Journal of Assisted Reproduction and Genetics
https://doi.org/10.1007/s10815-021-02106-3

REPRODUCTIVE PHYSIOLOGY AND DISEASE

CAG repeat instability in embryonic stem cells and derivative spermatogenic cells of transgenic Huntington's disease monkey

Sujittra Khampang[1,2] · Rangsun Parnpai[2] · Wiriya Mahikul[3] · Charles A. Easley IV[1,4,5] · In Ki Cho[1,6] · Anthony W.S. Chan[1,6]

Received: 7 December 2020 / Accepted: 8 February 2021

Abstract

Purpose The expansion of CAG (glutamine; Q) trinucleotide repeats (TNRs) predominantly occurs through male lineage in Huntington's disease (HD). As a result, offspring will have larger CAG repeats compared to their fathers, which causes an earlier onset of the disease called genetic anticipation. This study aims to develop a novel *in vitro* model to replicate CAG repeat instability in early spermatogenesis and demonstrate the biological process of genetic anticipation by using the HD stem cell model for the first time.

Methods HD rhesus monkey embryonic stem cells (rESCs) were cultured *in vitro* for an extended period. Male rESCs were used to derive spermatogenic cells *in vitro* with a 10-day differentiation. The assessment of CAG repeat instability was performed by GeneScan and curve fit analysis.

Results Spermatogenic cells derived from rESCs exhibit progressive expansion of CAG repeats with high daily expansion rates compared to the extended culture of rESCs. The expansion of CAG repeats is cell type–specific and size-dependent.

Conclusions Here, we report a novel stem cell model that replicates genome instability and CAG repeat expansion in *in vitro* derived HD monkey spermatogenic cells. The *in vitro* spermatogenic cell model opens a new opportunity for studying TNR instability and the underlying mechanism of genetic anticipation, not only in HD but also in other TNR diseases.

Keywords Huntington's disease · Spermatogenic cells · CAG repeat instability · HD rhesus monkey embryonic stem cells (rESCs) · Genetic anticipation

Introduction

Huntington's disease (HD) is an inherited trinucleotide repeat (TNR) disease-causing progressive motor and cognitive-behavioral impairment [1–3]. Even with significant progress in ameliorating HD symptoms or slowing down disease progression by the lowering of mutant huntingtin (HTT) protein levels using small hairpin RNA, microRNA, an antisense oligonucleotide (ASO), and gene editing [4–9], there is no cure available currently. HD progresses throughout life, until death 15–20 years after clinical onset [1–3]. The expansion of the CAG repeats in the first exon of the *HTT* gene is the primary cause of HD [1, 3, 10, 11]. In humans, the normal *HTT* allele has less than 26 CAG repeats, and the allele with full penetrance has more than 40 CAG repeats [1, 3, 10, 11]. The size of CAG repeats in the brain is inversely correlated with the age of onset and age of death in HD patients [1, 3, 10]. CAG repeat expansion has also been reported in different cell types, including lymphocytes, pluripotent stem cells (PSCs), fibroblasts, and gametes [12–22].

✉ In Ki Cho
inkicho@outlook.com

✉ Anthony W.S. Chan
awschan@hotmail.com

1 Division of Neuropharmacology and Neurologic Diseases, Yerkes National Primate Research Center, Atlanta, GA, USA

2 Embryo Technology and Stem Cell Research Center, School of Biotechnology, Suranaree University of Technology, Nakhon Ratchasima, Thailand

3 Faculty of Medicine and Public Health, HRH Princess Chulabhorn College of Medical Science, Chulabhorn Royal Academy, Bangkok, Thailand

4 Department of Environmental Health Science, College of Public Health, University of Georgia, Athens, GA, USA

5 Regenerative Bioscience Center, University of Georgia, Athens, GA, USA

6 Department of Human Genetics, Emory University, Atlanta, GA 30322, USA

Published online: 20 February 2021

Genetic anticipation was first described in fragile X syndrome (FXS) [23, 24]. It has also been reported in HD, spinocerebellar ataxia (SCA-1), and other TNR-related diseases [25–27]. The pathological size of the TNR in parents further expands in the germ cells and results in expanded TNR with earlier onset and increased disease severity in offspring, which is known as anticipation [28–32]. In HD, genetic anticipation predominantly occurs through the sperm rather than oocytes [17, 27, 33–38]. Although the underlying mechanism of paternal expansion is largely unknown, spermatogenesis is a continuous cycle of mitosis and meiosis with concurrent biological processes including DNA replication, replicative-repair, post-meiotic repair, and chromatin remodeling when the opportunities for errors are high [17, 20–22, 28, 39–42]. In HD, spermatogenesis prone to CAG expansion and expanded CAG repeats are transmitted to the progeny upon fertilization [17, 20, 22, 28, 38].

The paternal expansion of CAG repeats is a conserved and well-documented biological process in HD mouse models [17, 20, 37, 43] and human patients [22, 44, 45]. Also, we recently reported a progressive increase in CAG repeats in HD monkey sperm [13]. Unlike rodents with CAG repeat expansion primarily occurring in the post-meiotic stage [20, 21], in humans, CAG expansion occurs in both pre-meiotic cells (i.e., spermatogonia and spermatocytes) and post-meiotic spermatids [22]. Such discrepancy between humans and rodents may be due to differences in germ cell development and life span [39, 40, 46, 47]. Because of the limited availability and ethical usage of human tissues, as well as the developmental differences between rodent and human spermatogenesis, the progress in paternal germline instability and genetic anticipation research has been very limited.

Like HD in humans, CAG repeat expansion rates depend on the size of CAG repeats in HD monkeys [13]. We have reported a progressive CAG repeat expansion in HD monkey sperm at higher rates than in lymphocytes. We also observed that alleles with over 62 CAG repeats (i.e., the threshold of repeat instability) are more susceptible to expansion [13]. Because HD monkeys share similar CAG repeat behavior with HD human patients [12], here we utilized HD monkey ESCs (rESCs) to develop an *in vitro* model to replicate CAG repeat instability in spermatogenic cells for the first time. We examined CAG repeat instability of rESCs and their derivative spermatogenic cells of an HD monkey (see Fig. 1 for experimental flowchart). Similar to HD monkey sperm, *in vitro* derived spermatogenic cells have larger repeats and higher expansion rates compared to the long-term culture of rESCs. Also, consistent with prior reports in HD humans and HD monkeys, mutant *HTT* alleles with larger CAG repeats were more unstable. Here, we report a novel stem cell model that recapitulates CAG repeat behavior in *in vitro* derived spermatogenic cells. The *in vitro* spermatogenic cell model provides a unique opportunity for investigating the underlying mechanism of TNR instability and paternal genetic anticipation in TNR diseases, as well as being used in discovering new therapeutics or small molecules to slow down the expansion in spermatogenic cells

Materials and methods

HD monkey

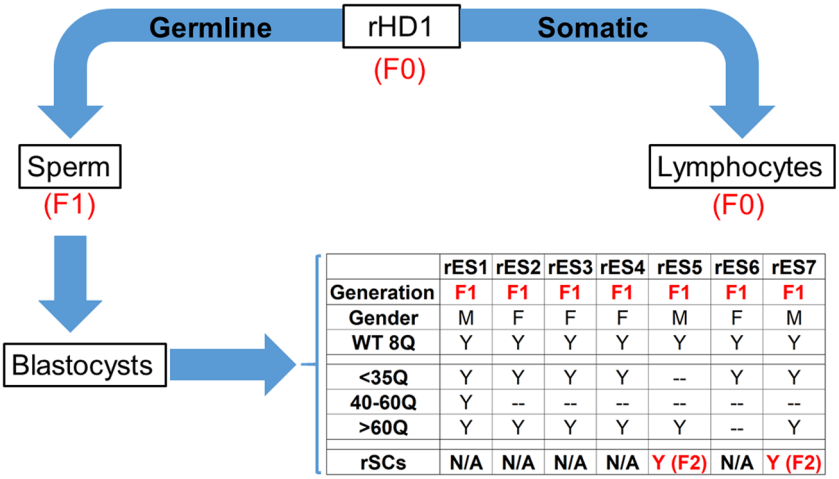
Male HD monkey, rHD1, was generated by transfecting mature rhesus macaque oocyte using a lentiviral vector expressing exon 1 of human *HTT* gene with 84 CAG repeats and green fluorescent protein (GFP) under the regulation of human polyubiquitin C (UBC) promoter. rHD1 was born with 29 CAG repeats based on assessable tissues [49]. Peripheral blood cells at 3 and 44 months and sperm samples at 44 months of age were collected and used in this study.

Derivation and culture of rESs1–7

Seven HD monkey embryonic stem cell lines (rESs1–7) and a wild-type embryonic stem cell line (rES-WT) were established from blastocysts derived from oocytes fertilized by using rHD1 sperm and WT monkey sperm as described by Putkhao and colleagues [48]. rESCs were cultured on mitomycin C inactivated mouse embryonic fibroblast feeder cells in Knockout DMEM (Gibco) with 20% Knockout serum replacement (KSR; Invitrogen) supplemented with 2-mM L-glutamine, 1x non-essential amino acid (NEAA; Lonza), and 4 ng/ml of human basic fibroblast growth factor (bFGF; R&D system) at 37°C with 5% CO_2. Colonies reached 70–80% confluences in approximately 4 days and were mechanically passaged onto the freshly prepared feeder cells.

In vitro derivation of spermatogenic cells

rES5, rES7, and rES-WT were *in vitro* differentiated into spermatogenic cells (rSCs) using mouse spermatogonial stem cell (SSC) culture medium as previously described [50]. Briefly, rESCs were cultured on STO-LIF feeder cells until 70–80% confluence when rESC culture medium was replaced with SSC medium composed of MEM alpha (Thermofisher, Inc.), 0.2% bovine serum albumin (Sigma), 5 μg/ml insulin (Sigma), 10 μg/ml transferrin (Sigma), 60-μM putrescine (Sigma), 2-mM L-glutamine (Invitrogen), 50-μM β-mercaptoethanol (Sigma), 1-ng/ml human basic fibroblast growth factor (bFGF; R&D System), 20-ng/ml glial-derived neurotrophic factor (GDNF, R&D System), 30-nM sodium selenite (Sigma), 2.36-μM palmitic acid (Sigma), 0.21-μM palmitoleic acid (Sigma), 0.88-μM stearic acid (Sigma), 1.02-μM oleic acid (Sigma), 2.71-μM linoleic acid (Sigma), 0.43-μM linolenic acid (Sigma), 10-mM HEPES (Sigma), and



	rES1	rES2	rES3	rES4	rES5	rES6	rES7
Generation	F1	F1	F1	F1	F1	F1	F1
Gender	M	F	F	F	M	F	M
WT 8Q	Y	Y	Y	Y	Y	Y	Y
<35Q	Y	Y	Y	Y	--	Y	Y
40-60Q	Y	--	--	--	--	--	--
>60Q	Y	Y	Y	Y	Y	--	Y
rSCs	N/A	N/A	N/A	N/A	Y (F2)	N/A	Y (F2)

Fig. 1 Experimental flow chart. Seven rESC lines (rESCs1–7; F1) [48] inherited mutant *HTT* transgenes from rHD1 (F0) [49] were established from rHD1 sperm (F1) derived blastocysts (F1). rES5 and rES7 were differentiated into F2 spermatogenic cells (SCs) *in vitro*. M: male; F: female, WT: wild type; rES: HD monkey embryonic stem cells; rSCs: HD monkey spermatogenic cells; N/A: not applied; Y: yes; N: no

0.5x penicillin/streptomycin (Thermofisher). Fresh SSC media were replaced every 2 days until day 10 of differentiation when cells were collected and preserved for different analyses described below [50].

Gene expression analysis by qRT-PCR

To examine the progression of differentiating spermatogenic cells, rES5, rES7, and rES-WT were collected at days 0 (ESCs before differentiation), 3, 5, 7, and 10 for determining the expression levels of spermatogenic cell markers. Total RNA was extracted using Trizol (Life Technologies). cDNA was synthesized using High-Capacity cDNA Reverse Transcription Kits (Applied Biosystems) using 500 ng of RNA samples. Quantitative mRNA expression was measured by using TaqMan Gene Expression Master Mix (Applied Biosystems) and rhesus macaque TaqMan primers for VASA (Rh02798028_m1), PLZF (Rh02834704_m1), and PIWIL1 (Rh04256262_m1) (Thermofisher). Bio-Rad CFX96 system was used for the reaction. Gene expression results were first normalized with endogenous control glyceraldehyde-3-phosphate dehydrogenase (GAPDH; Rh02621745_g1). The normalized data was then normalized to day 0 samples to determine the relative expression of rSCs across the 10-day differentiation by using ΔΔCt method. Three biological replicates and three technical replicates were performed for each sample.

Determining CAG repeat size

PCR and GeneScan analyses: To demonstrate the CAG repeat size, the polyglutamine tract was first amplified by polymerase chain reaction (PCR) followed by GeneScan analysis. Genomic DNA was isolated using Qiagen DNeasy Blood & Tissue Kits (Qiagen) following the manufacturer's instructions. PCR was carried out using primers targeting the flanking regions of CAG repeats in the exon1 of *HTT* gene: HD32F (forward sequences 5'-FAM-CTACGAGTCCCTCAAGTCCTTCCAGC-3') and MD177R (reverse sequence 5'-GACGCAGCAGCGGCTGTGCCTG-3'). Each reaction was performed using 100 ng of genomic DNA, 0.2 μM of forward and reverse primers, 1X PCR Buffer (Takara), 1-mM deoxyribonucleotide triphosphates (dNTP; Takara), 0.5-U Taq polymerase (Takara), 4 μM of Betain (Sigma), and nuclease-free water all added to bring up the total reaction to 50 μL. The PCR was run at 98 °C for 5 min, followed with 40 cycles of 96 °C for 5 min, 67 °C for 45 s, 72 °C for 1.5 min, and a final extension of 72 °C for 10 min.

GeneScan analysis of each PCR product was performed at the Emory Integrated Genomics Core. Of the PCR products, 1.5 μL were mixed with 0.5 μL of GeneScan™ 1000 ROX™ (ThermoFisher, Waltham, MA) size standard and 9.5 μL of Hi-Di™ Formamide (Applied Biosystems). The samples were denatured at 95°C for 5 min and run with 3130XL Genetic Analyser (Applied Biosystems).

Immunostaining

At the end of the 10-day spermatogenic cell differentiation, the SSC medium was removed and washed three times with phosphate-buffered saline solution (PBS: Lonza) before fixation with 4% paraformaldehyde (PFA) for 30 min at room temperature. Fixed cells were blocked with blocking solution containing 1X PBS, 0.20% Triton X-100 (Sigma), 2% bovine serum albumin (BSA; Sigma), 3-mM sodium azide (Sigma), 0.1% saponin (Sigma), and 5% normal goat serum (NGS) or donkey serum (Sigma) at 4°C overnight. The primary antibody, PLZF 10 μg/ml (R&D System), was diluted with blocking buffer and incubated with the cell sample overnight at 4°C. DNA was counterstained with Hoechst at 1 μg/ml for 10 min before imaging. Samples were examined by a BX51 epifluorescence microscope (Olympus, Inc.), and images were captured by CellSens software (Olympus, Inc.).

Data analysis

GeneScan data were analyzed by using GeneMarker® (SoftGenetics Version 2.2.0). The normal Gaussian distribution was used to fit into the raw GeneScan data. The curve fit results were then analyzed as previously described [13]. Briefly, raw data were imported into MATLAB (MatWorks R2019b) and analyzed using ipf.m function to determine the degree of best fit that were represented as an error less than 5% and overall fitness R^2 value of more than 0.97 (S1 Table). Curve fit results were then overlaid to the electrogram from GeneScan using Adobe Photoshop (Adobe). The mean of curve positions was used to calculate the number of repeats added, or the expansion of CAG repeats at the beginning and the end of the culture or differentiation as shown in Supplemental Table 1. CAG repeat size in rESs1–7 was divided into three allele clusters of <35Q, 40–60Q, and >60Q for further analysis. To calculate the number of repeats added to the mutant *HTT* alleles, we used the method described in previous studies [51, 52]. Briefly, the mean number of repeats at the end of culture was subtracted from the repeat number of original allele size in the earliest sample (S1 Table). The expansion of CAG repeats in rESs1–7, rSC5, and rSC7 was calculated using repeat size from the curve fit data. To determine daily expansion rates, the number of added repeats was divided by the number of days in culture.

Statistical analysis

Statistical significance was determined by the Mann-Whitney *U* test, and Wilcoxon signed ranks test were used for data with non-normal distribution. Correlation analysis of CAG repeat size and daily expansion rates rESCs and rSCs were performed by Spearman correlation analysis. The correlation level of each data set was indicated by Spearman's correlation coefficient (Spearman r: −1 to +1). Statistical significance denoted with $*p<0.05$, $**p<0.01$, $***p<0.001$.

Results

CAG repeat instability in HD monkey lymphocytes and sperm

HD monkey (rHD1) [48, 49, 53] lymphocytes (3 and 44 months of age) and sperm (44 months of age) were collected and the size of CAG repeats were determined. Because of the aggressive disease development in rHD1, we were not able to collect sperm samples at multiple time points. We have confirmed the normal *HTT* allele has around 8Q, and CAG mosaicism was observed in the mutant *HTT* alleles in both lymphocytes and sperm (Fig. 2 and Supplemental Fig. 1). Three clusters of mutant *HTT* alleles were observed in both samples (Supplemental Fig. 1a).

To examine the pattern of CAG mosaicism, we used the Gaussian distribution curve fit method to investigate the dynamics of mutant *HTT* alleles [54]. The deconvoluted individual alleles are represented as red curves, while the overall curve fit is represented as a black curve overlaid over the GeneScan electrogram (Fig. 2a). The alleles were grouped based on CAG repeat size and divided into three clusters; (1) <35Q, (2) 40Q–60Q, and (3) >60Q, accordingly (Fig. 2a). We observed minimal or no change in allele clusters <35Q (Fig. 2 and Supplemental Fig. 1). Similarly, the allele clusters of 40–60Q were also relatively stable in lymphocytes with no significant changes between 3 (prodromal stage) and 44 (symptomatic stage) months of age [48, 49, 53]. A shift of a smaller peak (~75Q) toward the larger CAG repeat in allele clusters >60Q was observed in lymphocytes (Fig. 2a and b). Compared to lymphocytes, allele clusters of 40–60Q in sperm had larger CAG repeats while the allele in the larger cluster of >60Q displayed higher CAG mosaicism. These observations suggest alleles with CAG repeats of over 60Q are extremely unstable in sperm, which supports the notion of tissue-specific CAG instability [38, 54]. Consistently, our recent study revealed similar CAG repeat expansion patterns in a longitudinal study of HD monkey lymphocytes and sperm [13]. Mutant *HTT* alleles in HD monkeys displayed different expansion patterns between sperm and lymphocytes, whereas lymphocytes tend to expand continuously with a small number of CAGs. In contrast, periodic expansion occurred in sperm, especially alleles with a larger size of CAG repeat as described by Møllersen and colleagues [31]. Our results are also consistent with prior studies that high CAG repeat instability and CAG mosaicism were found in male gametes [17, 27, 33–38] (Fig. 2a and c; Supplemental Table 1).

Assessment of stability of CAG repeats in HD monkey embryonic stem cells (rESCs) in *in vitro* culture

To derive rHD1 ESC lines, rHD1 sperm were used to produce F1 embryos by *in vitro* fertilization followed by the derivation of ESC lines as described previously [48]. A total of three males (rES1, 5, and 7) and four females (rES2, 3, 4, and 6) ESC lines were established [48]. These rESC lines were considered as the F1 progenies of rHD1 (Fig. 1). Among the seven rESC lines, rES1 was the only line that inherited all three mutant *HTT* allele clusters from rHD1 (Fig. 3a). rES2, 3, 4, and 7 inherited two mutant *HTT* allele clusters of <35Q and >60Q. rES5 and rES6 inherited allele clusters of >60Q and <35Q, respectively (Fig. 1 and Fig. 3a; Supplemental Fig. 1).

All cell lines were cultured and maintained under the same conditions for an extended period between 44 and 164 days (Table 1). To determine the effects of extended culture on CAG instability, samples were collected at the beginning

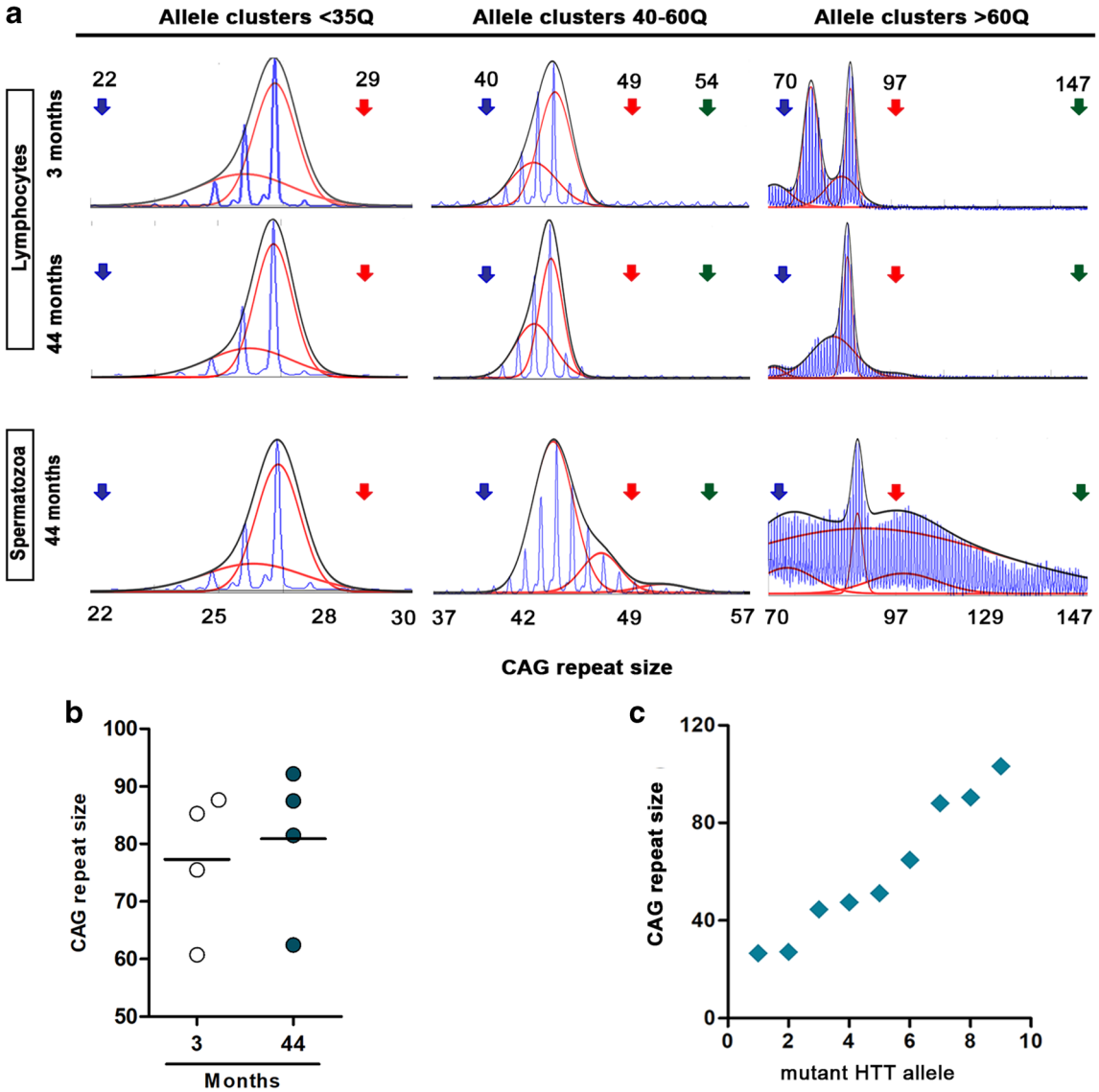


Fig. 2 CAG repeat instability in rHD1 lymphocytes and sperm. **a** Curve fit analysis of mutant *HTT* alleles. The cluster of CAG repeats are classified into allele clusters (1) <35Q, (2) 40–60Q, and (3) >60Q. Black lines represent the overall curve fit overlay onto the spectrum of each allele clusters. Red lines represent individual curves that fit to the expansion of mutant *HTT* alleles. The number of CAG repeats (Q) was denoted at the X-axis and by the arrows. **b** Distribution of CAG repeat size in lymphocytes at three and 44 months of age. The black horizontal line represents the median of each data set based on curve fit data on allele clusters of >60Q in (**a**). **c** CAG repeat mosaicism in sperm of rHD1 monkey based on curve fit data in (**a**)

and the end of the culture for GeneScan and curve fit analyses (Fig. 3a). To demonstrate the expansion of mutant *HTT* alleles, we used the number of repeats added as previously reported [51, 52] to distinguish the stable and expanded alleles. All rESC lines except rES5 inherited the allele clusters of <35Q, which remained stable up to 164 days in culture with no major change in CAG repeat size and the number of repeats added (Fig. 3a and b; Supplemental Fig. 2 and Supplemental Table 1). The mutant *HTT* allele cluster of 40–60Q was relatively stable after 72 days in culture (Fig. 3a; Supplemental Fig. 2 and Supplemental Table 1).

rESCs with allele clusters of >60Q showed a high instability with up to 11 repeats added after an extended culture period (Fig. 3a, b, and c; Supplemental Table 1).

In vitro derivation of spermatogenic cells from HD monkey ESCs

To investigate CAG repeat instability in spermatogenic cells, a novel stem cell model was used to derive HD monkey spermatogenic cells *in vitro*. Two male HD rESC lines (rES5, rES7) and one WT ESC line (rES-WT) were used to derive

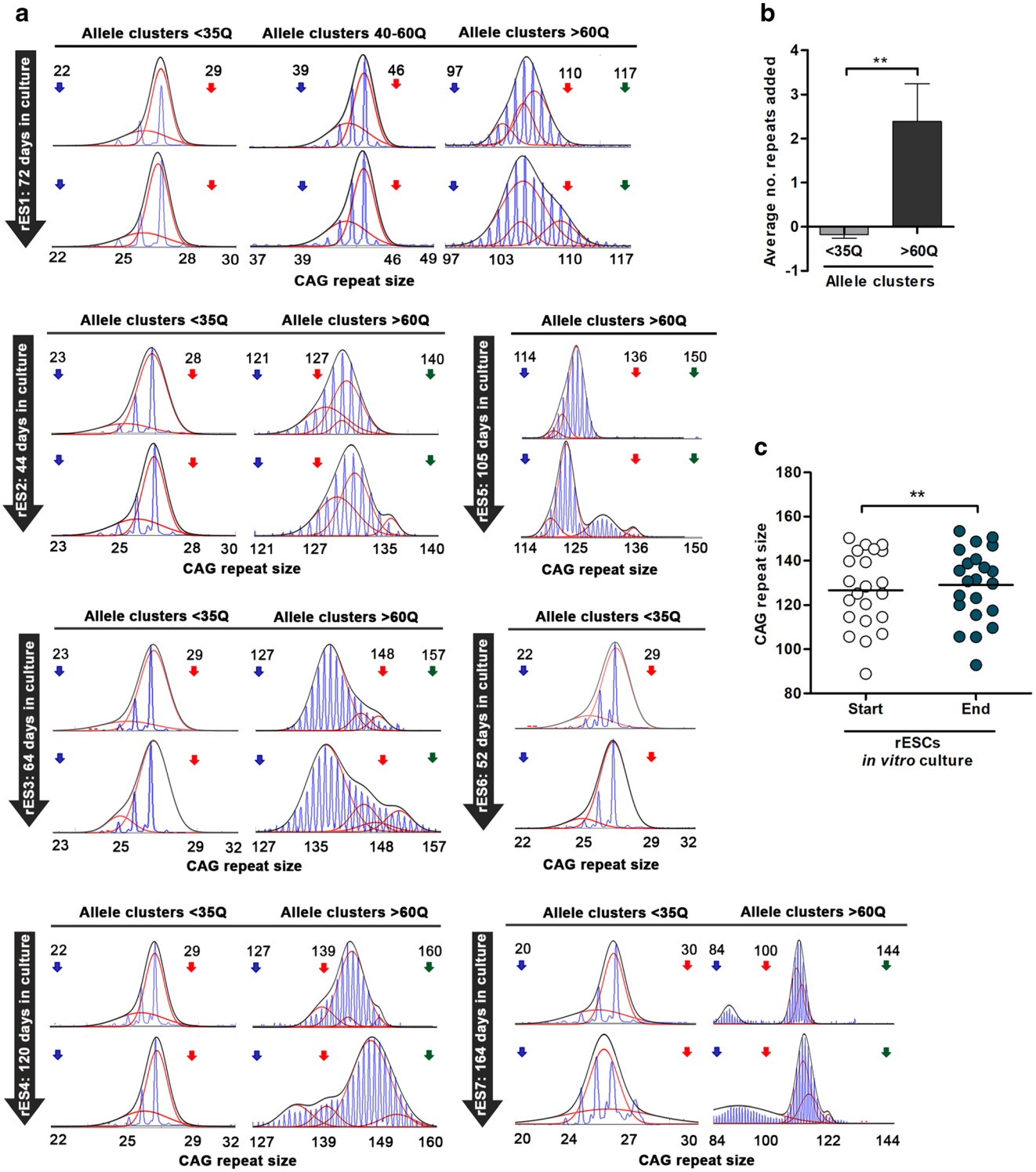


Fig. 3 Assessments of CAG repeat instability in the long-term culture of rESs1–7. **a** Curve fit analysis of CAG repeat expansion in rESs1–7 at the start and the end of *in vitro* culture. **b** The average number of repeats added in allele clusters of <35Q and >60Q in rESs1–7 after long-term culture. **c** Expansion of CAG repeats in allele cluster of >60Q. An open and solid circle represents CAG repeat size from an individual red curve in (**a**) *in vitro* culture beginning and end, respectively. A black horizontal line represents the median of each data set. $^{**}p$ <0.01

spermatogenic cells, rSC5, rSC7, and rSC-WT, respectively, using a 10-day spermatogonial stem cell (SSC) differentiation protocols (Fig. 4a) [50, 55, 56]. The elevated expression of DEAD-Box Helicase 4 (*VASA*), a germ cell–specific transcript; promyelocytic leukemia zinc finger (*PLZF*), a spermatogonia marker; and PIWI-like protein 1 (*PIWIL1*), a

Table 1 Summary of samples derived from rHD1

Name	Age (D/M) or passage (P)	Treatment	Culture time	Sample type	Polyglutamine (Q) size	Repeats added
Sperm	44 months	*In vivo*	N/A	Germ cells	26–103Q	N/A
Lymphocytes	3 months 44 months			Somatic cells	26–87Q 26–92Q	4.54Q
rES1	P10 P28	*In vitro* culture	72 days	rESs1–7	26–106Q 26–109Q	2.80Q
rES2	P7 P18		44 days		26–130Q 26–135Q	4.83Q
rES3	P12 P28		64 days		26–147Q 26–150Q	3.35Q
rES4	P22 P52		120 days		26–150Q 25–153Q	3.19Q
rES5	P6 P32		104 days		120–125Q 120–136Q	11.73Q
rES6	P38 P51		52 days		26–27Q 26–27Q	−0.22Q
rES7	P7 P48		164 days		26–114Q 25–124Q	9.75Q
rSC5	Day 0 Day 10	Spermatogenic cell differentiation	10 days	rSC derived from rES5	120–136Q 124–145Q	8.67Q
rSC7	Day 0 Day 10		10 days	rSC derived from rES7	25–124Q 26–131Q	7.42Q
rSC-WT	Day 0 Day 10		10 days	rSC derived from rES-WT	8Q 8Q	0

Q: CAG repeat; rES: HD monkey embryonic stem cell; rSC: HD monkey spermatogenic cell; D: day; M: month; P: number of passages; WT: wild type; N/A: not applied

spermatocyte marker (Fig. 4b, c, and d) suggest the progressive development toward spermatogenic lineage during the 10-day differentiation (Supplemental Fig. 3a). Based on the expression patterns of spermatogenic cell markers, $PLZF^{+}$ spermatogonia-like cells were the main spermatogenic cell population after 10-day differentiation (Supplemental Fig. 3).

CAG repeat instability of *in vitro* derived spermatogenic cells

To determine if *in vitro* directed differentiation toward spermatogenic cells provokes CAG repeat instability, we performed GeneScan analysis on spermatogenic cells (rSCs) derived from rES5, rES7, and rES-WT. Normal *HTT* alleles with 8Q were detected in rES-WT as well as in rES5 and rES7 with no CAG expansion observed after spermatogenic cell differentiation (Supplemental Fig. 1c, Table 1). We then performed curve fit analyses on spermatogenic cells (rSCs) derived from rES5 (Fig. 5a) and rES7 (Fig. 5b). During the 10-day spermatogenic differentiation, rSC5 and rSC7 added up to 8.67 and 7.42 CAG repeats in allele clusters of >60Q, respectively (Fig. 5a and Supplemental Table 1). The number of CAG added in allele cluster of >60Q was much higher than <35Q cluster (Fig. 5c). Also, CAG repeat size in the allele cluster of >60Q was significantly increased after the differentiation (125.4 ± 5.390, $P = 0.012$) (Fig. 5d). These findings echo the correlation between CAG repeat instability and CAG repeat size in *in vitro* culture of rESCs (Fig. 3).

To further investigate the correlation between CAG repeat instability and time in culture, and if *in vitro* differentiating spermatogenic cells promote CAG repeat expansion, we compared the average number of repeats added between rESC cultures (44–164 days) and 10-day *in vitro* spermatogenic cell differentiation in small and large allele clusters of mutant *HTT* gene. Spermatogenic cells derived from rES5 and rES7 during the 10-day differentiation process showed a significant increase in the number of added CAG repeats compared to the long-term culture rESCs in alleles of <60Q (0.6150 ± 0.01833, $P = 0.017$) (Fig. 6a left panel) and >60Q (5.273 ± 1.043, $P = 0.043$) (Fig. 6a right panel) (Fig. 6a). The expansion rates (i.e., the number of CAG repeats added per day) of CAG repeats were used to examine the effect of culture time between long-term culture of rESCs and 10-day spermatogenic cell differentiation. Spermatogenic cells, rSC5 and rSC7, demonstrated the largest daily expansion rates compared to the long-term culture of rESs1–7 (44–165 days) and their parent rES5 and 7 (rES5-105 days and rES7-164 days) (Fig. 6b). Allele clusters of >60Q in spermatogenic cells showed a

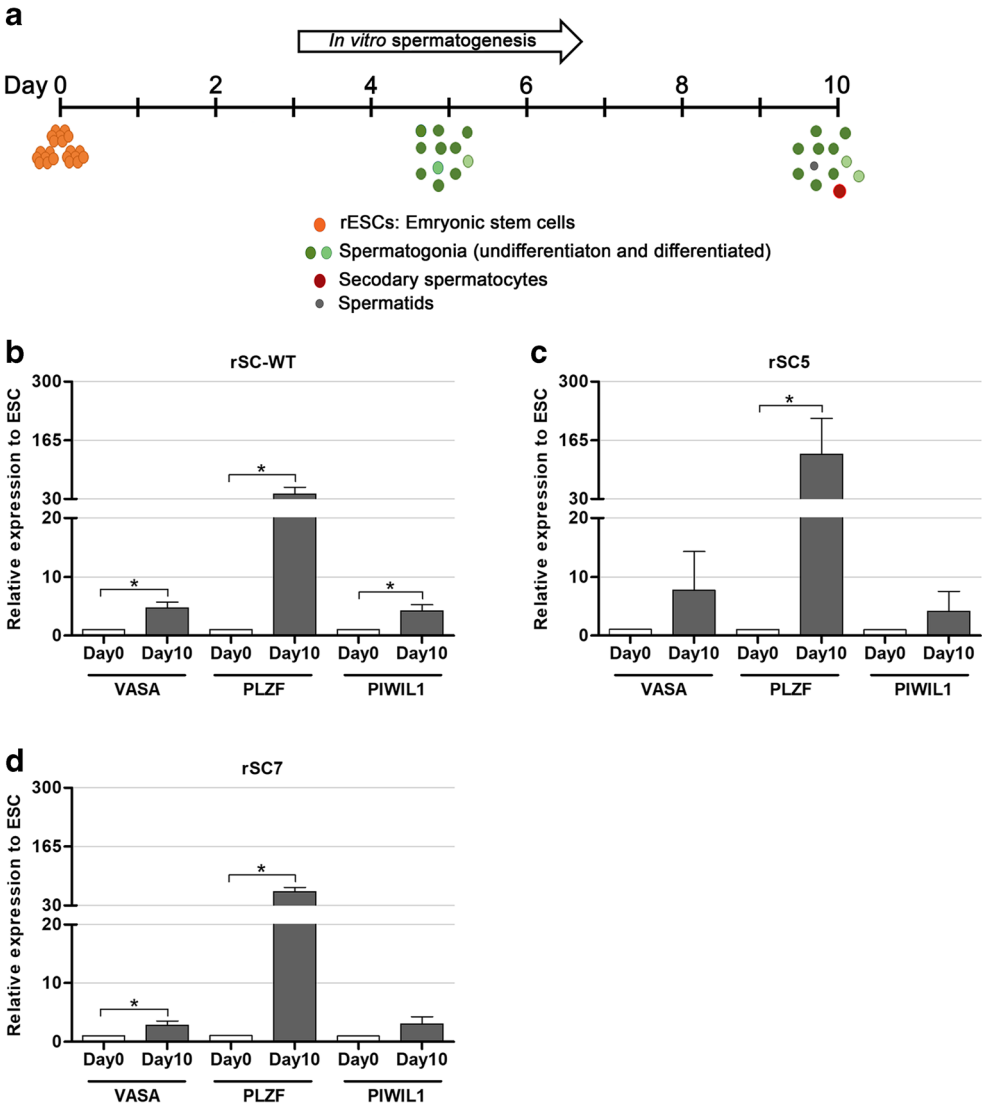


Fig. 4 *In vitro* spermatogenic cell differentiation of rES5 and rES7. **a** Experimental timeline of spermatogenic cell differentiation. **b**-d Expression of spermatogenic cell markers on day 0 (before spermatogenic cell differentiation) and day 10 (after 10-day spermatogenic cell differentiation) of spermatogenic cell differentiation of rSC-WT (**b**), rSC5 (**c**), and rSC7 (**d**) by quantitative real-time PCR (qRT-PCR). *p<0.05 and n=3

strong positive correlation between expansion rates and CAG repeat size (Fig. 6c, $R^2 = 0.8571$, $P = 0.0054$, solid square), but such correlation was not observed in rESCs (Fig. 6c, $R^2 = -0.0728$, $P = 0.3737$ open circle). In contrast, small allele clusters of less than 60Q were relatively stable in both rESCs and spermatogenic cells (Fig. 6c). Our data suggest the link between cell type, CAG repeat size, and CAG repeat expansion in HD spermatogenic cells.

Discussion

The purpose of this study was to examine whether *in vitro* derivation of spermatogenic cells from HD monkey ESCs can replicate TNR instability observed in sperm of HD patients. This novel stem cell model could provide an unprecedented platform for investigating CAG repeat instability and the underlying mechanism of genetic anticipation via the male germline. We used a transgenic HD monkey, rHD1, as our model for this proof of principle study. rHD1 is the first transgenic HD monkey that developed progressive clinical impairment paralleling human HD [49, 53, 57–61]. Based on GeneScan analysis of lymphocytes and sperm, we confirmed that there were three mutant *HTT* allele clusters (1) <35Q, (2) 40–60Q, and (3) >60Q in both tissues. rHD1 sperm exhibited high CAG mosaicism in mutant *HTT* allele clusters of >60Q (Fig. 2a and c). However, the same allele clusters in lymphocytes only showed small CAG repeat expansion between three and 44 months of age (Fig. 2a and b). On the other hand, allele clusters of <35Q and 40–60Q were relatively stable in both lymphocytes and sperm (Fig. 1). These findings were consistent with previous studies in humans [12, 35–37], rodent HD models [20, 21, 31, 43], and HD monkeys [13] in which alleles with large CAG repeats are more unstable and prone for expansions in male germ cells than in somatic cells (Fig. 2b and c).

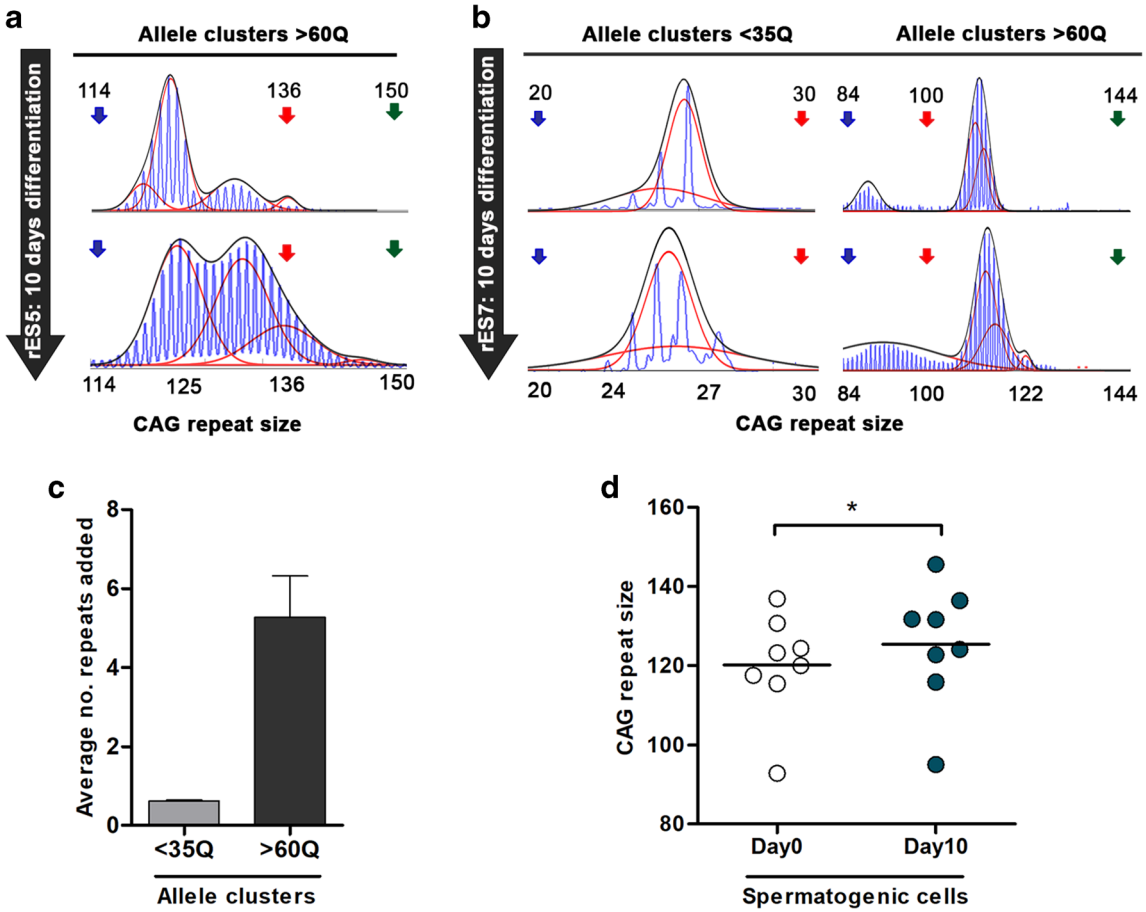


Fig. 5 Assessments of CAG repeat instability in *in vitro* derived spermatogenic cells. **a** and **b** Curve fit analysis of rES5 and rES7 spermatogenic cell differentiation. **c** An average number of repeats added in allele clusters of <35Q and >60Q in spermatogenic cells derived from rES5 and rES7. **d** Expansion of CAG repeats in allele clusters of >60Q in spermatogenic cells derived from rES5 and rES7. The open circle represents CAG repeat size from the individual red curve in "**a**" at day 0 before differentiation. The solid circle represents the CAG repeat size at day10 of spermatogenic cell differentiation. The black line represents the median of each data set. *p <0.05

A critical aspect of studying CAG repeat instability and genetic anticipation is the assessment of CAG repeat size in offspring. Although rodents are a useful model for studying genetic anticipation, the fundamental differences in spermatogenesis [39, 40, 46] and DNA repair response [62, 63] from humans have limited rodents from precisely replicating biological events that are critical for regulating CAG repeat stability during spermatogenesis. Multiple DNA repair pathways have been shown to play critical roles in regulating CAG repeat stability [18, 19, 25, 64–66]. Also, the manipulation of DNA repair enzyme activities, such as Msh2/3 of the mismatch repair pathway, has therapeutic benefit in HD mice [66, 67]. These observations further supported the importance of DNA repair mechanisms in regulating CAG repeat stability that could result in expansion as well as possible contraction if proper manipulation can be achieved [18, 19, 66, 67]. However, the observation of differentially expressed DNA repair genes in species with longer lifespan such as humans [62, 63] suggested the importance of a model that is closer to humans not only on the regulation of spermatogenesis but also the DNA repair responses that influences CAG repeat stability. Since it is ethically unacceptable to create human embryos for biological research, a non-human primate model such as rHD1 that replicates the progressive development of human HD is a viable animal model for bridging the gap between HD rodents and human HD patients.

We have established seven rhesus macaque ES cell lines (rESs1–7) from embryos conceived by using rHD1 sperm [48]. The inheriting patterns of the mutant *HTT* allele clusters suggest that they were segregated independently (S1 Fig). A clear inheritance was demonstrated in rHD1 (F0) and his progeny rESs1–7 cell lines (F1) (Fig. 3a). The CAG mosaicism in allele clusters of >60Q was observed in rHD1 sperm with an expansion of more than 100Q (Fig. 2a). Allele clusters of >60Q were inherited in six of the seven rES cell lines (Fig. 3a) and were more unstable compared to allele clusters of <35Q and 40–60Q (Fig. 3a). This evidence supports the size-dependency of CAG repeat instability. Because of the unique inheritance of the mutant *HTT* alleles in rESs1–7, we can examine if CAG repeat instability is size-dependent, and if

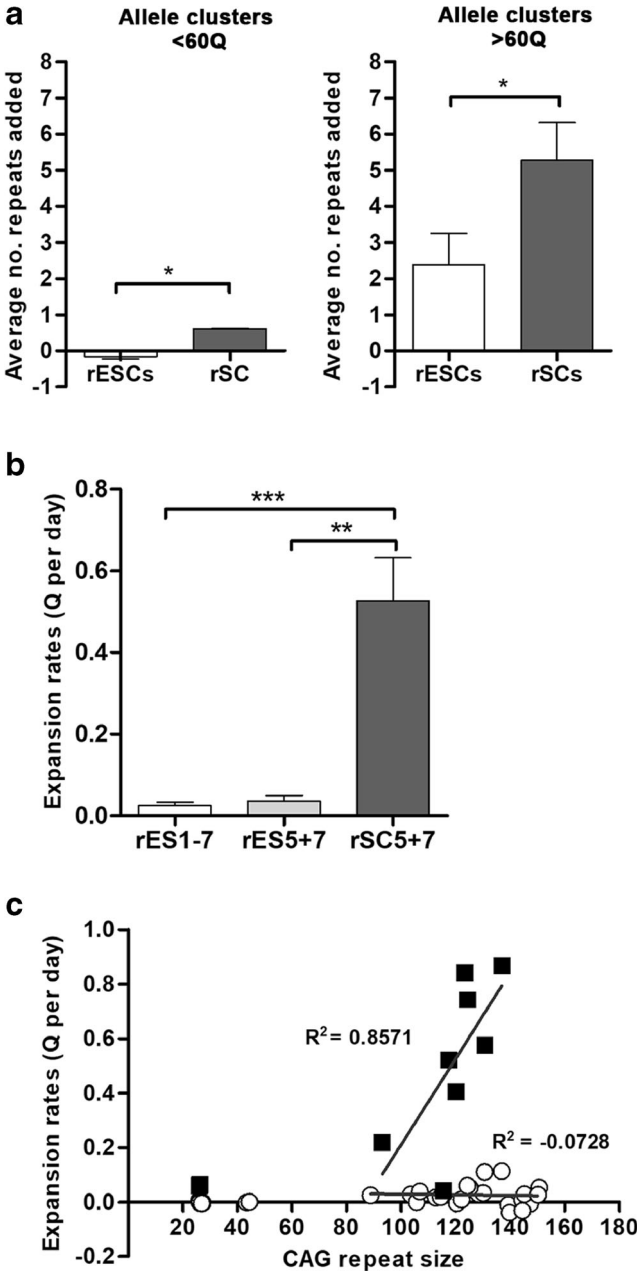


Fig. 6 CAG repeat instability and expansion rates in rESs1–7 and spermatogenic cells derived from rES5 and rES7. **a** An average number of repeats added in mutant *HTT* alleles of <60Q and >60Q clusters between rESCs and spermatogenic cells derived from rES5 and rES7 was presented. **b** Daily expansion rates of long-term cultured rESs1–7, long-term cultured rES5 and 7, and *in vitro* derived spermatogenic cells (rSC5 and 7). **c** Analysis between CAG repeat size and daily expansion rates of rESs1–7 (p = 0.3737, open circles) and *in vitro* derived spermatogenic cells (p = 0.0054, solid squares). *p <0.05, **p <0.01, ***p<0.001

long-term culture could provoke CAG repeat instability and lead to expansion or contraction in male and female sibling cell lines that inherited different mutant *HTT* alleles from the same sire.

Most *in vitro* studies on CAG repeat instability uses stable cell lines overexpressing mutant *HTT* or pluripotent stem cells (PSCs) derived from HD patients (HD-PSCs) [12, 14, 15, 18, 19, 21]. *In vitro* differentiation of neural cells from human HD-PSCs has been used to determine if neuronal cell types are more susceptible to CAG repeat expansion [19, 68]. Interestingly, a recent report revealed a discrepancy in CAG repeat instability and expansion between induced PSCs (iPSCs) and the parent fibroblasts [18, 19]. Donor HD fibroblasts were more prone to CAG repeat expansion in culture [18]. On the contrary, iPSCs derived from the same donor fibroblasts were relatively stable [18, 19]. The discrepancy between the HD-iPSCs and their parent fibroblasts has been linked to the reprogramming processes that alter the epigenetic landscape of the DNA repair response system and global epigenetic architecture [19]. While the underlying mechanisms that regulate the biological functions in iPSCs and parent fibroblasts have yet to be determined, this observation has raised the question of whether iPSCs, ESCs, or primary culture of patient cells is best for capturing CAG repeat instability merits further investigation.

Besides CAG repeat length, time is another factor that contributes to CAG repeat instability and expansion. *In vitro* culture of human HD fibroblasts and HD-PSCs, or mammalian cells overexpressing mutant *HTT* gene had resulted in CAG repeat expansion and were positively correlating to the length of the culture time [69–72]. Similar to aging, long-term culture increases oxidative stress-induced oxidative damages, specifically, double-strand break (DSB), which could accumulate during tissue culture and mediate CAG repeat instability [25, 65, 69]. We observed a similar correlation between CAG repeat instability, CAG repeat size, and time in the culture of rESCs. During 44 to 164 days of culture, CAG repeat expansion predominantly occurred in alleles with over 60Q but not in the allele clusters of <35Q or 40–60Q, regardless of the length of culture (Fig. 3). We did not observe gender-specific patterns on CAG repeat instability nor the levels of expansion in rESCs (Fig. 3). We showed a strong positive correlation between CAG repeat size and expansion rate per day in alleles with large CAG repeats during the *in vitro* derivation of spermatogenic cells (Fig. 6b and c). This finding suggested the increase of CAG repeat instability and the increase of CAG mosaicism in spermatogenic cells derived from rESs. By examining the number of repeats added to mutant *HTT* alleles, we demonstrated allele clusters with more than 60Q are more susceptible to CAG repeat expansion with significantly more CAG repeats added than the smaller alleles (Fig. 3b and c). This finding echoed the effect of age and the size of CAG repeats in a 4-year longitudinal study on lymphocytes and sperm of a different cohort of HD monkeys [13]. Our data further support that the larger the CAG repeats, the more it is prone to expansion than alleles with smaller CAG repeat size [13]. The largely expanded CAG repeat sequences

are sites of increased chromosomal fragility with an increasing tendency to form secondary structures such as hairpins, which are considered as a substrate of expansion [21]. The formation of a secondary structure poses additional stress to the local region that favors DNA breakage and triggers DNA repair responses [20, 21, 25, 65]. Thus, the formation of a secondary structure, the accumulation of DNA damage, and the accumulation of mutant HTT further impair DNA repair response result in elevated CAG repeat instability and expansion [25, 30, 65, 70, 72, 73]. A similar observation was observed in the cell culture of human and rodent cell models [12, 14–16, 18, 19, 21]; however, not in the context of cell lines derived from siblings that share a similar inheritance. The unique genetic configurations of rESs1–7 provide a novel *in vitro* platform for investigating the differential regulation on size-dependent instability of CAG repeats. One additional advantage of the HD monkey model is that additional cell lines, including ESCs, iPSCs, and primary cultures, can be established to examine if the origins of cell lines constitutes a discrepancy in the biology of genome stability and genetic anticipation.

CAG repeat instability in male germ cells remains challenging to examine without the use of rodent HD models or human HD testicular tissues [20, 22, 74–78]. To overcome the limitations of currently available models, we developed a stem cell model to replicate CAG repeat behavior that leads to genetic anticipation in differentiating spermatogenic cells. *In vitro* derived spermatogenic cells are the second generation (F2) of rHD1 (F0) while rHD1 sperm used for the production of F1 embryos and F1 ESCs are the first generation (F1) (Fig. 1). By *in vitro* differentiating rESCs to spermatogenic cells, we can examine CAG repeat instability in two generations of spermatogenic cells. Based on the expression of the spermatogenic cell markers in all of the rSCs after 10 days of differentiation, we confirmed the progression of differentiating spermatogenic cells with a majority of $PLZF^+$ mitotically active ("A" pale spermatogonia) or inactive spermatogonia ("A" dark spermatogonia)-like cells (Fig. 4b, c, and d) [40, 79, 80]. Immunostaining further confirms the expression of $PLZF^+$ spermatogonia-like cells (Supplemental Fig. 3). The increased expression of *PIWIL1* in rSCs after 10 days of differentiation suggested a small fraction of differentiating spermatogonia progressed towards secondary spermatocytes (Fig. 4b, c, and d) and spermatids [50].

Prior studies have suggested the discrepancy in CAG repeat expansion between rodents and humans. In mice, CAG repeat expansion occurs in post-meiotic mouse spermatogenic cells [20, 74–78]. However, in humans, CAG repeat expansion could occur in pre-meiotic cells (spermatogonia and spermatocytes), during meiosis in spermatocytes, and in post-meiotic spermatids, and the repair of double-strand breaks (DSB) and sperm elongation, as well as in spermatids when the chromatin is re-packaged with protamines [20, 22, 75, 77]. Therefore, if *in vitro* derived HD monkey spermatogenic cells replicated CAG repeat instability and promote expansions, CAG repeat expansion was likely to occur in pre-meiotic $PLZF^+$ spermatogonia-like cells. Although in our model, only a fraction of cells develop to the secondary spermatocyte and spermatid stage, our *in vitro* platform provides a new opportunity for investigating CAG repeat instability and the underlying mechanisms of genetic anticipation in spermatogenic cells *in vitro* for the first time. One of our current efforts is to optimize differentiation toward post-meiotic spermatogenic cells to capture a comprehensive view of how CAG repeat instability and genetic anticipation are regulated throughout spermatogenesis. Our *in vitro* spermatogenic cell model also opens a new opportunity for isolating specific spermatogenic cell types in spermatogenesis; thus, cell type–specific susceptibility and the underlying mechanism of CAG repeat instability in developing spermatogenic cells can be further investigated. Nonetheless, our data suggest that *in vitro* derived spermatogenic cells can replicate CAG repeat instability and expansion in early spermatogenesis, and it is a viable model for studying the dynamic regulation of CAG repeat instability and genetic anticipation that merit continues development.

In vitro derived spermatogenic cells, rSC5 and rSC7, have a significant increase of CAG repeat size (Fig. 5d) and a higher daily expansion rate compared to the long-term culture of rESs1–7, including their parent rES5 and rES7 (Fig. 6b). These results suggest tissue-specific instability [13, 35, 38, 81–83] and echo the observation of high CAG mosaicism in rHD1 sperm (Fig. 2) as well as a prior study in HD monkey lymphocytes and sperm [13]. Both ESCs and lymphocytes exhibit a similar pattern in CAG repeat expansion and are relatively stable compared to *in vitro* derived spermatogenic cells and sperm [13]. A recent study using human HD ESC-derived cardiomyocytes showed high CAG repeat stability with no expansion [15]. Interestingly, the three HD ESC lines used in the study have 38 and 43 CAG repeats, and no expansion was observed after 30 or 60 days of cardiomyocyte differentiation. Although the stability of CAG repeats in human HD ESC culture was not discussed, the teratoma studies indicated that human HD ESCs and the resultant teratomas have the same CAG repeat size. It was also fair to consider that HD ESC lines should have been cultured for weeks if not months from ESC derivation before performing teratoma experiments. Thus, human HD ESCs were stable in culture and in differentiated cardiomyocytes [15]. Human HD lymphoblasts also revealed a similar pattern of DNA instability. Most lymphoblastoid cell lines examined were below 41Q with little or no CAG repeat instability, and moderate instability was observed in cell lines with 42–59Q. Those with 60–120Q have high CAG repeat instability with high mutation penetrance after cultured for more than 6 months [12]. The high stability found in lymphoblastoid cell lines was consistent with the cardiomyocyte study where cell lines carrying 38 to 43Q, which were below the threshold repeat number of the instability of 60Q, exhibited high stability [12]. Similarly, a longitudinal study on HD monkey

lymphocytes and sperm echoes human lymphoblasts studies with a threshold repeat number of 62Q [13]. These studies further support our findings that mutant *HTT* alleles with <35Q and 40–60Q were relatively stable in long-term cultures or upon differentiation to spermatogenic cells. However, the increase in CAG repeat instability and periodic CAG repeat expansion was observed in cells carrying alleles with over 60Q in ESCs and spermatogenic cells (Fig. 3 and Fig. 5).

A recent study by Mollica and colleagues showed hypermethylation of DNA repair genes and down-regulation of DNA repair genes, and resulted in CAG repeat expansion in human HD fibroblasts [18]. The four DNA repair genes, *APEX1*, *BRCA1*, *RPA1*, and *RPA3* that were downregulated in HD cells, are also known as effectors of TNR stability [19, 84–86], 5-azacytidine-induced hypomethylation of the aforementioned DNA repair genes and stabilized CAG repeat from expansion [18]. This study suggested epigenetic control in regulating CAG repeat instability via the modulation of DNA repair machinery and merits further investigation, especially in paternal germ cells when dynamic changes in epigenetic landscape occur during spermatogenesis [41, 87–89]. Besides influencing DNA repair gene functions, changes in the epigenetic landscape also play a key role in the reprogramming of somatic cells to PSCs. Upon the derivation of iPSCs from human HD fibroblasts, Mollica and colleagues [19] observed induced CAG repeat stability in iPSCs and their derivative neural stem cells. They showed global upregulation of 5-hydroxymethylation (5-hmC) that resulted in increased levels of ten-eleven translocation 1 and 2 (TET1/2), which upregulate the expression levels of the aforementioned DNA repair genes and stabilize CAG repeats in iPSCs and neural stem cells [19]. These studies raise an interesting and important aspect on whether iPSCs, ESCs, or primary fibroblasts share similar if not identical epigenetic regulatory mechanisms and functions that might lead to different biological consequences such as CAG repeat stability.

The HD monkey model offers a unique opportunity that could overcome the ethical limitations in the derivation of human ESCs and the limitations of using human tissues. Primary cell cultures of HD monkeys can be established from multiple tissues from the same donors of ESCs and iPSCs. ESCs and iPSCs can also be derived from parents and their offspring, which is particularly important for examining intergenerational effect, and investigating CAG repeat instability and genetic anticipation among siblings. Moreover, longitudinal studies using samples collected over time could provide additional information on the impact of time, especially for diseases such as HD that progress as individual ages. This study utilizing lymphocytes at 3 and 44 months of age (Fig. 2) and a prior study on progressive CAG repeat expansion in lymphocytes and sperm [19] suggest that HD monkey is a viable model for investigating how CAG repeats behave over time in a cell type–specific manner. Although this study is primarily focused on the stability of CAG repeats at the *HTT* locus in *in vitro* differentiating spermatogenic cells because of its high relevance to HD pathogenesis and the constitution of genetic anticipation, other genomic domains and repetitive DNA loci including SINE or LINE (Interspersed Nuclear Elements) merits in-depth investigation in a future study to examine if *in vitro* culture or spermatogenic differentiation provoke non-specific genome modeling besides the *HTT* locus. The *in vitro* spermatogenic cell model provides an opportunity to dissect the mechanism of CAG repeat instability in differentiating male germ cells for the first time that have resulted in genetic anticipation impacting future generations in humans. Thus, novel therapeutics and approaches can be developed to stabilize CAG repeats, suppress or reverse CAG repeat expansion, and ameliorate or stop the vicious cycle of genetic anticipation. The *in vitro* spermatogenic cell model also opens a new opportunity for studying TNR instability and the underlying mechanism of genetic anticipation, not only in HD but also in other TNR diseases.

Conclusion

Studies on male germ cell genome instability or genetic anticipation have been limited to mouse models and human tissues. Due to the physiological differences between rodent and primate spermatogenesis, and the limited availability of human tissues, a stem cell model that could replicate CAG repeat behavior in differentiating spermatogenic cells would provide a unique opportunity for studying the underlying mechanisms of genome instability in male germline that result in periodic CAG repeat expansion and genetic anticipation affecting future generations. Our *in vitro* spermatogenic cell model recapitulates CAG repeat instability with much higher expansion rates compared to the extended culture of rESCs. Also, similar to the prior studies in rodents, humans, and non-human primates, CAG repeat instability in *in vitro* derived spermatogenic cells is positively correlated to the size of the repeats. While the current spermatogenic differentiation protocol has yet to be fully optimized, our goal is to optimize the current protocol to differentiate toward advanced spermatogenic cell types such as elongated spermatids. Thus, the susceptibility of spermatogenic cell types to CAG repeat expansion can be determined, and a therapeutic approach targeting specific spermatogenic cells can be developed to stabilize or reverse CAG repeat expansion to stop the vicious cycle of genetic anticipation.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10815-021-02106-3.

Acknowledgements All materials and experiments were performed at the Yerkes National Primate Research Center (YNPRC). We would like to thank YNPRC staff, Jinjing Yang, Siran Tian, Dr. Kanchana Punyawai, and other Dr. Chan's Lab members.

Author contributions Anthony W. S. Chan, In Ki Cho, and Sujittra Khampang conceptualized the study, designed experimentations, and wrote the manuscript. Sujittra Khampang and In Ki Cho performed experiments, data collection, and data analysis. Anthony W. S. Chan and Charles A. Easley IV supervised the spermatogenic cell differentiation, characterization, and experimental procedures. Wiriya Mahikul supervised the data analysis and statistical interpretation. Rangsun Parnpai supervised and reviewed the manuscript. All authors reviewed and approved the manuscript.

Funding YNPRC is supported by the Office of Research and Infrastructure Program (ORIP)/OD P51OD11132. The Transgenic Huntington's Disease Monkey Resource "THDMR" and this study were supported in part by grants awarded by the ORIP/NIH (OD010930) and NINDS/NIH (NS101701) to AWSC. This study was also supported in part by OD020182 and Georgia Partners in Regenerative Medicine Seed Grant to AWSC and CAE. We also received the support by the Emory University Research Council and Arthur and Sarah Merrill Foundation to AWSC and IKC. Sujittra Khampang and Rangsun Parnpai are supported by SUT-PhD scholarship, Suranaree University of Technology, Nakhorachasrima, Thailand.

Data availability N/A

Code availability N/A

Declarations

Disclaimer This work was prepared while Anthony W.S. Chan was employed at Yerkes National Primate Research Center and Emory University. The opinions expressed in this article are the author's own and do not reflect the view of the National Institutes of Health, the Department of Health and Human Services, or the United States government.

Ethics approval N/A

Consent to participate N/A

Consent for publication N/A

Conflict of interest The authors declare no competing interests.

References

1. Bates GP, Dorsey R, Gusella JF, Hayden MR, Kay C, Leavitt BR, et al. Huntington disease. Nat Rev Dis Primers. 2015;1:15005.
2. Ghosh R, Tabrizi SJ. Clinical features of Huntington's disease. Adv Exp Med Biol. 2018;1049:1–28.
3. Ross CA, Aylward EH, Wild EJ, Langbehn DR, Long JD, Warner JH, et al. Huntington disease: natural history, biomarkers and prospects for therapeutics. Nat Rev Neurol. 2014;10(4):204–16.
4. Aronin N, DiFiglia M. Huntingtin-lowering strategies in Huntington's disease: antisense oligonucleotides, small RNAs, and gene editing. Mov Disord. 2014;29(11):1455–61.
5. Didiot MC, Hall LM, Coles AH, Haraszti RA, Godinho BM, Chase K, et al. Exosome-mediated delivery of hydrophobically modified siRNA for Huntingtin mRNA silencing. Mol Ther. 2016;24(10):1836–47.
6. Gagnon KT, Pendergraff HM, Deleavey GF, Swayze EE, Potier P, Randolph J, et al. Allele-selective inhibition of mutant huntingtin expression with antisense oligonucleotides targeting the expanded CAG repeat. Biochemistry. 2010;49(47):10166–78.
7. Johnson E, Chase K, McGowan S, Mondo E, Pfister E, Mick E, et al. Safety of striatal infusion of siRNA in a transgenic Huntington's disease mouse model. J Huntingtons Dis. 2015;4(3):219–29.
8. Skotte NH, Southwell AL, Ostergaard ME, Carroll JB, Warby SC, Doty CN, et al. Allele-specific suppression of mutant huntingtin using antisense oligonucleotides: providing a therapeutic option for all Huntington disease patients. PLoS One. 2014;9(9):e107434.
9. Yang S, Chang R, Yang H, Zhao T, Hong Y, Kong HE, et al. CRISPR/Cas9-mediated gene editing ameliorates neurotoxicity in mouse model of Huntington's disease. J Clin Invest. 2017;127(7):2719–24.
10. McColgan P, Tabrizi SJ. Huntington's disease: a clinical review. Eur J Neurol. 2018;25(1):24–34.
11. Rawlins MD, Wexler NS, Wexler AR, Tabrizi SJ, Douglas I, Evans SJ, et al. The prevalence of Huntington's disease. Neuroepidemiology. 2016;46(2):144–53.
12. Cannella M, Maglione V, Martino T, Ragona G, Frati L, Li GM, et al. DNA instability in replicating Huntington's disease lymphoblasts. BMC Med Genet. 2009;10:11.
13. Clever F, Cho IK, Yang J, Chan AWS. Progressive polyglutamine repeat expansion in peripheral blood cells and sperm of transgenic Huntington's disease monkeys. J Huntingtons Dis. 2019;8(4):443–8.
14. Hung CL, Maiuri T, Bowie LE, Gotesman R, Son S, Falcone M, et al. A patient-derived cellular model for Huntington's disease reveals phenotypes at clinically relevant CAG lengths. Mol Biol Cell. 2018;29(23):2809–20.
15. Jacquet L, Neueder A, Foldes G, Karagiannis P, Hobbs C, Jolinon N, et al. Three Huntington's disease specific mutation-carrying human embryonic stem cell lines have stable number of CAG repeats upon in vitro differentiation into cardiomyocytes. PLoS One. 2015;10(5):e0126860.
16. Manley K, Pugh J, Messer A. Instability of the CAG repeat in immortalized fibroblast cell cultures from Huntington's disease transgenic mice. Brain Res. 1999;835(1):74–9.
17. McMurray CT. Mechanisms of trinucleotide repeat instability during human development. Nat Rev Genet. 2010;11(11):786–99.
18. Mollica PA, Reid JA, Ogle RC, Sachs PC, Bruno RD. DNA Methylation leads to DNA repair gene down-regulation and trinucleotide repeat expansion in patient-derived Huntington disease cells. Am J Pathol. 2016;186(7):1967–76.
19. Mollica PA, Zamponi M, Reid JA, Sharma DK, White AE, Ogle RC, et al. Epigenetic alterations mediate iPSC-induced normalization of DNA repair gene expression and TNR stability in Huntington's disease cells. J Cell Sci. 2018;131(13).
20. Simard O, Gregoire MC, Arguin M, Brazeau MA, Leduc F, Marois I, et al. Instability of trinucleotidic repeats during chromatin remodeling in spermatids. Hum Mutat. 2014;35(11):1280–4.
21. Usdin K, House NC, Freudenreich CH. Repeat instability during DNA repair: insights from model systems. Crit Rev Biochem Mol Biol. 2015;50(2):142–67.
22. Yoon SR, Dubeau L, de Young M, Wexler NS, Arnheim N. Huntington disease expansion mutations in humans can occur before meiosis is completed. Proc Natl Acad Sci U S A. 2003;100(15):8834–8.
23. Heitz D, Devys D, Imbert G, Kretz C, Mandel JL. Inheritance of the fragile X syndrome: size of the fragile X premutation is a major determinant of the transition to full mutation. J Med Genet. 1992;29(11):794–801.
24. Sherman SL, Jacobs PA, Morton NE, Froster-Iskenius U, Howard-Peebles PN, Nielsen KB, et al. Further segregation analysis of the fragile X syndrome with special reference to transmitting males. Hum Genet. 1985;69(4):289–99.

25. Jones L, Houlden H, Tabrizi SJ. DNA repair in the trinucleotide repeat disorders. Lancet Neurol. 2017;16(1):88–96.
26. Murmann AE, Yu J, Opal P, Peter ME. Trinucleotide repeat expansion diseases, RNAi, and cancer. Trends Cancer. 2018;4(10):684–700.
27. Orr HT, Zoghbi HY. Trinucleotide repeat disorders. Annu Rev Neurosci. 2007;30:575–621.
28. Kovtun IV, McMurray CT. Trinucleotide expansion in haploid germ cells by gap repair. Nat Genet. 2001;27(4):407–11.
29. Kraus-Perrotta C, Lagalwar S. Expansion, mosaicism and interruption: mechanisms of the CAG repeat mutation in spinocerebellar ataxia type 1. Cerebellum Ataxias. 2016;3:20.
30. Massey T, McAllister B, Jones L. Methods for assessing DNA repair and repeat expansion in Huntington's disease. Methods Mol Biol. 1780;2018:483–95.
31. Mollersen L, Rowe AD, Larsen E, Rognes T, Klungland A. Continuous and periodic expansion of CAG repeats in Huntington's disease R6/1 mice. PLoS Genet. 2010;6(12): e1001242.
32. Snell RG, MacMillan JC, Cheadle JP, Fenton I, Lazarou LP, Davies P, et al. Relationship between trinucleotide repeat expansion and phenotypic variation in Huntington's disease. Nat Genet. 1993;4(4): 393–7.
33. Duyao M, Ambrose C, Myers R, Novelletto A, Persichetti F, Frontali M, et al. Trinucleotide repeat length instability and age of onset in Huntington's disease. Nat Genet. 1993;4(4):387–92.
34. Kremer B, Almqvist E, Theilmann J, Spence N, Telenius H, Goldberg YP, et al. Sex-dependent mechanisms for expansions and contractions of the CAG repeat on affected Huntington disease chromosomes. Am J Hum Genet. 1995;57(2):343–50.
35. MacDonald ME, Barnes G, Srinidhi J, Duyao MP, Ambrose CM, Myers RH, et al. Gametic but not somatic instability of CAG repeat length in Huntington's disease. J Med Genet. 1993;30(12):982–6.
36. Myers RH, MacDonald ME, Koroshetz WJ, Duyao MP, Ambrose CM, Taylor SA, et al. De novo expansion of a (CAG)n repeat in sporadic Huntington's disease. Nat Genet. 1993;5(2):168–73.
37. Ranen NG, Stine OC, Abbott MH, Sherr M, Codori AM, Franz ML, et al. Anticipation and instability of IT-15 (CAG)n repeats in parent-offspring pairs with Huntington disease. Am J Hum Genet. 1995;57(3):593–602.
38. Telenius H, Almqvist E, Kremer B, Spence N, Squitieri F, Nichol K, et al. Somatic mosaicism in sperm is associated with intergenerational (CAG)n changes in Huntington disease. Hum Mol Genet. 1995;4(2):189–95.
39. Ehmcke J, Wistuba J, Schlatt S. Spermatogonial stem cells: questions, models and perspectives. Hum Reprod Update. 2006;12(3): 275–82.
40. Fayomi AP, Orwig KE. Spermatogonial stem cells and spermatogenesis in mice, monkeys and men. Stem Cell Res. 2018;29:207–14.
41. McCarrey JR. EPIGENETICS. The epigenome–a family affair. Science. 2015;350(6261):634–5.
42. McMurray CT, Kortun IV. Repair in haploid male germ cells occurs late in differentiation as chromatin is condensing. Chromosoma. 2003;111(8):505–8.
43. Neto JL, Lee JM, Afridi A, Gillis T, Guide JR, Dempsey S, et al. Genetic contributors to intergenerational CAG repeat instability in Huntington's disease knock-in mice. Genetics. 2017;205(2):503–16.
44. Aziz NA, van Belzen MJ, Coops ID, Belfroid RD, Roos RA. Parent-of-origin differences of mutant HTT CAG repeat instability in Huntington's disease. Eur J Med Genet. 2011;54(4):e413–8.
45. Norremolle A, Sorensen SA, Fenger K, Hasholt L. Correlation between magnitude of CAG repeat length alterations and length of the paternal repeat in paternally inherited Huntington's disease. Clin Genet. 1995;47(3):113–7.
46. Hermann BP, Cheng K, Singh A, Roa-De La Cruz L, Mutoji KN, Chen IC, et al. The mammalian spermatogenesis single-cell transcriptome, from spermatogonial stem cells to spermatids. Cell Rep. 2018;25(6):1650–67.e8.
47. Luetjens CM, Weinbauer GF, Wistuba J. Primate spermatogenesis: new insights into comparative testicular organisation, spermatogenic efficiency and endocrine control. Biol Rev Camb Philos Soc. 2005;80(3):475–88.
48. Putkhao K, Kocerha J, Cho IK, Yang J, Parnpai R, Chan AW. Pathogenic cellular phenotypes are germline transmissible in a transgenic primate model of Huntington's disease. Stem Cells Dev. 2013;22(8):1198–205.
49. Yang SH, Cheng PH, Banta H, Piotrowska-Nitsche K, Yang JJ, Cheng EC, et al. Towards a transgenic model of Huntington's disease in a non-human primate. Nature. 2008;453(7197):921–4.
50. Easley CA, Phillips BT, McGuire MM, Barringer JM, Valli H, Hermann BP, et al. Direct differentiation of human pluripotent stem cells into haploid spermatogenic cells. Cell Rep. 2012;2(3):440–6.
51. Gazy I, Hayward B, Potapova S, Zhao X, Usdin K. Double-strand break repair plays a role in repeat instability in a fragile X mouse model. DNA Repair. 2019;74:63–9.
52. Zhao X, Zhang Y, Wilkins K, Edelmann W, Usdin K. MutLγ promotes repeat expansion in a Fragile X mouse model while EXO1 is protective. PLoS Genet. 2018;14(10):e1007719.
53. Chan AW, Xu Y, Jiang J, Rahim T, Zhao D, Kocerha J, et al. A two years longitudinal study of a transgenic Huntington disease monkey. BMC Neurosci. 2014;15:36.
54. Mollersen L, Rowe AD, Illuzzi JL, Hildrestrand GA, Gerhold KJ, Tveteras L, et al. Neil1 is a genetic modifier of somatic and germline CAG trinucleotide repeat instability in R6/1 mice. Hum Mol Genet. 2012;21(22):4939–47.
55. Steves AN, Bradner JM, Fowler KL, Clarkson-Townsend D, Gill BJ, Turry AC, et al. Ubiquitous flame-retardant toxicants impair spermatogenesis in a human stem cell model. iScience. 2018;3: 161–76.
56. Steves AN, Turry A, Gill B, Clarkson-Townsend D, Bradner JM, Bachli I, et al. Per- and polyfluoroalkyl substances impact human spermatogenesis in a stem-cell-derived model. Syst Biol Reprod Med. 2018;64(4):225–39.
57. Chan AW, Jiang J, Chen Y, Li C, Prucha MS, Hu Y, et al. Progressive cognitive deficit, motor impairment and striatal pathology in a transgenic Huntington disease monkey model from infancy to adulthood. PLoS One. 2015;10(5):e0122335.
58. Kocerha J, Liu Y, Willoughby D, Chidamparam K, Benito J, Nelson K, et al. Longitudinal transcriptomic dysregulation in the peripheral blood of transgenic Huntington's disease monkeys. BMC Neurosci. 2013;14:88.
59. Lallani SB, Villalba RM, Chen Y, Smith Y, Chan AWS. Striatal interneurons in transgenic nonhuman primate model of Huntington's disease. Sci Rep. 2019;9(1):3528.
60. Meng Y, Jiang J, Bachevalier J, Zhang X, Chan AW. Developmental whole brain white matter alterations in transgenic Huntington's disease monkey. Sci Rep. 2017;7(1):379.
61. Raper J, Bosinger S, Johnson Z, Tharp G, Moran SP, Chan AW. Increased irritability, anxiety, and immune reactivity in transgenic Huntington's disease monkeys. Brain Behav Immun. 2016;58:181–90.
62. MacRae SL, Croken MM, Calder RB, Aliper A, Milholland B, White RR, et al. DNA repair in species with extreme lifespan differences. Aging (Albany NY). 2015;7(12):1171–84.
63. Mouse Genome Sequencing C, Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, et al. Initial sequencing and comparative analysis of the mouse genome. Nature. 2002;420(6915):520–62.
64. Lai Y, Budworth H, Beaver JM, Chan NL, Zhang Z, McMurray CT, et al. Crosstalk between MSH2-MSH3 and polbeta promotes trinucleotide repeat expansion during base excision repair. Nat Commun. 2016;7:12465.
65. Massey TH, Jones L. The central role of DNA damage and repair in CAG repeat diseases. Dis Model Mech. 2018;11(1).

66. Wheeler VC, Lebel LA, Vrbanac V, Teed A, te Riele H, MacDonald ME. Mismatch repair gene Msh2 modifies the timing of early disease in Hdh(Q111) striatum. Hum Mol Genet. 2003;12(3):273–81.
67. Pinto RM, Dragileva E, Kirby A, Lloret A, Lopez E, St Claire J, et al. Mismatch repair genes Mlh1 and Mlh3 modify CAG instability in Huntington's disease mice: genome-wide and candidate approaches. PLoS Genet. 2013;9(10):e1003930.
68. Ooi J, Langley SR, Xu X, Utami KH, Sim B, Huang Y, et al. Unbiased profiling of isogenic huntington disease hPSC-derived CNS and peripheral cells reveals strong cell-type specificity of CAG length effects. Cell Rep. 2019;26(9):2494–508.e7.
69. Cardozo-Pelaez F, Song S, Parthasarathy A, Hazzi C, Naidu K, Sanchez-Ramos J. Oxidative DNA damage in the aging mouse brain. Mov Disord. 1999;14(6):972–80.
70. Goold R, Flower M, Moss DH, Medway C, Wood-Kaczmar A, Andre R, et al. FAN1 modifies Huntington's disease progression by stabilizing the expanded HTT CAG repeat. Hum Mol Genet. 2019;28(4):650–61.
71. Jonson I, Ougland R, Klungland A, Larsen E. Oxidative stress causes DNA triplet expansion in Huntington's disease mouse embryonic stem cells. Stem Cell Res. 2013;11(3):1264–71.
72. Maiuri T, Mocle AJ, Hung CL, Xia J, van Roon-Mom WM, Truant R. Huntingtin is a scaffolding protein in the ATM oxidative DNA damage response complex. Hum Mol Genet. 2017;26(2):395–406.
73. Ferlazzo ML, Sonzogni L, Granzotto A, Bodgi L, Lartin O, Devic C, et al. Mutations of the Huntington's disease protein impact on the ATM-dependent signaling and repair pathways of the radiation-induced DNA double-strand breaks: corrective effect of statins and bisphosphonates. Mol Neurobiol. 2014;49(3):1200–11.
74. Brykczynska U, Hisano M, Erkek S, Ramos L, Oakeley EJ, Roloff TC, et al. Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. Nat Struct Mol Biol. 2010;17(6):679–87.
75. Budworth H, McMurray CT. A brief history of triplet repeat diseases. Methods Mol Biol. 2013;1010:3–17.
76. Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR. Distinctive chromatin in human sperm packages genes for embryo development. Nature. 2009;460(7254):473–8.
77. Kovtun IV, Goellner G, McMurray CT. Structural features of trinucleotide repeats associated with DNA expansion. Biochem Cell Biol. 2001;79(3):325–36.
78. Siklenka K, Erkek S, Godmann M, Lambrot R, McGraw S, Lafleur C, et al. Disruption of histone methylation in developing sperm impairs offspring health transgenerationally. Science. 2015;350(6261):aab2006.
79. Buaas FW, Kirsh AL, Sharma M, McLean DJ, Morris JL, Griswold MD, et al. Plzf is required in adult male germ cells for stem cell self-renewal. Nat Genet. 2004;36(6):647–52.
80. Hermann BP, Sukhwani M, Simorangkir DR, Chu T, Plant TM, Orwig KE. Molecular dissection of the male germ cell lineage identifies putative spermatogonial stem cells in rhesus macaques. Hum Reprod. 2009;24(7):1704–16.
81. Kovtun IV, Therneau TM, McMurray CT. Gender of the embryo contributes to CAG instability in transgenic mice containing a Huntington's disease gene. Hum Mol Genet. 2000;9(18):2767–75.
82. Telenius H, Kremer B, Goldberg YP, Theilmann J, Andrew SE, Zeisler J, et al. Somatic and gonadal mosaicism of the Huntington disease gene CAG repeat in brain and sperm. Nat Genet. 1994;6(4):409–14.
83. Wheeler VC, Persichetti F, McNeil SM, Mysore JS, Mysore SS, MacDonald ME, et al. Factors associated with HD CAG repeat instability in Huntington disease. J Med Genet. 2007;44(11):695–701.
84. Beaver JM, Lai Y, Xu M, Casin AH, Laverde EE, Liu Y. AP endonuclease 1 prevents trinucleotide repeat expansion via a novel mechanism during base excision repair. Nucleic Acids Res. 2015;43(12):5948–60.
85. Crespan E, Hubscher U, Maga G. Expansion of CAG triplet repeats by human DNA polymerases lambda and beta in vitro, is regulated by flap endonuclease 1 and DNA ligase 1. DNA Repair (Amst). 2015;29:101–11.
86. Mason AG, Tome S, Simard JP, Libby RT, Bammler TK, Beyer RP, et al. Expression levels of DNA replication and repair genes predict regional somatic repeat instability in the brain but are not altered by polyglutamine disease protein expression or age. Hum Mol Genet. 2014;23(6):1606–18.
87. Liu Y, Zhang Y, Yin J, Gao Y, Li Y, Bai D, et al. Distinct H3K9me3 and DNA methylation modifications during mouse spermatogenesis. J Biol Chem. 2019;294(49):18714–25.
88. Lombo M, Fernandez-Diez C, Gonzalez-Rojo S, Herraez MP. Genetic and epigenetic alterations induced by bisphenol A exposure during different periods of spermatogenesis: from spermatozoa to the progeny. Sci Rep. 2019;9(1):18029.
89. McSwiggin HM, O'Doherty AM. Epigenetic reprogramming during spermatogenesis and male factor infertility. Reproduction. 2018;156(2):R9–R21.