**OMTN, Volume *35***

## Supplemental information

# *CircHTT(2,3,4,5,6)* — co-evolving with the *HTT* CAG-repeat tract — modulates Huntington's disease phenotypes

**Jasmin Morandell, Alan Monziani, Martina Lazioli, Deborah Donzel, Jessica Döring, Claudio Oss Pegorar, Angela D'Anzi, Miguel Pellegrini, Andrea Mattiello, Dalia Bortolotti, Guendalina Bergonzoni, Takshashila Tripathi, Virginia B. Mattis, Marina Kovalenko, Jessica Rosati, Christoph Dieterich, Erik Dassi, Vanessa C. Wheeler, Zdenka Ellederová, Jeremy E. Wilusz, Gabriella Viero, and Marta Biagioli**

A

| Database | circID in database | Genome assembly | Genomic location |
|---|---|---|---|
| CIRCpedia | HSA_CIRCpedia_48392 | GRCh37/hg19 | chr4:3088665-3109150 |
| circBase | hsa_circ_0001392 | GRCh37/hg19 | chr4:3088665-3109150 |
| circBank | hsa_circHTT_004 | GRCh37/hg19 | chr4:3088665-3109150 |

*circHTT(2,3,4,5,6)*

B

*HTT* exon 2

*HTT* exon 3

*HTT* exon 4

*HTT* exon 5

*HTT* exon 6

C

*D. melanogaster*: 302 bp, Exon 1, Exon 2, Exon 3, 61 bp

*C. elegans*: 600 bp, Exon 1, Exon 2, Exon 3, Exon 4, 46 bp

human *HTT* exons 2-6 homologous sequence

**Figure S1 - Sequences of exons 2-6 of *HTT* orthologues are well conserved among vertebrate species. A**, List of publicly available circRNA databases and their respective ID for the same circular RNA stemming from the human *HTT* locus, chr4: 3088665-3109150, then called *circHTT(2,3,4,5,6)*. **B**, UCSC genome browser sequence alignments of *HTT* orthologues for selected vertebrate species demonstrates high degree of conservation in exons 2-6; **C**, Schematic representation of sequence alignment of human *HTT* exons 2-6 onto gene orthologues of *D. melanogaster* and *C. elegans.* Homologous sequences (green) are spread out over a different number of incomplete exons in both species, i.e. lower degree of conservation.

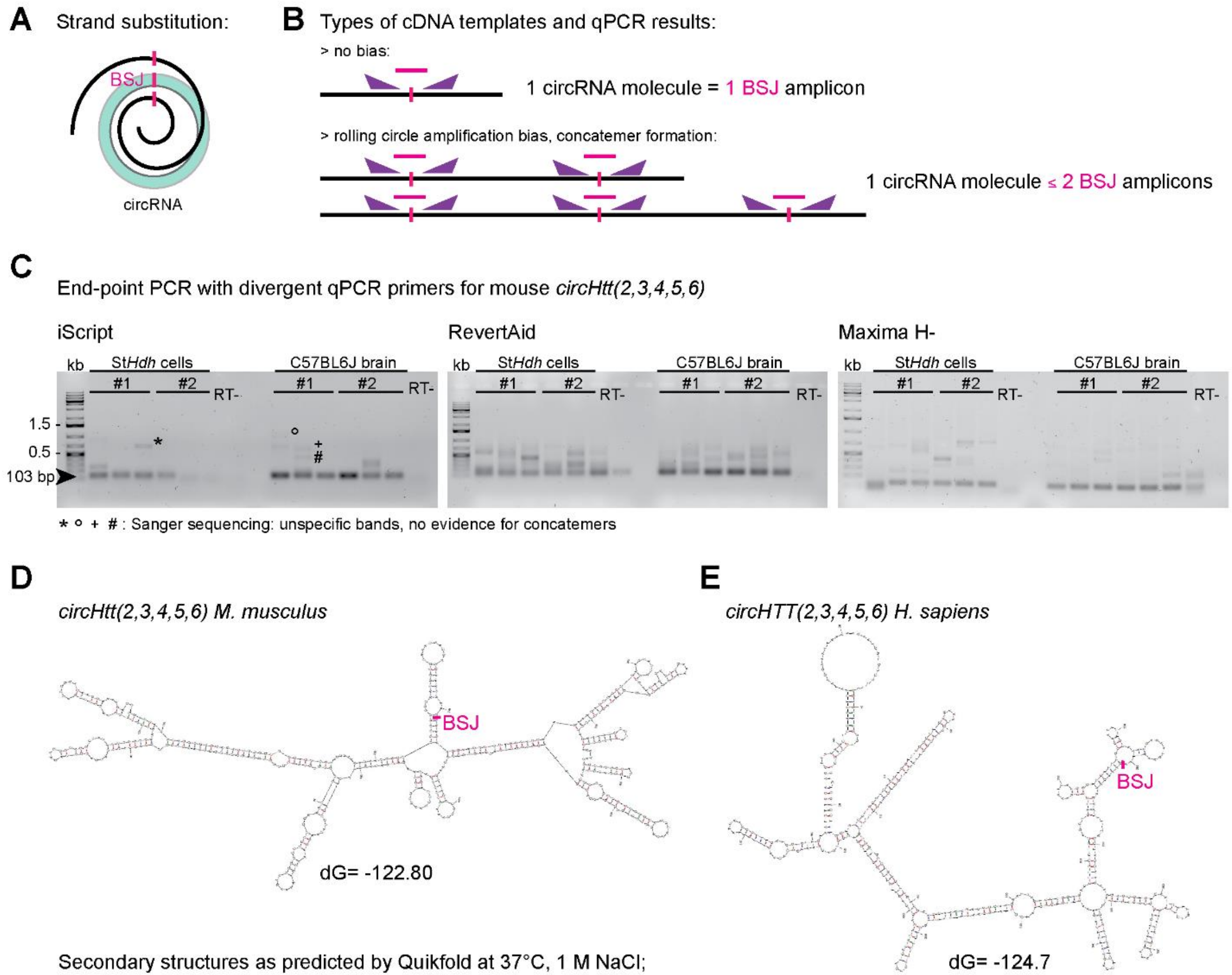


**Figure S2 - Rolling circle amplification bias unlikely in *circHtt(2,3,4,5,6)*/*circHTT(2,3,4,5,6)*.** **A**,**B**, Schematic representation of the rolling circle amplification bias in the assessment of circRNA abundance by RT-qPCR (divergent primers in purple). **C**, End-point PCR experiments using *circHtt(2,3,4,5,6)* qPCR primers and long elongation times (30 sec) on cDNA generated from total RNA of mouse St*Hdh,* as well as brain samples, using three different, commercially available, reverse transcriptase (RT) kits (iScript by Bio-Rad #1708890, RevertAid First Strand cDNA Synthesis Kit by Thermo Scientific #K1621, Maxima H Minus by Thermo Scientific #EP0753). The iScript reverse transcriptase resulted in least unspecific amplicons as reveled by agarose gel electrophoresis (left) as opposed to the other two RTs; Unspecific bands from the iScript reactions (*, °, +, #) were excised and Sanger sequenced, to test for evidence of rolling circle amplification/concatemers. None of the amplicons corresponded to *circHtt(2,3,4,5,6)* concatemeric sequences; **D**,**E**, CircRNA secondary structure prediction using the Quickfold tool for mouse and human *circHtt(2,3,4,5,6)/circHTT(2,3,4,5,6)* predicts highly structured circRNA conformations at the lowest free energy levels (location of the BSJ indicated in pink).

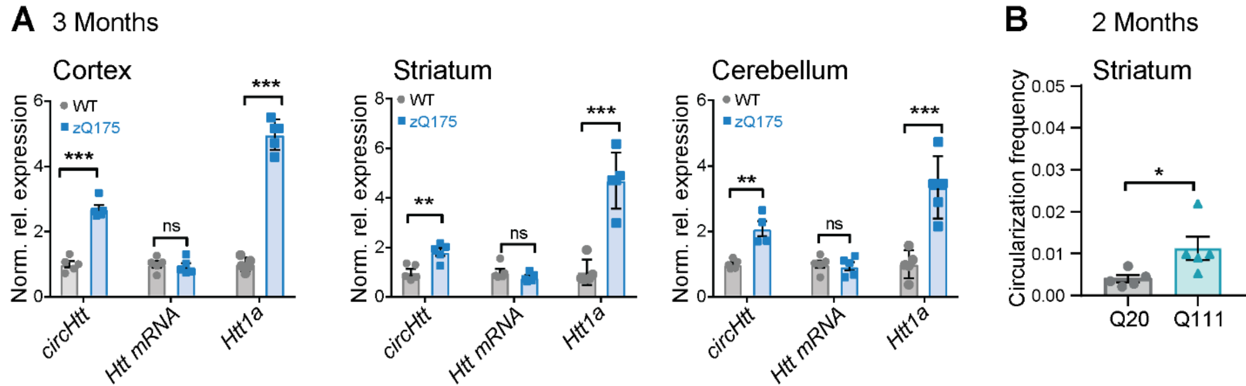


**Figure S3 - *CircHtt(2,3,4,5,6)* expression levels and circularization frequency in the zQ175 and Q111 mouse models for HD. A**, RT-qPCR on total RNA from brain samples of indicated brain regions (cortex, striatum, cerebellum) of 3 months old adult wild-type and zQ175 ($n$= 5 biological replicates per tissue and genotype, one-way ANOVA with Sidak`s multiple comparisons testing, $P < 0.001$ =***, $P < 0.01$ =**, $P < 0.05$ =*, ns= not significant). **B**, *CircHtt(2,3,4,5,6)* circularization frequency in the striatum of 2 months old Q20 and Q111 mice ($n$= *5* per genotype; the relative level of expression of *circHtt(2,3,4,5,6)* and linear isoforms was first calculated normalizing on the *Pgk1* housekeeping gene and subsequently the circularization frequency as ratio between back-splicing an linear splicing was computed, unpaired two-sided t-test, $P < 0.05$ =*).

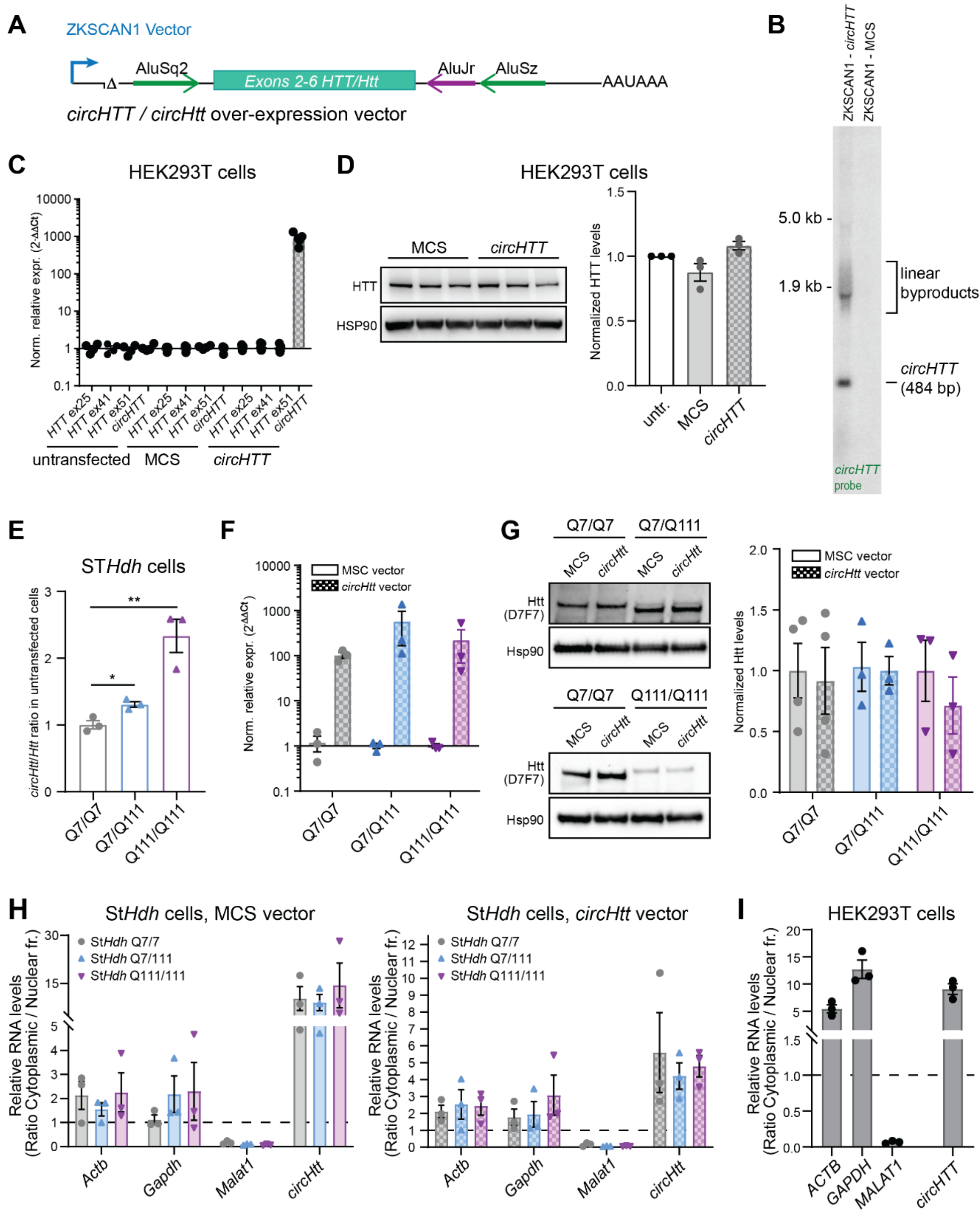


**Figure S4 - Over-expression of *circHtt(2,3,4,5,6)*/*circHTT(2,3,4,5,6)* does not alter *Htt*/*HTT* transcript or protein levels in HEK293T cells or the ST*Hdh* striatal cell model system for HD. A**, Schematic representation of the ZKSCAN1 vector containing an expression cassette composed by two short, artificial introns and an internal multiple cloning site. Linear sequences from exons 2-6 from human and mouse *HTT*/*Htt* respectively were cloned into the multiple cloning

site using EcoRV and SacII. Upon transfection the sequence becomes simultaneously over-expressed and circularized. **B**, Northern blot analysis of HEK293T cells over-expressing ZKSCAN1 *circHTT(2,3,4,5,6)* or empty vector (MCS) revealed a clear band of the expected size (484 nt) when incubated with an exon 5-targeting probe (24 hours exposure). **C**, *CircHTT(2,3,4,5,6)* over-expression in HEK293T cells upon transfection of the ZKSCAN1 vector does not alter the abundance of linear *HTT* mRNA (as assessed by qPCR analysis employing primers targeting exons 25, 41 or 51 respectively, *n*= 3 independent transfections). **D**, Western blot analysis of huntingtin protein levels in ZKSCAN1 *circHTT(2,3,4,5,6)* over-expressing HEK293T cells (left: representative blot, right: quantification, *n*= 3 independent transfections). **E**, *CircHtt(2,3,4,5,6)*/*Htt* ratio as assessed by RT-qPCR analysis of cDNA derived from total RNA of ST*Hdh* striatal cell lines from Q7/Q7 wild-type, Q7/Q111 heterozygous and Q111/Q111 homozygous knock-in mouse models for HD (*n*= 3 per genotype, levels were normalized against *Pgk1* and ratio between the circular and linear transcripts was calculated, one-way ANOVA with Sidak`s multiple comparisons testing, $P < 0.01$ =**, $P < 0.05$ =*, ns= not significant). **F**, RT-qPCR based assessment of *circHtt(2,3,4,5,6)* over-expression in polyclonal Q7/Q7, Q7/Q111 and Q111/Q111 ST*Hdh* cells (*n*= 3, transcript levels normalized on *Actb* and MCS empty vector control); **G**, Huntingtin protein levels in *circHtt(2,3,4,5,6)* over-expressing polyclonal Q7/Q7, Q7/Q111 and Q111/Q111 ST*Hdh* cells compared to the MCS empty vector (left: representative blot, right: quantification, *n*= 3 biological replicates). **H**, Subcellular fractionation followed by RNA extraction, cDNA synthesis and RT-qPCR analysis of marker RNAs (*Actb*, *Gapdh*, cytosol; *Malat1* for nuclear enriched transcripts) revealed clear segregation of *circHtt(2,3,4,5,6)* to the cytosolic cell compartment at physiologic levels in the MCS empty vector polyclonal cells (left), as well as *circHtt(2,3,4,5,6)* over-expressing cells (right) (data presented as ratio cytoplasmic/nuclear fraction, *n*= 3 per genotype and condition); **I**, Subcellular fractionation analysis of marker RNAs (*ACTB*, *GAPDH*, cytosol; *MALAT1* for nuclear enriched transcripts) in human HEK293T cells, revealing clear cytoplasmic localization of *circHTT(2,3,4,5,6)* in physiologic conditions (data presented as ratio cytoplasmic/nuclear fraction, *n*= 3).

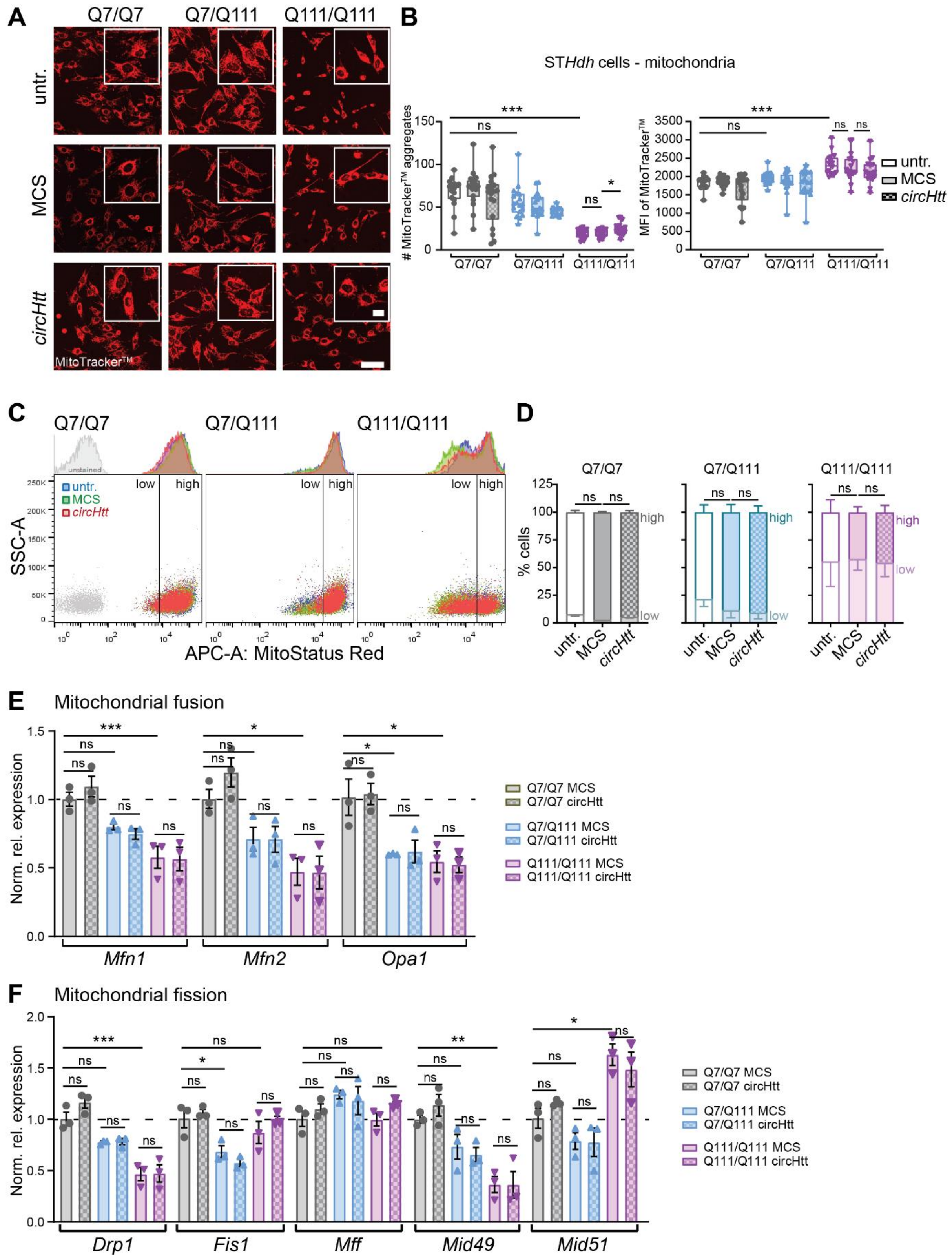


**Figure S5 - Mitochondrial morphology and function in the ST*Hdh* cells over-expressing *circHtt(2,3,4,5,6).*** **A**, Representative images of MitoTracker$^{TM}$ (red). **B**, Quantification of average

number of MitoTracker$^{TM}$ aggregates/cell (left) and mean fluorescent intensity (MFI) of MitoTracker$^{TM}$ signal per cell (right) (A,B, *n(cells)*= 8000-10000 cells per genotype and condition over three biological replicates, each dot represents the average value of all cells from the individual replica wells; outliers were removed using the ROUT (Q=1%) method, followed by one-way ANOVA with Sidak`s multiple comparisons testing (parametric data) and Kruskal-Wallis with Dunn`s multiple comparisons testing (nonparametric data), $P < 0.001$ =***, $P < 0.01$ =**, $P < 0.05$ =*, scale bars indicate 100 µm in overview and 25 µm in close ups); **C-D**, MitoStatus Red - Flow Cytometry, representative graph illustrating the intensity distribution of cells stained with MitoStatus, i.e. for ST*Hdh* Q7/Q7, ST*Hdh* Q7/Q111 and ST*Hdh* Q111/Q111; (unstained negative control in grey, stained untransfected cells in blue, stained MCS empty control in green, and stained cells overexpressing ZKSCAN1 *circHtt(2,3,4,5,6)* in red; each quadrant, corresponding to a specific genotype, is delineated by a gate separating low-intensity cell populations and high-intensity cell populations) (C). Quantification of percentage of cells belonging to these two distinct populations out of a total of 10,000 events detected (*n*= 4 biologic replicates, two-way ANOVA with Sidak`s multiple comparisons testing, ns= not significant); **E**,**F**, RT-qPCR based assessment of gene expression of regulators of mitochondrial fusion, i.e. *Mfn1*, *Mfn2* and *Opa1* (E) and fission, i.e. *Drp1*, *Fis1*, *Mff*, *Mid49*, *Mid51* (F) (*n*= 3, target gene expression level normalized on geometric mean of *Actb* and *Pgk1* levels and the Q7/7 MCS control; $2^{-DDCt}$ levels are reported, one-way ANOVA with Sidak`s multiple comparisons testing, $P < 000.1$=***, $P < 0.01$ =**, $P < 0.05$ =*, ns= not significant);

## Supplemental Tables:

**Table S1:** circRNAs stemming from the *HTT* locus

**Table S2:** Primers used in this study

**Table S3:** RBP-sites on *circHTT*