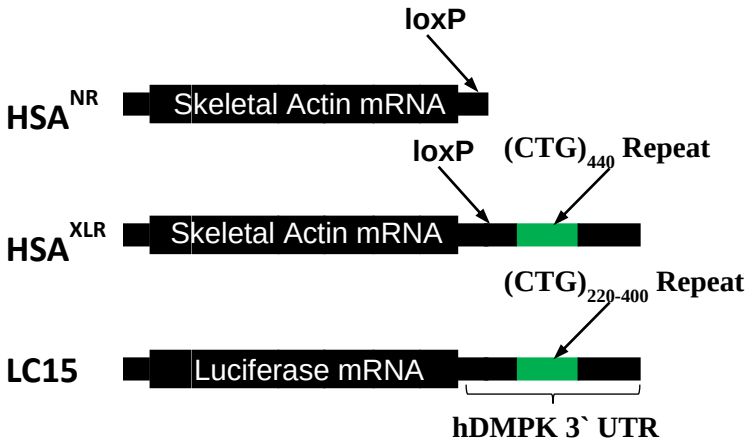
Supplemental Table 1

Table of exonic loci having ≥ 25 nt of CTG-CAG repeats expressed in human skeletal muscle. The STR Catalog Viewer (42) was used to examine sequence data from the 1000 Genomes Project (41). Shown are 16 loci having (1) ≥ 25 nt of CTG-CAG repeats for the major allele in at least 100 genomes; (2) > 90% purity of the repeat tract; (3) orientation of the repeat tract for expression of CUG repeats in an exon; and (4) absence of "A" nt interruptions (which are terminated by ddUTP). For repeats having bi-directional transcription, the listed gene is for the transcription unit expressing CUG-repeats. To estimate relative expression of CUG-repeats from each locus, the size of the major allele size was multiplied by reads per kilobase of transcript per million mapped reads from 430 muscle samples in GTEx data (43).

Chromosome	Size of major allele in 1000 genomes	Frequency of heterozygosity in 1000 genomes	Presence of "A" interruption	RPKM in skeletal muscle in GTEx	Relative CUG-repeat expression	CUG expression for exonic repeats	Gene
chr1	32	0.0489	+	1.2	1.2	38.4	PTBP2
chr2	31	0.1627		0.03	0.03	0.93	PDCD1
chr2	26	0.344		0.01	0.01	0.26	GALNT5
chr6	29	0.7029		3.4	3.4	98.6	NOTCH4
chr6	27	0.2087		3.4	3.4	91.8	CNPY3
chr7	34	0.0827		52	52	1768	BZW2
chr7	25	0.1853		4.5	4.5	112.5	BPGM
chr8	25	0.2271		3.8	3.8	95	ZNF395
chr8	25	0.1067	+	0.1	0.1	2.5	PXDNL
chr11	26	0.4082		0.6	0.6	15.6	DCHS1
chr16	47	0.6504		0.02	0.02	0.94	JPH3
chr17	25	0.1053		1.9	1.9	47.5	GHDC
chr18	25	0.422		0.4	0.4	10	LDLRAD4
chr19	17	0.6063		42	42	714	DMPK
chr19	25	0.5428		9	9	225	FCGRT
chr20	23	0.4621		16.4	16.4	377.2	CTSA
chrX	58	0.7303		0.7	0.7	40.6	CASK

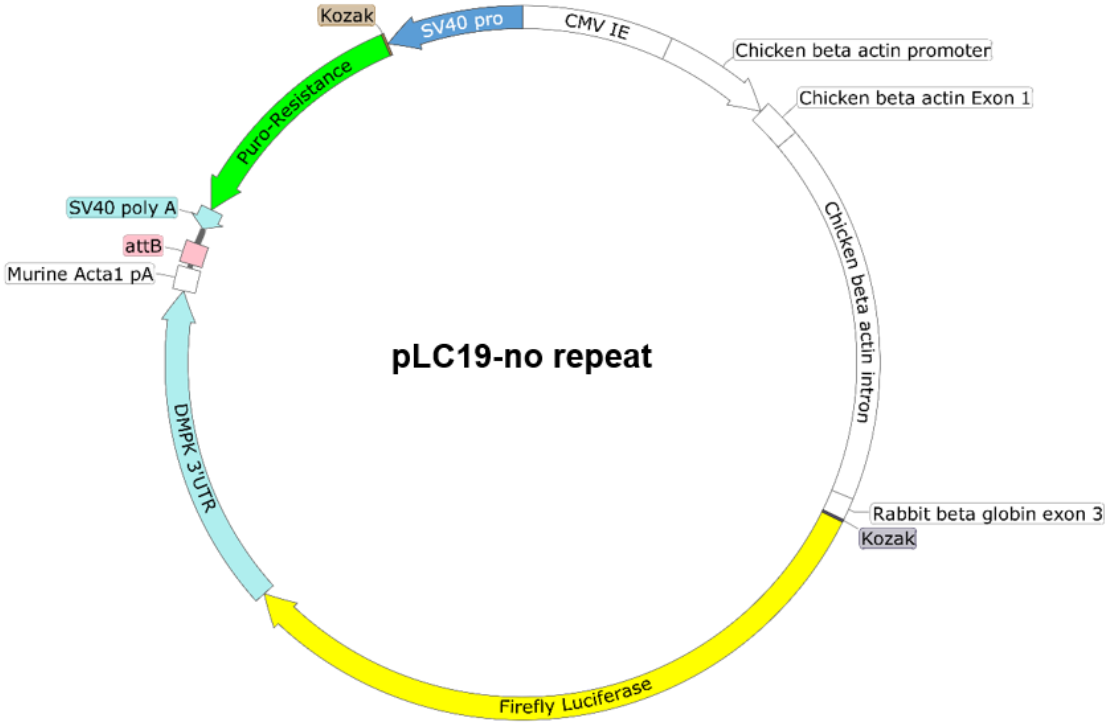
Supplemental Figure 1.

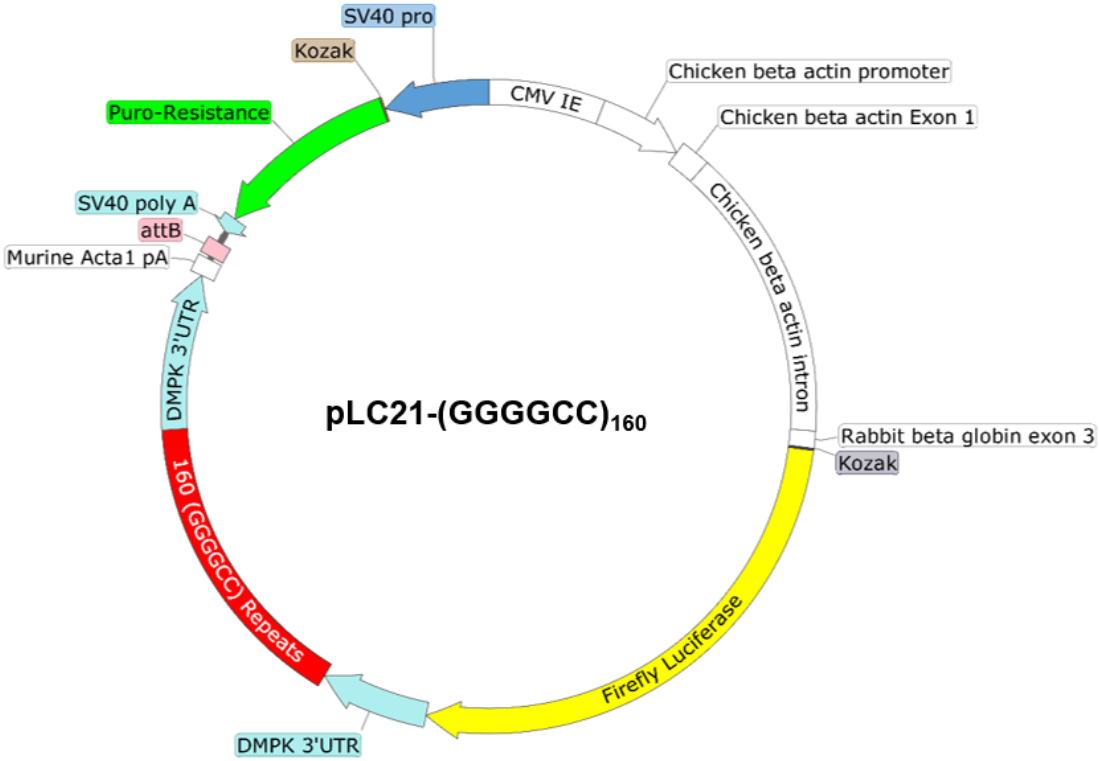
Diagram of transcripts produced by HSA^{NR}, HSA^{XLR}, and LC15 transgenic mice. Expression of HSA^{NR} and HSA^{XLR} is controlled by hACTA1 promoter. HSA^{NR} transgene produces hACTA1 mRNA with a truncated 3` untranslated region (UTR). $HSA^{XLR}$ produces full-length hDMPK with a 3` UTR containing a $(CTG)_{440}$ repeat (green bar). The LC15 transgene consists of luciferase cDNA fused to the hDMPK 3` UTR containing a $(CTG)_{220\text{-}400}$ repeat (green bar).



Supplemental Figure 2.

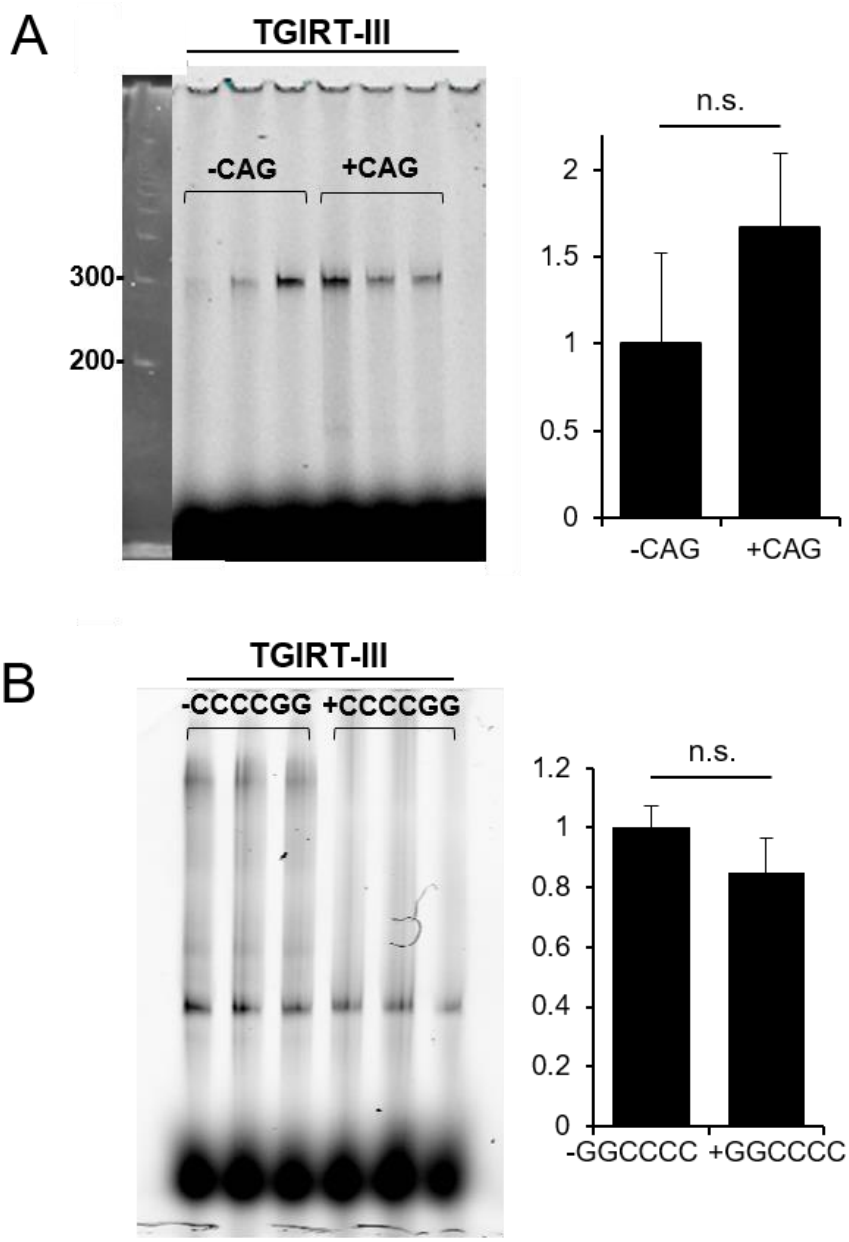
Diagrams of minigene constructs LC19-no repeat and LC21-$(GGGGCC)_{160}$. Expression of luciferase is driven by the CMV-enhancer and chicken beta-actin promoter. The luciferase cDNA is fused to the 3` UTR of *DMPK* containing no repeats or $(GGGGCC)_{160}$ repeats. Stable transfection of N2a neuroblastoma cells was obtained by cotransfection with plasmid expressing phi C31 integrase.


Kozak
SV40 pro
CMV IE
Chicken beta actin promoter
Chicken beta actin Exon 1
Puro-Resistance
SV40 poly A
attB
Murine Acta1 pA
Chicken beta actin intron
pLC19-no repeat
DMPK 3'UTR
Rabbit beta globin exon 3
Kozak
Firefly Luciferase


SV40 pro
Kozak
CMV IE
Chicken beta actin promoter
Chicken beta actin Exon 1
Puro-Resistance
SV40 poly A
attB
Murine Acta1 pA
Chicken beta actin intron
DMPK 3'UTR
pLC21-(GGGGCC)160
Rabbit beta globin exon 3
Kozak
160 (GGGGCC) Repeats
Firefly Luciferase
DMPK 3'UTR

Supplemental Figure 3.

Polyacrylamide TBE-urea gel electrophoresis analysis cDNAs generated by TGIRT-III using $r(CUG)_{100}$ or $r(GGGGCC)_{40}$ templates. (**A**) Reverse transcription of $r(CUG)_{100}$ template was performed with end-labeled primer that anneals 3' of the repeat tract, in the presence (+CAG) or absence (-CAG) of RNA/DNA hybrid $(CAG)_{6+1}$ primers. Quantification of corresponding bands on the right. (**B**) As in (A), using $r(GGGGCC)_{40}$ template with (+CCCCGG) or without (-CCCCGG) RNA/DNA hybrid $(GGCCCC)_{3+3}$ primers.



Supplemental Figure 4.

Reverse transcription of r(CUG)$_{100}$ with RNA/DNA hybrid primers. Representative chromatograms of 3` FAM-labeled cDNA products generated by TGIRT-III or SS-III using 5`-cagcagcagcagcagCAGC-3` (rna/DNA) primer and FAM-labeled ddUTP.