associated with a higher general intelligence (IQ) (Lee et al., 2017). In addition, larger CAG repeat numbers in *HTT* were associated with increased gray matter volume in the globus pallidus (Muhlau et al., 2012). Furthermore, organisms with a more complex central nervous system possess a higher number of repeats in *HTT* (Cattaneo et al., 2005; Tartari et al., 2008). However, CAG repeat numbers in *HTT* beyond 35 repeats cause HD, which is characterized among other things by cognitive deterioration. These facts imply a curvilinear association between cognitive function and the number of CAG repeats in *HTT* with the suggestion of an optimal number of CAG repeats. Indeed, here, we found that CAG repeat numbers in the longer *HTT* allele had a nonlinear association with cognitive function in older adults with relatively small and large CAG repeat numbers being associated with a lower cognitive function and the optimal CAG repeat number being around 22 repeats. This nonlinear association has also been suggested by other groups (Lee et al., 2018).

Comparable with *HTT*, the number of CAG repeats in the X-linked *AR* gene increases as organisms gain a more complex central nervous system throughout evolution (Choong et al., 1998; Djian et al., 1996; Hong et al., 2006). In addition, very short CAG repeat sequences in *AR* were associated with very severe mental retardation (Kooy et al., 1999). However, in older men, larger CAG repeat numbers in *AR* were found to accelerate cognitive decline (Gardiner et al., 2019b; Yaffe et al., 2003). Thus, these findings suggest a complex association between cognitive function and CAG repeat variations in *AR*, which is reflected in our findings. CAG repeat variations in the longer *AR* allele had a nonlinear association with cognitive function and the volumes of the amygdala, the putamen, and the thalamus. Both relatively small and relatively large CAG repeat numbers were associated with a higher cognitive function

Table 4
The association between MRI variables and the CAG repeat number in PDAGs

Outcome	Variable[a]	β-coefficient[d]	SE	t	*p*-value	95% CI		R^2
Increase in atrophy								
	ATXN2_s	$1.04*10^{-1}$	$5.82*10^{-2}$	1.79	0.075	$-1.04*10^{-2}$	$2.18*10^{-1}$	$5.30*10^{-3}$
	ATXN2_l	$-5.83*10^{-2}$	$5.48*10^{-2}$	−1.06	0.288	$-1.66*10^{-1}$	$4.94*10^{-2}$	
	ATXN2_sl	**$2.24*10^{-1}$**	**$1.06*10^{-1}$**	**2.12**	**0.035**[e]	**$1.62*10^{-2}$**	**$1.62*10^{-2}$**	
	ATXN2_s[b]	$8.22*10^{-2}$	$4.21*10^{-2}$	1.95	0.052	$-5.66*10^{-4}$	$1.65*10^{-1}$	
	ATXN2_l[b]	$-5.14*10^{-2}$	$5.61*10^{-2}$	−0.92	0.360	$-1.62*10^{-1}$	$5.89*10^{-2}$	
	ATXN2_sl[b]	**$1.44*10^{-1}$**	**$5.88*10^{-2}$**	**2.44**	**0.015**[e]	**$2.81*10^{-2}$**	**$2.59*10^{-1}$**	
Subcortical structures								
Thalamus								
	ATXN2_s	−1040.46	299.23	−3.48	0.001	−1628.72	−452.21	$1.92*10^{-2}$
	ATXN2_l	−25.54	61.43	−0.42	0.678	−146.31	95.23	
	ATXN2_s2	**−245.40**	**67.15**	**−3.65**	**$0.292*10^{-3}$**[e]	**−377.10**	**−113.39**	
	ATXN2_s[b]	−1074.56	363.73	−2.95	0.003	−1789.95	−359.18	
	ATXN2_l[b]	−10.65	67.34	−0.16	0.874	−143.11	121.80	
	ATXN2_s2[b]	**−246.22**	**81.53**	**−3.02**	**0.003**[e]	**−406.58**	**−85.86**	
	AR_l	−38.11	26.23	−1.45	0.147	−89.68	13.45	$1.77*10^{-2}$
	AR_l2	**20.34**	**7.64**	**2.66**	**0.008**[e]	**5.31**	**35.37**	
	AR_l[b]	−66.02	26.36	−2.50	0.013	−117.87	−14.16	
	AR_l2[b]	**21.20**	**7.66**	**2.77**	**0.006**[e]	**6.13**	**36.26**	
Nucleus accumbens								
	ATXN7_s	−1513.85	88.52	−17.10	$0.436^{e}10^{-39}$	−1688.48	−1339.21	$4.03*10^{-2}$
	ATXN7_l	21.31	17.29	1.23	0.219	−12.80	55.41	
	ATXN7_s2	599.53	21.66	27.68	$0.478^{e}10^{-67}$	556.80	642.26	
	ATXN7_sl	**320.05**	**35.37**	**9.05**	**$0.188*10^{-15}$**[e]	**250.27**	**389.83**	
	ATXN7_s[b]	−1564.11	96.44	−16.22	$0.269*10^{-34}$	−1754.60	−1373.62	
	ATXN7_l[b]	14.86	17.47	0.85	0.396	−19.64	49.37	
	ATXN7_s2[b]	617.60	38.68	15.97	$0.124*10^{-34}$	541.20	694.01	
	ATXN7_sl[b]	**326.38**	**43.21**	**7.55**	**$0.336*10^{-11}$**[e]	**241.02**	**411.73**	
Amygdala								
	ATXN2_s	−10.94	29.72	−0.37	0.713	−69.36	47.48	$2.16*10^{-2}$
	ATXN2_l	−214.06	78.92	−2.71	0.007	−369.22	−58.90	
	ATXN2_l2	45.54	21.39	2.13	0.034	3.49	87.60	
	ATXN2_sl	**34.65**	**10.18**	**3.40**	**0.001**[e]	**14.63**	**54.67**	
	ATXN2_s[a]	−33.95	27.27	−1.24	0.214	−87.58	19.69	
	ATXN2_l[b]	195.89	89.94	−2.18	0.030	−372.80	−18.99	
	ATXN2_l2[b]	39.42	23.51	1.68	0.094	−6.81	85.66	
	ATXN2_sl[b]	**46.42**	**11.08**	**4.19**	**$0.357*10^{-4}$**[e]	**24.62**	**68.22**	
	CACNA1A_s	76.64	23.93	3.20	0.001	29.59	123.69	$5.37*10^{-2}$
	CACNA1A_l	−43.38	37.83	−1.15	0.252	−117.75	31.00	
	CACNA1A_s2	16.06	6.60	2.43	0.015	3.09	29.04	
	CACNA1A_sl	**−33.74**	**9.94**	**−3.39**	**0.001**[e]	**−53.28**	**−14.20**	
	CACNA1A_s[b]	75.32	25.78	2.92	0.004[e]	24.63	126.02	
	CACNA1A_l[b]	−36.80	40.46	−0.91	0.346	−116.38	42.78	
	CACNA1A_s2[b]	17.04	6.92	2.46	0.014	3.43	30.66	
	CACNA1A_sl[b]	**−37.61**	**10.23**	**−3.68**	**$0.274*10^{-3}$**[e]	**−57.74**	**−17.48**	
	AR_s ♀	−32.12	21.60	−1.49	0.139	−74.76	10.52	$4.59*10^{-2}$
	AR_l ♀	12.00	19.57	0.61	0.541	−26.63	50.63	
	AR_l2 ♀	**17.01**	**7.03**	**2.42**	**0.017**[e]	**3.13**	**30.88**	
	AR_s ♀[c]	−41.08	24.45	−1.68	0.095	−89.39	7.23	
	AR_l ♀[c]	16.92	21.56	0.78	0.434	−25.69	59.52	
	AR_l2 ♀[c]	**22.37**	**7.05**	**3.17**	**0.002**[e]	**8.44**	**36.30**	
	AR_s	**−31.62**	**11.06**	**−2.86**	**0.004**[e]	**−53.37**	**−9.88**	$2.00*10^{-2}$
	AR_s[b]	**−33.25**	**12.21**	**−2.72**	**0.007**[e]	**−57.27**	**−9.24**	
	AR_l	−26.32	10.95	−2.40	0.017	−47.85	−4.79	$2.07*10^{-2}$
	AR_l2	**6.15**	**3.01**	**2.05**	**0.042**[e]	**2.38*10−1**	**12.06**	
	AR_l[b]	−29.67	12.16	−2.44	0.015	−53.60	−5.75	
	AR_l2[b]	**7.38**	**3.24**	**2.28**	**0.023**[e]	**1.00**	**13.75**	
Brain stem								
	ATXN2_s	21.61	114.39	0.19	0.850	−203.26	246.49	$1.06*10^{-2}$
	ATXN2_l	**−283.18**	**106.49**	**−2.66**	**0.008**[e]	**−492.52**	**−73.83**	
	ATXN2_s[b]	67.65	122.75	0.55	0.582	−173.76	309.07	
	ATXN2_l[b]	**−277.09**	**123.48**	**−2.24**	**0.025**[e]	**−519.95**	**−34.23**	
	ATXN7_s	**−748.72**	**361.24**	**−2.07**	**0.039**[e]	**−1458.91**	**−38.52**	**$9.20*10^{-3}$**
	ATXN7_l	8.23	126.19	0.07	0.948	−239.86	256.33	
	ATXN7_s[b]	**−1078.27**	**283.32**	**−3.81**	**$0.167*10^{-3}$**[e]	**−1635.55**	**−512.00**	
	ATXN7_l[b]	97.66	119.33	0.82	0.414	−137.05	332.38	

Bold indicates the variables that drive the significant association.
Key: CAG, cytosine adenine guanine; CI, confidence interval; l, relatively longer allele; l2, quadratic term relatively longer allele; MRI, magnetic resonance imaging; s, relatively shorter allele; PDAGs, polyglutamine disease–associated genes; R^2, estimated variance explained by the significant association; s2, quadratic term relatively shorter allele; sl, interaction term relatively shorter and longer allele; SE, standard error.
[a] Only the variables of PDAGs are reported in the cases where the omnibus test was statistically significant after multiple testing correction.
[b] Corrected for age, sex, and population structure using principal components.
[c] Corrected for age and population structure using principal components.
[d] This column indicates the change in volume of the MRI variable per unit CAG repeat size increase.
[e] *p*-value statistically significant.

and larger volumes of the respective subcortical structures. The first part of this nonlinear association is reflected in the association between the CAG repeat number in the shorter *AR* allele and the rate of cognitive decline and the volume of the amygdala. Relatively short CAG repeat numbers in the shorter *AR* allele were associated with a slower decline in cognitive function and a larger volume of the amygdala. As the CAG repeat number in the shorter *AR* allele increased, the rate of cognitive decline increased and the volume of the amygdala decreased, similar to the first part of the nonlinear association between the longer *AR* allele and cognitive function. This complex association between cognitive function and CAG repeat variation over the entire "normal" range in *AR* is intriguing and warrants more detailed investigation in future experiments.

CAG repeat numbers in *TBP* exceeding 44 repeats are associated with the development of spinocerebellar ataxia type 17 (SCA17), an autosomal dominant inherited neurodegenerative disease characterized by cerebellar ataxia, involuntary movements, psychiatric symptoms, and cognitive decline eventually resulting in dementia (Zuhlke and Burk, 2007). Previously, we demonstrated that larger CAG repeat numbers in *TBP* within the "normal" range were associated with a higher risk of lifetime depression (Gardiner et al., 2017b). This finding supports the notion that the cutoff of disease causing CAG repeats in SCA17 might not be as rigorous as previously suggested and that depressed individuals with a CAG repeat number within this large "normal" CAG repeat range could have a less severe form of SCA17 characterized only by psychiatric symptoms. Here, we found that larger CAG repeat numbers in the longer *TBP* allele were associated with a decreased cognitive function. Similarly, this finding suggests that the pathology caused by larger CAG repeat numbers in *TBP* operates in a more continuous and gradual fashion rather than on a dichotomous scale.

The aging brain is characterized by several cellular and molecular changes, many of which overlap with pathways affected in polyglutamine disorders (Burke and Barnes, 2006; Hands et al., 2008). For instance, processes dysfunctional in HD, such as transcriptional abnormalities, dysregulation of the chaperone network, alterations of cellular protein degradation systems, mitochondrial deficits, unbalanced redox-homeostasis, and changes in axonal transport and synaptic function have all been linked to aging (Hands et al., 2008). Aside from the polyglutamine disease spinocerebellar ataxia type 2, polymorphisms in *ATXN2* have also been associated with other neurodegenerative diseases, such as amyotrophic lateral sclerosis and progressive supranuclear palsy (Pulst et al., 1996; Ross et al., 2011; Sproviero et al., 2017). In addition, in a large GWAS, the *ATXN2* locus was associated with longevity (Fortney et al., 2015). Therefore, we can hypothesize that polymorphisms in *ATXN2* increase the susceptibility for certain neurodegenerative diseases and neurodegeneration in general through similar cellular pathways and perhaps also resulting in alterations in different cerebral structures, such as the volume of the brain stem, the putamen, and the thalamus. Yeast strains lacking Sgf73 are exceptionally long lived. Sgf73 is a yeast orthologue of ataxin-7 (ATXN7), the protein encoded by *ATXN7* (Helmlinger et al., 2006; McCormick et al., 2014). ATXN7 is a member of the TBP (TATA-binding protein)-free TAF (TBP-associated factor) complex (TFTC) and the SPT3/TAF9/GCN5 acetyltransferase (STAGA) complex, the co-activators required for the transcription of RNA polymerase II-dependent genes (Helmlinger et al., 2006). We found that larger CAG repeat sequences in *ATXN7* were associated with a smaller volume of the brain stem. This finding suggests that polyglutamine elongations might very well affect ATXN7 function within TFTC and STAGA and consequently affect the volume of cerebral structures and consequentially longevity. Several studies have also shown that aged neurons suffer from an increased Ca^{2+} conduction (Foster and Norris, 1997; Toescu et al., 2004). *CACNA1A* encodes the α1A subunit of P/Q-type voltage-dependent calcium channel (Zhuchenko et al., 1997). This subunit contains the pore forming structure of the calcium channel and is responsible for channel gating, permeability, and voltage-dependent activation and inactivation (Catterall, 2000). Although the role of the polyglutamine domain in the function of the channel continues to be debated, the associations we found between the CAG repeat number in *CACNA1A* and the volume of the brain stem in the older adults indicate a potential modifying role. Collectively, these findings illustrate the close similarities between characteristics of the aging brain and the pathophysiology of polyglutamine diseases. Combined with the associations we found between the CAG repeat number in *ATXN2*, *CACNA1A*, *ATXN7*, *TBP*, *HTT*, and *AR*, and cognitive function, the volume of the brain stem as well as several other subcortical structures in older adults, these findings implicate an important role for PDAGs in the normal aging process of the human brain.

In total, we found 79 individuals with a CAG repeat number in the pathological range of at least 1 PDAG. None of the included participants had been diagnosed with a polyglutamine disease. Unfortunately, however, follow-up data on these participants were not available. Therefore, we were neither able to determine whether these participants would have developed a polyglutamine disease at a later time nor distinguish between the cognitive decline as an early sign of a late-onset polyglutamine disease or part of "normal" aging. However, the fact that we found the association between cognitive function in old age and CAG repeat polymorphisms in different PDAGs to extend well into the normal range suggests that independent of polyglutamine disease development, CAG repeat polymorphisms are associated with cognitive function in old age.

A limitation of our work is that we conducted our study in a large cohort containing participants from 3 European countries, a relatively heterogeneous population. To adjust for this, we corrected for potential population stratification. Results from heterogeneous cohorts can be generalized more easily. Nonetheless, to increase the robustness of our findings, additional research should be conducted in other study populations. Another limitation involves the fact that MRI data were only available for a subset of our cohort and the available data consisted solely of structural measurements. Unfortunately, we did not have diffusion tensor imaging data or functional measurements at our disposal, which could potentially have led to more insight into the biological foundation of the significant associations we found between CAG repeat polymorphisms and cognitive function.

In conclusion, we found that CAG repeat number variations in *TBP*, *HTT*, and *AR* were significantly associated with cognitive function in older adults, jointly explaining nearly 0.49% of the variance. Furthermore, CAG repeat variations in *ATXN2*, *CACNA1A*, *ATXN7*, and *AR* were associated with the volume of the brain stem and several subcortical structures. Our results demonstrate the importance of tandem repeat polymorphisms as novel, but hitherto underappreciated, modifiers of cognitive aging and emphasize the role of PDAGs in healthy brain function.

Disclosure

Dr. R.A.C. Roos reported being an advisor for UniQure. No other disclosures were reported.

Acknowledgements

The authors would like to express our gratitude to all individuals who participated in the PROSPER study as well as all research staff for collecting the data.

This study was supported by a VENI-grant (#91615080) from the Netherlands Organization of Scientific Research, a Marie Sklodowska-Curie Individual Fellowship grant from the European Union (Horizon, 2020, #701130; NAA) and an investigator-initiated grant obtained from Bristol-Myers Squibb.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neurobiolaging.2019.08.002.

References

Aziz, N.A., Jurgens, C.K., Landwehrmeyer, G.B., van Roon-Mom, W.M., van Ommen, G.J., Stijnen, T., Roos, R.A., 2009. Normal and mutant HTT interact to affect clinical severity and progression in Huntington disease. Neurology 73, 1280–1285.

Benjamini, Y., Hochberg, Y., 1995. Controling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B (Methodological) 57, 289–300.

Benton, C.S., de Silva, R., Rutledge, S.L., Bohlega, S., Ashizawa, T., Zoghbi, H.Y., 1998. Molecular and clinical studies in SCA-7 define a broad clinical spectrum and the infantile phenotype. Neurology 51, 1081–1086.

Bettencourt, C., Hensman-Moss, D., Flower, M., Wiethoff, S., Brice, A., Goizet, C., Stevanin, G., Koutsis, G., Karadima, G., Panas, M., Yescas-Gomez, P., Garcia-Velazquez, L.E., Alonso-Vilatela, M.E., Lima, M., Raposo, M., Traynor, B., Sweeney, M., Wood, N., Giunti, P., Network, S., Durr, A., Holmans, P., Houlden, H., Tabrizi, S.J., Jones, L., 2016. DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases. Ann. Neurol. 79, 983–990.

Blazer, D.G., Yaffe, K., Karlawish, J., 2015. Cognitive aging: a report from the Institute of Medicine. Jama 313, 2121–2122.

Burke, S.N., Barnes, C.A., 2006. Neural plasticity in the ageing brain. Nat. Rev. Neurosci. 7, 30–40.

Cattaneo, E., Zuccato, C., Tartari, M., 2005. Normal huntingtin function: an alternative approach to Huntington's disease. Nat. Rev. Neurosci. 6, 919–930.

Catterall, W.A., 2000. Structure and regulation of voltage-gated Ca2+ channels. Annu. Rev. Cell Dev Biol 16, 521–555.

Choong, C.S., Kemppainen, J.A., Wilson, E.M., 1998. Evolution of the primate androgen receptor: a structural basis for disease. J. Mol. Evol. 47, 334–342.

Chung, M.Y., Ranum, L.P., Duvick, L.A., Servadio, A., Zoghbi, H.Y., Orr, H.T., 1993. Evidence for a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type I. Nat. Genet. 5, 254–258.

Cigolle, C.T., Langa, K.M., Kabeto, M.U., Tian, Z., Blaum, C.S., 2007. Geriatric conditions and disability: the health and retirement study. Ann. Intern. Med. 147, 156–164.

David, G., Abbas, N., Stevanin, G., Durr, A., Yvert, G., Cancel, G., Weber, C., Imbert, G., Saudou, F., Antoniou, E., Drabkin, H., Gemmill, R., Giunti, P., Benomar, A., Wood, N., Ruberg, M., Agid, Y., Mandel, J.L., Brice, A., 1997. Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. Nat. Genet. 17, 65–70.

Davies, G., Lam, M., Harris, S.E., Trampush, J.W., Luciano, M., Hill, W.D., Hagenaars, S.P., Ritchie, S.J., Marioni, R.E., Fawns-Ritchie, C., Liewald, D.C.M., Okely, J.A., Ahola-Olli, A.V., Barnes, C.L.K., Bertram, L., Bis, J.C., Burdick, K.E., Christoforou, A., DeRosse, P., Djurovic, S., Espeseth, T., Giakoumaki, S., Giddaluru, S., Gustavson, D.E., Hayward, C., Hofer, E., Ikram, M.A., Karlsson, R., Knowles, E., Lahti, J., Leber, M., Li, S., Mather, K.A., Melle, I., Morris, D., Oldmeadow, C., Palviainen, T., Payton, A., Pazoki, R., Petrovic, K., Reynolds, C.A., Sargurupremraj, M., Scholz, M., Smith, J.A., Smith, A.V., Terzikhan, N., Thalamuthu, A., Trompet, S., van der Lee, S.J., Ware, E.B., Windham, B.G., Wright, M.J., Yang, J., Yu, J., Ames, D., Amin, N., Amouyel, P., Andreassen, O.A., Armstrong, N.J., Assareh, A.A., Attia, J.R., Attix, D., Avramopoulos, D., Bennett, D.A., Bohmer, A.C., Boyle, P.A., Brodaty, H., Campbell, H., Cannon, T.D., Cirulli, E.T., Congdon, E., Conley, E.D., Corley, J., Cox, S.R., Dale, A.M., Dehghan, A., Dick, D., Dickinson, D., Eriksson, J.G., Evangelou, E., Faul, J.D., Ford, I., Freimer, N.A., Gao, H., Giegling, I., Gillespie, N.A., Gordon, S.D., Gottesman, R.F., Griswold, M.E., Gudnason, V., Harris, T.B., Hartmann, A.M., Hatzimanolis, A., Heiss, G., Holliday, E.G., Joshi, P.K., Kahonen, M., Kardia, S.L.R., Karlsson, I., Kleineidam, L., Knopman, D.S., Kochan, N.A., Konte, B., Kwok, J.B., Le Hellard, S., Lee, T., Lehtimaki, T., Li, S.C., Liu, T., Koini, M., London, E., Longstreth Jr., W.T., Lopez, O.L., Loukola, A., Luck, T., Lundervold, A.J., Lundquist, A., Lyytikainen, L.P., Martin, N.G., Montgomery, G.W., Murray, A.D., Need, A.C., Noordam, R., Nyberg, L., Ollier, W., Papenberg, G., Pattie, A., Polasek, O., Poldrack, R.A., Psaty, B.M., Reppermund, S., Riedel-Heller, S.G., Rose, R.J., Rotter, J.I., Roussos, P., Rovio, S.P., Saba, Y., Sabb, F.W., Sachdev, P.S., Satizabal, C.L., Schmid, M., Scott, R.J., Scult, M.A., Simino, J., Slagboom, P.E., Smyrnis, N., Soumare, A., Stefanis, N.C., Stott, D.J., Straub, R.E., Sundet, K., Taylor, A.M., Taylor, K.D., Tzoulaki, I., Tzourio, C., Uitterlinden, A., Vitart, V., Voineskos, A.N., Kaprio, J., Wagner, M., Wagner, H., Weinhold, L., Wen, K.H., Widen, E., Yang, Q., Zhao, W., Adams, H.H.H., Arking, D.E., Bilder, R.M., Bitsios, P., Boerwinkle, E., Chiba-Falek, O., Corvin, A., De Jager, P.L., Debette, S., Donohoe, G., Elliott, P., Fitzpatrick, A.L., Gill, M., Glahn, D.C., Hagg, S., Hansell, N.K., Hariri, A.R., Ikram, M.K., Jukema, J.W., Vuoksimaa, E., Keller, M.C., Kremen, W.S., Launer, L., Lindenberger, U., Palotie, A., Pedersen, N.L., Pendleton, N., Porteous, D.J., Raikkonen, K., Raitakari, O.T., Ramirez, A., Reinvang, I., Rudan, I., Dan, R., Schmidt, R., Schmidt, H., Schofield, P.W., Schofield, P.R., Starr, J.M., Steen, V.M., Trollor, J.N., Turner, S.T., Van Duijn, C.M., Villringer, A., Weinberger, D.R., Weir, D.R., Wilson, J.F., Malhotra, A., McIntosh, A.M., Gale, C.R., Seshadri, S., Mosley Jr., T.H., Bressler, J., Lencz, T., Deary, I.J., 2018. Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. Nat. Commun. 9, 2098.

Deary, I.J., Johnson, W., Houlihan, L.M., 2009. Genetic foundations of human intelligence. Hum. Genet. 126, 215–232.

Del-Favero, J., Krols, L., Michalik, A., Theuns, J., Lofgren, A., Goossens, D., Wehnert, A., Van den Bossche, D., Van Zand, K., Backhovens, H., van Regenmorter, N., Martin, J.J., Van Broeckhoven, C., 1998. Molecular genetic analysis of autosomal dominant cerebellar ataxia with retinal degeneration (ADCA type II) caused by CAG triplet repeat expansion. Hum. Mol. Genet. 7, 177–186.

Djian, P., Hancock, J.M., Chana, H.S., 1996. Codon repeats in genes associated with human diseases: fewer repeats in the genes of nonhuman primates and nucleotide substitutions concentrated at the sites of reiteration. Proc. Natl. Acad. Sci. U S A. 93, 417–421.

Fan, H.C., Ho, L.I., Chi, C.S., Chen, S.J., Peng, G.S., Chan, T.M., Lin, S.Z., Harn, H.J., 2014. Polyglutamine (PolyQ) diseases: genetics to treatments. Cell Transpl. 23, 441–458.

Fortney, K., Dobriban, E., Garagnani, P., Pirazzini, C., Monti, D., Mari, D., Atzmon, G., Barzilai, N., Franceschi, C., Owen, A.B., Kim, S.K., 2015. Genome-wide scan informed by age-related disease identifies loci for exceptional human longevity. PLoS Genet. 11, e1005728.

Foster, T.C., Norris, C.M., 1997. Age-associated changes in Ca(2+)-dependent processes: relation to hippocampal synaptic plasticity. Hippocampus 7, 602–612.

Gardiner, S.L., van Belzen, M.J., Boogaard, M.W., van Roon-Mom, W.M.C., Rozing, M.P., van Hemert, A.M., Smit, J.H., Beekman, A.T.F., van Grootheest, G., Schoevers, R.A., Oude Voshaar, R.C., Roos, R.A.C., Comijs, H.C., Penninx, B.W.J.H., van der Mast, R.C., Aziz, N.A., 2017a. Huntington gene repeat size variations affect risk of lifetime depression. Translational Psychiatry 7, 1277.

Gardiner, S.L., van Belzen, M.J., Boogaard, M.W., van Roon-Mom, W.M.C., Rozing, M.P., van Hemert, A.M., Smit, J.H., Beekman, A.T.F., van Grootheest, G., Schoevers, R.A., Oude Voshaar, R.C., Comijs, H.C., Penninx, B., van der Mast, R.C., Roos, R.A.C., Aziz, N.A., 2017b. Large normal-range TBP and ATXN7 CAG repeat lengths are associated with increased lifetime risk of depression. Transl Psychiatry 7, e1143.

Gardiner, S.L., van Belzen, M.J., Boogaard, M.W., van Roon-Mom, W.M.C., Rozing, M.P., van Hemert, A.M., Smit, J.H., Beekman, A.T.F., van Grootheest, G., Schoevers, R.A., Oude Voshaar, R.C., Comijs, H.C., Penninx, B.W.J.H., van der Mast, R.C., Roos, R.A.C., Aziz, N.A., 2017c. Large normal-range TBP and ATXN7 CAG repeat lengths are associated with increased lifetime risk of depression. Transl Psychiatry 7, e1143.

Gardiner, S.L., van Belzen, M.J., Boogaard, M.W., van Roon-Mom, W.M.C., Rozing, M.P., van Hemert, A.M., Smit, J.H., Beekman, A.T.F., van Grootheest, G., Schoevers, R.A., Oude Voshaar, R.C., Roos, R.A.C., Comijs, H.C., Penninx, B., van der Mast, R.C., Aziz, N.A., 2017d. Huntington gene repeat size variations affect risk of lifetime depression. Transl Psychiatry 7, 1277.

Gardiner, S.L., de Mutsert, R., Trompet, S., Boogaard, M.W., van Dijk, K.W., Jukema, P.J.W., Slagboom, P.E., Roos, R.A.C., Pijl, H., Rosendaal, F.R., Aziz, N.A., 2018. Repeat length variations in polyglutamine disease-associated genes affect body mass index. Int. J. Obes. (Lond) 43, 440–449.

Gardiner, S.L., Boogaard, M.W., Trompet, S., de Mutsert, R., Rosendaal, F.R., Gussekloo, J., Jukema, J.W., Roos, R.A.C., Aziz, N.A., 2019a. Prevalence of carriers of intermediate and pathological polyglutamine disease-associated alleles among large population-based cohorts. JAMA Neurol. 76, 650–656.

Gardiner, S.L., Harder, A.V.E., Campman, Y.J.M., Trompet, S., Gussekloo, J., van Belzen, M.J., Boogaard, M.W., Roos, R.A.C., Jansen, I.E., Pijnenburg, Y.A.L., Scheltens, P., van der Flier, W.M., Aziz, N.A., 2019b. Repeat length variations in ATXN1 and AR modify disease expression in Alzheimer's disease. Neurobiol. Aging 73, 230.e239-230.e217.

Giunti, P., Sweeney, M.G., Harding, A.E., 1995. Detection of the Machado-Joseph disease/spinocerebellar ataxia three trinucleotide repeat expansion in families with autosomal dominant motor disorders, including the drew family of Walworth. Brain 118 (Pt 5), 1077–1085.

Hands, S., Sinadinos, C., Wyttenbach, A., 2008. Polyglutamine gene function and dysfunction in the ageing brain. Biochim. Biophys. Acta 1779, 507–521.

Hannan, A.J., 2010a. Tandem repeat polymorphisms: modulators of disease susceptibility and candidates for 'missing heritability. Trends Genet. 26, 59–65.

Hannan, A.J., 2010b. TRPing up the genome: tandem repeat polymorphisms as dynamic sources of genetic variability in health and disease. Discov. Med. 10, 314–321.

Hannan, A.J., 2018. Tandem repeats mediating genetic plasticity in health and disease. Nat. Rev. Genet. 19, 286–298.

He, W., Goodkind, D., Kowal, P., 2016. An Aging World: 2015. U.S. Government Publishing Office, Washington, DC.

Helmlinger, D., Hardy, S., Eberlin, A., Devys, D., Tora, L., 2006. Both normal and polyglutamine- expanded ataxin-7 are components of TFTC-type GCN5 histone acetyltransferase- containing complexes. Biochem. Soc. Symp. 73, 155–163.

Hong, K.W., Hibino, E., Takenaka, O., Hayasaka, I., Murayama, Y., Ito, S., Inoue-Murayama, M., 2006. Comparison of androgen receptor CAG and GGN repeat length polymorphism in humans and apes. Primates 47, 248–254.

Johansson, J., Forsgren, L., Sandgren, O., Brice, A., Holmgren, G., Holmberg, M., 1998. Expanded CAG repeats in Swedish spinocerebellar ataxia type 7 (SCA7) patients: effect of CAG repeat length on the clinical manifestation. Hum. Mol. Genet. 7, 171–176.

Koide, R., Kobayashi, S., Shimohata, T., Ikeuchi, T., Maruyama, M., Saito, M., Yamada, M., Takahashi, H., Tsuji, S., 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.

Kooy, R.F., Reyniers, E., Storm, K., Vits, L., van Velzen, D., de Ruiter, P.E., Brinkmann, A.O., de Paepe, A., Willems, P.J., 1999. CAG repeat contraction in the androgen receptor gene in three brothers with mental retardation. Am. J. Med. Genet. 85, 209–213.

La Spada, A.R., Wilson, E.M., Lubahn, D.B., Harding, A.E., Fischbeck, K.H., 1991. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. Nature 352, 77–79.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J.P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, Y., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Marra, M.A., Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chissoe, S.L., Wendl, M.C., Delehaunty, K.D., Miner, T.L., Delehaunty, A., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R.A., Muzny, D.M., Scherer, S.E., Bouck, J.B., Sodergren, E.J., Worley, K.C., Rives, C.M., Gorrell, J.H., Metzker, M.L., Naylor, S.L., Kucherlapati, R.S., Nelson, D.L., Weinstock, G.M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D.R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H.M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R.W., Federspiel, N.A., Abola, A.P., Proctor, M.J., Myers, R.M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D.R., Olson, M.V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G.A., Athanasiou, M., Schultz, R., Roe, B.A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W.R., de la Bastide, M., Dedhia, N., Blocker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J.A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D.G., Burge, C.B., Cerutti, L., Chen, H.C., Church, D., Clamp, M., Copley, R.R., Doerks, T., Eddy, S.R., Eichler, E.E., Furey, T.S., Galagan, J., Gilbert, J.G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L.S., Jones, T.A., Kasif, S., Kaspryzk, A., Kennedy, S., Kent, W.J., Kitts, P., Koonin, E.V., Korf, I., Kulp, D., Lancet, D., Lowe, T.M., McLysaght, A., Mikkelsen, T., Moran, J.V., Mulder, N., Pollara, V.J., Ponting, C.P., Schuler, G., Schultz, J., Slater, G., Smit, A.F., Stupka, E., Szustakowki, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y.I., Wolfe, K.H., Yang, S.P., Yeh, R.F., Collins, F., Guyer, M.S., Peterson, J., Felsenfeld, A., Wetterstrand, K.A., Patrinos, A., Morgan, M.J., de Jong, P., Catanese, J.J., Osoegawa, K., Shizuya, H., Choi, S., Chen, Y.J., Szustakowki, J., 2001. Initial sequencing and analysis of the human genome. Nature 409, 860–921.

Lee, T., Henry, J.D., Trollor, J.N., Sachdev, P.S., 2010. Genetic influences on cognitive functions in the elderly: a selective review of twin studies. Brain Res. Rev. 64, 1–13.

Lee, J.K., Ding, Y., Conrad, A.L., Cattaneo, E., Epping, E., Mathews, K., Gonzalez-Alegre, P., Cahill, L., Magnotta, V., Schlaggar, B.L., Perlmutter, J.S., Kim, R.E., Dawson, J.D., Nopoulos, P., 2017. Sex-specific effects of the Huntington gene on normal neurodevelopment. J. Neurosci. Res. 95, 398–408.

Lee, J.K., Conrad, A., Epping, E., Mathews, K., Magnotta, V., Dawson, J.D., Nopoulos, P., 2018. Effect of trinucleotide repeats in the Huntington's gene on intelligence. EBioMedicine 31, 47–53.

Lin, C.H., Lin, E., Lane, H.Y., 2017. Genetic biomarkers on age-related cognitive decline. Front Psychiatry 8, 247.

Maciel, P., Gaspar, C., DeStefano, A.L., Silveira, I., Coutinho, P., Radvany, J., Dawson, D.M., Sudarsky, L., Guimaraes, J., Loureiro, J.E., Nezarati, M.M., Corwin, L.I., Lopes-Cendes, I., Rooke, K., Rosenberg, R., MacLeod, P., Farrer, L.A., Sequeiros, J., Rouleau, G.A., 1995. Correlation between CAG repeat length and clinical features in Machado-Joseph disease. Am. J. Hum. Genet. 57, 54–61.

Maruyama, H., Nakamura, S., Matsuyama, Z., Sakai, T., Doyu, M., Sobue, G., Seto, M., Tsujihata, M., Oh-i, T., Nishio, T., et al., 1995. Molecular features of the CAG repeats and clinical manifestation of Machado-Joseph disease. Hum. Mol. Genet. 4, 807–812.

Matilla, T., Volpini, V., Genis, D., Rosell, J., Corral, J., Davalos, A., Molins, A., Estivill, X., 1993. Presymptomatic analysis of spinocerebellar ataxia type 1 (SCA1) via the expansion of the SCA1 CAG-repeat in a large pedigree displaying anticipation and parental male bias. Hum. Mol. Genet. 2, 2123–2128.

Matsuyama, Z., Kawakami, H., Maruyama, H., Izumi, Y., Komure, O., Udaka, F., Kameyama, M., Nishio, T., Kuroda, Y., Nishimura, M., Nakamura, S., 1997. Molecular features of the CAG repeats of spinocerebellar ataxia 6 (SCA6). Hum. Mol. Genet. 6, 1283–1287.

McCormick, M.A., Mason, A.G., Guyenet, S.J., Dang, W., Garza, R.M., Ting, M.K., Moller, R.M., Berger, S.L., Kaeberlein, M., Pillus, L., La Spada, A.R., Kennedy, B.K., 2014. The SAGA histone deubiquitinase module controls yeast replicative lifespan via Sir2 interaction. Cell Rep 8, 477–486.

Mirkin, S.M., 2007. Expandable DNA repeats and human disease. Nature 447, 932–940.

Muhlau, M., Winkelmann, J., Rujescu, D., Giegling, I., Koutsouleris, N., Gaser, C., Arsic, M., Weindl, A., Reiser, M., Meisenzahl, E.M., 2012. Variation within the Huntington's disease gene influences normal brain structure. PLoS One 7, e29809.

Nagafuchi, S., Yanagisawa, H., Sato, K., Shirayama, T., Ohsaki, E., Bundo, M., Takeda, T., Tadokoro, K., Kondo, I., Murayama, N., et al., 1994. Dentatorubral and pallidoluysian atrophy expansion of an unstable CAG trinucleotide on chromosome 12p. Nat. Genet. 6, 14–18.

Nakagawa, S., Schielzeth, H., 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods Ecol. Evol. 4, 9.

Nakamura, K., Jeong, S.Y., Uchihara, T., Anno, M., Nagashima, K., Nagashima, T., Ikeda, S., Tsuji, S., Kanazawa, I., 2001. SCA17, a novel autosomal dominant cerebellar ataxia caused by an expanded polyglutamine in TATA-binding protein. Hum. Mol. Genet. 10, 1441–1448.

Orr, H.T., Chung, M.Y., Banfi, S., Kwiatkowski Jr., T.J., Servadio, A., Beaudet, A.L., McCall, A.E., Duvick, L.A., Ranum, L.P., Zoghbi, H.Y., 1993. Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. Nat. Genet. 4, 221–226.

Paulson, H.L., Shakkottai, V.G., Clark, H.B., Orr, H.T., 2017. Polyglutamine spinocerebellar ataxias - from genes to potential treatments. Nat. Rev. Neurosci. 18, 613–626.

Plomin, R., DeFries, J.C., McClearn, G.E., McGuffin, P., 2008. Behavioral Genetics, fifth ed. Worth Publishers, New York.

Pulst, S.M., Nechiporuk, A., Nechiporuk, T., Gispert, S., Chen, X.N., Lopes-Cendes, I., Pearlman, S., Starkman, S., Orozco-Diaz, G., Lunkes, A., DeJong, P., Rouleau, G.A., Auburger, G., Korenberg, J.R., Figueroa, C., Sahba, S., 1996. Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. Nat. Genet. 14, 269–276.

Pusswald, G., Tropper, E., Kryspin-Exner, I., Moser, D., Klug, S., Auff, E., Dal-Bianco, P., Lehrner, J., 2015. Health-related quality of life in patients with subjective cognitive decline and mild cognitive impairment and its relation to activities of daily living. J. Alzheimers Dis. 47, 479–486.

Ross, O.A., Rutherford, N.J., Baker, M., Soto-Ortolaza, A.I., Carrasquillo, M.M., DeJesus-Hernandez, M., Adamson, J., Li, M., Volkening, K., Finger, E., Seeley, W.W., Hatanpaa, K.J., Lomen-Hoerth, C., Kertesz, A., Bigio, E.H., Lippa, C., Woodruff, B.K., Knopman, D.S., White 3rd, C.L., Van Gerpen, J.A., Meschia, J.F., Mackenzie, I.R., Boylan, K., Boeve, B.F., Miller, B.L., Strong, M.J., Uitti, R.J., Younkin, S.G., Graff-Radford, N.R., Petersen, R.C., Wszolek, Z.K., Dickson, D.W., Rademakers, R., 2011. Ataxin-2 repeat-length variation and neurodegeneration. Hum. Mol. Genet. 20, 3207–3212.

Sanpei, K., Takano, H., Igarashi, S., Sato, T., Oyake, M., Sasaki, H., Wakisaka, A., Tashiro, K., Ishida, Y., Ikeuchi, T., Koide, R., Saito, M., Sato, A., Tanaka, T., Hanyu, S., Takiyama, Y., Nishizawa, M., Shimizu, N., Nomura, Y., Segawa, M., Iwabuchi, K., Eguchi, I., Tanaka, H., Takahashi, H., Tsuji, S., 1996. Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and cloning technique, DIRECT. Nat. Genet. 14, 277–284.

Shao, J., Diamond, M.I., 2007. Polyglutamine diseases: emerging concepts in pathogenesis and therapy. Hum. Mol. Genet. 16 (Spec No. 2), R115–R123.

Shepherd, J., Blauw, G.J., Murphy, M.B., Cobbe, S.M., Bollen, E.L., Buckley, B.M., Ford, I., Jukema, J.W., Hyland, M., Gaw, A., Lagaay, A.M., Perry, I.J., Macfarlane, P.W., Meinders, A.E., Sweeney, B.J., Packard, C.J., Westendorp, R.G., Twomey, C., Stott, D.J., 1999. The design of a prospective study of pravastatin in the elderly at risk (PROSPER). PROSPER study group. PROspective study of pravastatin in the elderly at risk. Am. J. Cardiol. 84, 1192–1197.

Shepherd, J., Blauw, G.J., Murphy, M.B., Bollen, E.L., Buckley, B.M., Cobbe, S.M., Ford, I., Gaw, A., Hyland, M., Jukema, J.W., Kamper, A.M., Macfarlane, P.W., Meinders, A.E., Norrie, J., Packard, C.J., Perry, I.J., Stott, D.J., Sweeney, B.J., Twomey, C., Westendorp, R.G., 2002. Pravastatin in elderly individuals at risk of vascular disease (PROSPER): a randomised controlled trial. Lancet 360, 1623–1630.

Shiffman, D., Trompet, S., Louie, J.Z., Rowland, C.M., Catanese, J.J., Iakoubova, O.A., Kirchgessner, T.G., Westendorp, R.G., de Craen, A.J., Slagboom, P.E., Buckley, B.M., Stott, D.J., Sattar, N., Devlin, J.J., Packard, C.J., Ford, I., Sacks, F.M., Jukema, J.W., 2012. Genome-wide study of gene variants associated with differential cardiovascular event reduction by pravastatin therapy. PLoS One 7, e38240.

Singer, J.J., MacGregor, A.J., Cherkas, L.F., Spector, T.D., 2006. Genetic influences on cognitive function using the cambridge neuropsychological test automated battery. Intelligence 34, 421–428.

Sproviero, W., Shatunov, A., Stahl, D., Shoai, M., van Rheenen, W., Jones, A.R., Al-Sarraj, S., Andersen, P.M., Bonini, N.M., Conforti, F.L., Van Damme, P., Daoud, H., Del Mar Amador, M., Fogh, I., Forzan, M., Gaastra, B., Gellera, C., Gitler, A.D., Hardy, J., Fratta, P., La Bella, V., Le Ber, I., Van Langenhove, T., Lattante, S., Lee, Y.C., Malaspina, A., Meininger, V., Millecamps, S., Orrell, R., Rademakers, R., Robberecht, W., Rouleau, G., Ross, O.A., Salachas, F., Sidle, K., Smith, B.N.,

Soong, B.W., Soraru, G., Stevanin, G., Kabashi, E., Troakes, C., van Broeckhoven, C., Veldink, J.H., van den Berg, L.H., Shaw, C.E., Powell, J.F., Al-Chalabi, A., 2017. ATXN2 trinucleotide repeat length correlates with risk of ALS. Neurobiol. Aging 51, 178.e171–178.e179.

Stuitje, G., van Belzen, M.J., Gardiner, S.L., van Roon-Mom, W.M.C., Boogaard, M.W., Tabrizi, S.J., Roos, R.A.C., Aziz, N.A., 2017. Age of onset in Huntington's disease is influenced by CAG repeat variations in other polyglutamine disease-associated genes. Brain 140, e42.

Tartari, M., Gissi, C., Lo Sardo, V., Zuccato, C., Picardi, E., Pesole, G., Cattaneo, E., 2008. Phylogenetic comparison of Huntington homologues reveals the appearance of a primitive polyQ in sea urchin. Mol. Biol. Evol. 25, 330–338.

ten Dam, V.H., Box, F.M., de Craen, A.J., van den Heuvel, D.M., Bollen, E.L., Murray, H.M., van Buchem, M.A., Westendorp, R.G., Blauw, G.J., 2005. Lack of effect of pravastatin on cerebral blood flow or parenchymal volume loss in elderly at risk for vascular disease. Stroke 36, 1633–1636.

Tezenas du, M.S., Durr, A., Bauer, P., Figueroa, K.P., Ichikawa, Y., Brussino, A., Forlani, S., Rakowicz, M., Schols, L., Mariotti, C., van de Warrenburg, B.P., Orsi, L., Giunti, P., Filla, A., Szymanski, S., Klockgether, T., Berciano, J., Pandolfo, M., Boesch, S., Melegh, B., Timmann, D., Mandich, P., Camuzat, A., Goto, J., Ashizawa, T., Cazeneuve, C., Tsuji, S., Pulst, S.M., Brusco, A., Riess, O., Brice, A., Stevanin, G., 2014. Modulation of the age at onset in spinocerebellar ataxia by CAG tracts in various genes. Brain 137 (Pt 9), 2444–2455.

Toescu, E.C., Verkhratsky, A., Landfield, P.W., 2004. Ca2+ regulation and gene expression in normal brain aging. Trends Neurosci. 27, 614–620.

Trompet, S., van Vliet, P., de Craen, A.J., Jolles, J., Buckley, B.M., Murphy, M.B., Ford, I., Macfarlane, P.W., Sattar, N., Packard, C.J., Stott, D.J., Shepherd, J., Bollen, E.L., Blauw, G.J., Jukema, J.W., Westendorp, R.G., 2010. Pravastatin and cognitive function in the elderly. Results of the PROSPER study. J. Neurol. 257, 85–90.

Yaffe, K., Edwards, E.R., Lui, L.Y., Zmuda, J.M., Ferrell, R.E., Cauley, J.A., 2003. Androgen receptor CAG repeat polymorphism is associated with cognitive function in older men. Biol. Psychiatry 54, 943–946.

Zhuchenko, O., Bailey, J., Bonnen, P., Ashizawa, T., Stockton, D.W., Amos, C., Dobyns, W.B., Subramony, S.H., Zoghbi, H.Y., Lee, C.C., 1997. Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the alpha 1A-voltage-dependent calcium channel. Nat. Genet. 15, 62–69.

Zuhlke, C., Burk, K., 2007. Spinocerebellar ataxia type 17 is caused by mutations in the TATA-box binding protein. Cerebellum (London, England) 6, 300–307.