Supporting Information

O'Donnell et al. 10.1073/pnas.1302504110

SI Methods

Sequence Analysis of Tetracycline-*ORFeus* Insertions. Sequence reads were processed with phred and the fasta files were first aligned to the genome with blastall version 2.2.22 (mouse genome build mm9, 37.1) (1–3). All files containing mouse genomic DNA (gDNA) were aligned to each other to determine whether the entire construct had been sequenced and finally to the two inverse PCR (iPCR) primers. If the sequences overlapped in the expected configuration, the insertion site was determined. Perl scripts are available upon request.

Quantitative Real-Time PCR. All quantitative real-time PCR analysis was performed using the Step One Plus Real Time PCR system (Life Technologies). Analysis of ORF2 mRNA levels in tissues was performed by SYBR green analysis of cDNA samples. Primers used for ORF2 were forward (5′-TCGGCAAGGAGGAAGTGAAGATCAG-3′) and reverse (5′-GCTCTTGTTGCTGTTGATCTTGTAG-3′); *Actin* forward (5′-CGGTTCCGATGCCCTGAGGCTCTT-3′) and reverse (5′-CGTCACACTTCATGATGGAATTGA-3′). ORF2 expression was normalized to actin and represented relative to the animal treated with doxycycline during embryogenesis for each tissue. All SYBR green PCR assays were performed in triplicate.

Blood Isolation. Submandibular bleeding of mice was performed as previously described using a sterile lancet (4). Blood was harvested with EDTA K microvette tubes (Sarstedt) and immediately processed to isolate genomic DNA with the DNeasy Blood and Tissue Kit (Qiagen).

1. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410.
2. Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* 8(3):186–194.
3. Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 8(3):175–185.
4. Golde WT, Gollobin P, Rodriguez LL (2005) A rapid, simple, and humane method for submandibular bleeding of mice using a lancet. *Lab Anim (NY)* 34(9):39–43.

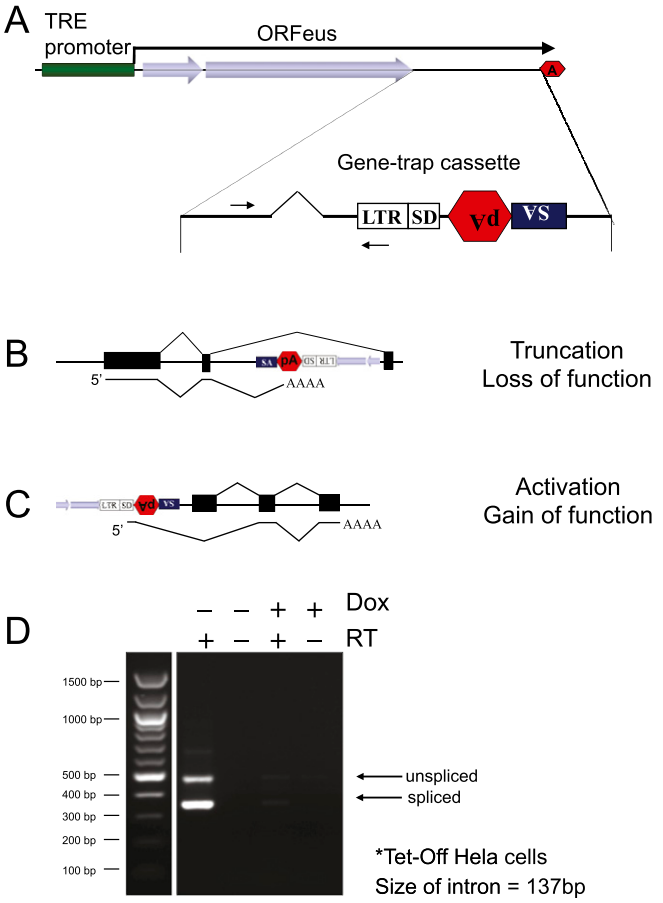

Fig. S1. Design of the tet-*ORFeus* gene trap transgene. (*A*) A mutagenic gene trap cassette is inserted in the 3′ untranslated region of the tet-*ORFeus* element. A chimeric intron (inverted "V") and intron flanking primers (horizontal arrows) used to detect intron splicing are indicated. Red stop signs indicate polyadenylation signals; the one in the sense orientation is used to define the 3′ end of the *ORFeus* transcript and the inverted one is part of the gene trap. LTR, long terminal repeat promoter from MSCV; SA, splice acceptor; SD, splice donor; pA, polyadenylation signal. (*B*) Schematic of an *ORFeus* gene trap insertion into an intron that results in mRNA truncation. (*C*) An *ORFeus* gene trap insertion upstream of a gene. Transcription from the MSCV 5′-LTR and splicing of the splice donor to a downstream exon is depicted, resulting in a gain of function allele. Note that a similar effect can be obtained by insertion into an intron; the result would be overexpression of a protein fragment, which might have a dominant negative phenotype. (*D*) RT-PCR analysis of the tet-*ORFeus* gene trap transgene 72 h after transfection into tet-OFF Hela cells and treatment (+) or absence (-) of 1 mg/ml doxycycline. Intron flanking primers were used to assay for intron splicing. Size ladder and PCR products were run on the same gel but irrelevant lanes were removed for clarity (denoted by white space between ladder and gel).

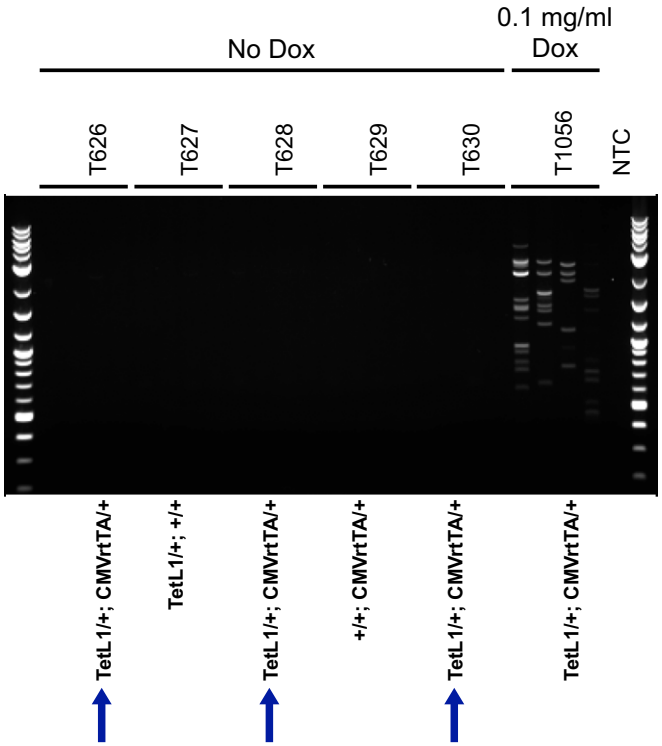

Fig. S2. No evidence for leaky tet-*ORFeus* transgene expression in vivo. Inverse PCR assay with tail genomic DNA isolated from line 058 animals that never received doxycycline (T626-T630). T1056, which received 0.1 mg/mL doxycycline during embryogenesis, serves as a positive control for the iPCR assay. Genotypes are labeled below the gel image. Blue arrows indicate double-transgenic animals lacking tet-*ORFeus* insertions.

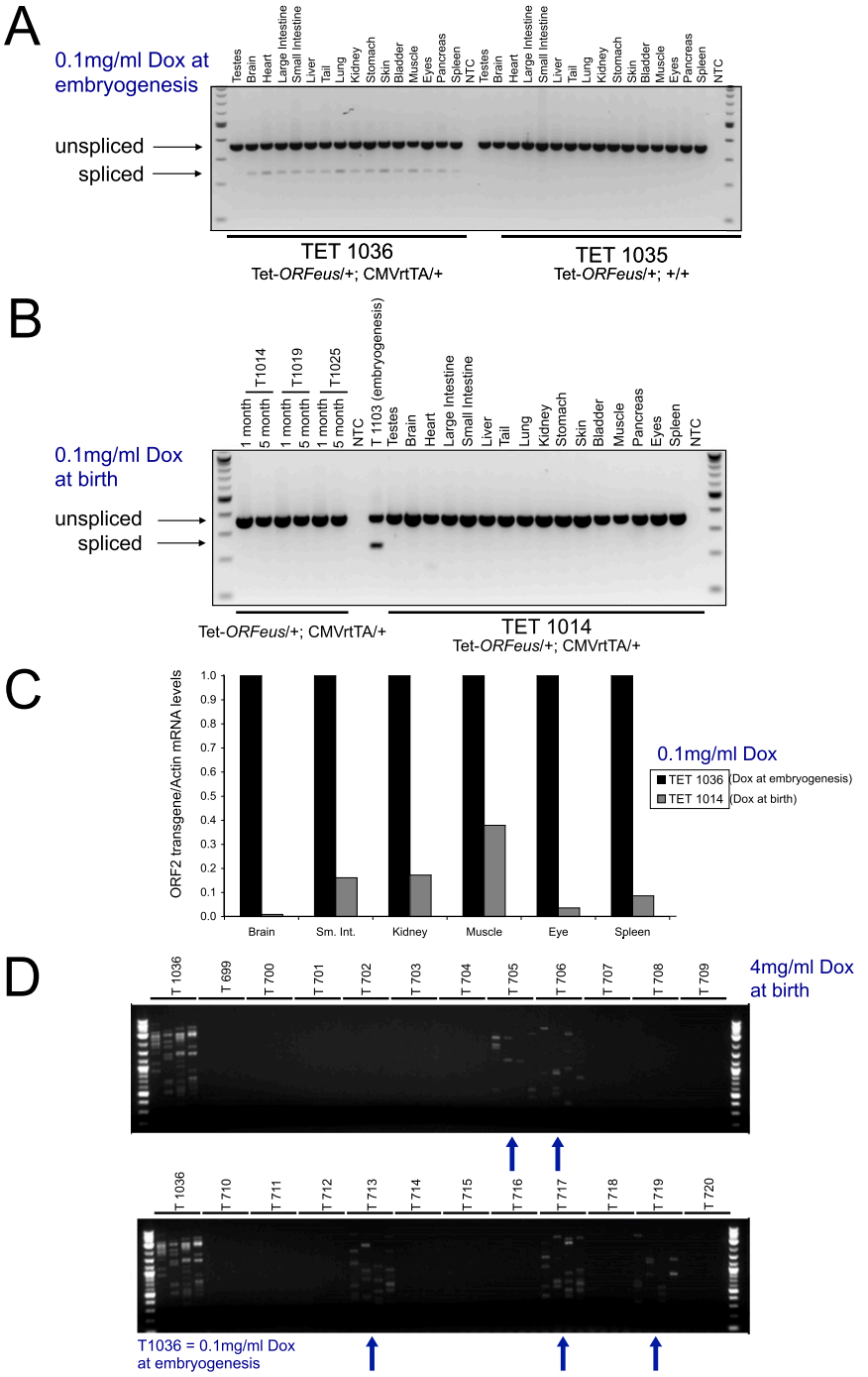

Fig. S3. Retrotransposition is less efficient when mice are treated with doxycycline postnatally compared with doxycycline treatment during embryogenesis. (*A*) Gene-trap splicing assay with gDNA isolated from double-transgenic (T1036) or single-transgenic (T1035) tissues when animals were administered doxycycline during embryogenesis. (*B*) Gene-trap splicing assay with gDNA isolated from tissues of double-transgenic animals given 0.1 mg/mL doxycycline at birth. Tail snips were isolated at 1 and 5 mo for double-transgenic mice T1014, T1019, and T1025. T1103, which was treated with doxycycline during embryogenesis, serves as a positive control. (*C*) Quantitative real-time PCR analysis of ORF2 mRNA relative to actin mRNA in tissues from a double-transgenic animal treated with doxycycline during embryogenesis (T1036) and a double-transgenic animal treated with doxycycline at birth (T1014). Data were normalized to the values obtained with T1036. (*D*) iPCR assay with tail gDNA isolated from animals given 4 mg/mL doxycycline at birth. T1036 received 0.1 mg/mL doxycycline during embryogenesis. Blue arrows highlight double-transgenic animals with tet-*ORFeus* insertions. T705 and T706 are line 058 animals; T713 is a line 095 animal; T717 and T719 are line 185 animals.

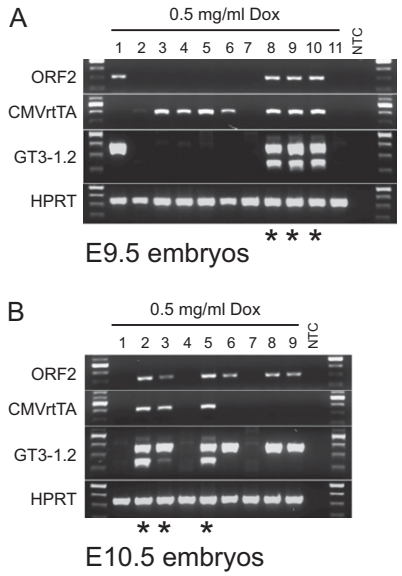

Fig. S4. Retrotransposition in E9.5–E10.5 embryos. Genotyping PCR (ORF2 and CMVrtTA) and gene-trap splicing assay (GT3-1.2) with gDNA isolated from E9.5 (*A*) and E10.5 (*B*) embryos. Animals were treated with 0.5 mg/mL doxycycline beginning at conception. Numbers represent individual embryos. Asterisks below the gel image indicate double-transgenic embryos. Hypoxanthine phosphoribosyltransferase 1 (HPRT) serves as a control for equal loading of template gDNA.

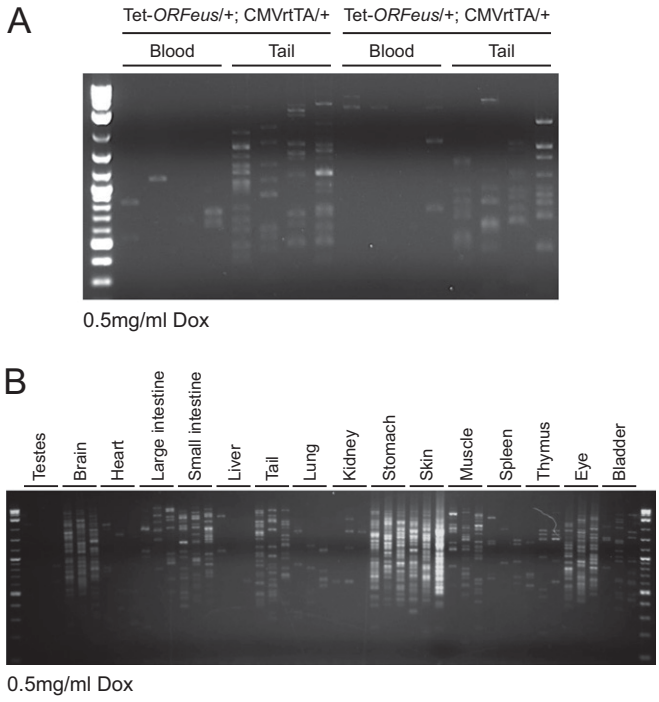

Fig. S5. Analysis of *ORFeus* insertions in blood and tail (*A*) and a panel of tissues (*B*) in double-transgenic line 095 animals treated with 0.5 mg/mL doxycycline during embryogenesis. Three to four independent iPCR assays were performed for each sample.

Table S1. Copy number of tet-*ORFeus* transgenic lines

Line ID no.	Approximate copy no.*	No. of integration loci†	Retrotransposition activity‡
TET 058	50	1	Yes
TET 095	1	1	Yes
TET 185	10	1	Yes
TET 161	15	1	No
TET 162	5	1	No

*Copy number was determined by quantitative Southern blotting.
†No. of integration loci was determined by Southern blotting.
‡Retrotransposition activity was determined by the presence of the "spliced" band in the intron-spanning PCR assay and by the iPCR assay.

Table S2. Locations of cloned tet-*ORFeus* insertions

Chromosome	Integration site*	Orientation†	Position to gene	Gene name	Gene description
X	3770631	–	Intergenic	Gm3706	Predicted gene 3706
X	13531317	+	Intergenic	Cask	Calcium/calmodulin-dependent serine protein kinase
X	15378435	+	Intergenic	AL732400.1	microRNA
X	30142600	+	Intergenic	Gm2777	Predicted gene 2777
X	41543038	–	Intergenic	Dcaf12l2	DDB1- and CUL4-associated factor 12-like 2
X	51674702	–	Intergenic	Gm14596	X-linked lymphocyte-regulated protein PM1-like
X	51856780	–	Intron 1	Gm16430	Predicted gene 16430
X	52110455	+	Intergenic	Slxl1	Slx-like 1
X	52111695	–	Intergenic	Gm14590	X-linked lymphocyte-regulated protein PM1-like isoform 2
X	58841094	–	Intergenic	Ldoc1	Leucine zipper, down-regulated in cancer 1
X	104471306	+	Intron 3	Gpr174	G-protein–coupled receptor 174
1	7263236	+	Intergenic	Gm16899	LincRNA gene
1	10523786	–	Intron 2	Cpa6	Carboxypeptidase A6
1	12405021	–	Intergenic	Gm17644	LincRNA gene
1	141545449	–	Intron 4	Zcchc16	Zinc-finger CCHC domain containing 16
1	147213431	+	Intergenic	LOC383570	—
1	4996691	–	Intron 2	Rgs20	Regulator of G protein signaling 20
2	44017964	–	Intron 10	Arhgap15	Rho GTPase-activating protein 15
2	67472782	+	Intergenic	Xirp2	Xin actin-binding repeat containing 2
2	70620018	+	Intron 2	Tlk1	Tousled-like kinase 1
2	109554330	+	Intron 1	Bdnf	Brain-derived neurotrophic factor, isoform 2
2	125700808	–	Intron 1	Galk2	*N*-acetylgalactosamine kinase
3	19867094	+	Intron 3	Cp	Ceruloplasmin isoform b precursor
3	117851146	+	Intergenic	Snx7	Sorting nexin 7
4	109231345	–	Intergenic	4930522H14Rik	RIKEN cDNA 4930522H14 gene
4	117023793	+	Intron 2	Rnf220	Ring-finger protein 220
4	15498869	–	Intergenic	LOC100040095	—
4	27750203	+	Intergenic	Epha7	Eph receptor A7
4	71396833	–	Intergenic	Tle1	Transducin-like enhancer of split 1
4	23401989	+	Intergenic	Gm11884	LincRNA gene
4	9428016	–	Intron 20	Asph	Aspartyl β-hydroxylase isoform 1
4	72963524	–	Intergenic	LOC433719	Uncharacterized protein
5	19211678	+	Intron 2	Magi2	Membrane-associated guanylate kinase, WW and PDZ domain-containing 2
5	43493091	–	Intergenic	Cpeb2	Cytoplasmic polyadenylation element-binding protein 2
5	84058498	+	Intergenic	Tecrl	Trans-2,3-enoyl-CoA reductase-like
5	149943955	–	Intron 2	5730422E09Rik	LincRNA gene
6	62931348	+	Intergenic	Arap2	ArfGAP with RhoGAP domain, ankyrin repeat, and PH domain 2
6	104207079	–	Intergenic	Cntn6	Contactin 6
6	12763776	–	Intergenic	Thsd7a	Thrombospondin, type I, domain-containing 7A
6	18112294	+	Intergenic	Cftr	Cystic fibrosis transmembrane conductance regulator
6	39654832	–	Intron 1	Braf	Braf-transforming gene
6	69996997	+	Intergenic	Igkv7-33	Ig kappa chain variable 7–33
7	74312280	+	Intergenic	Adamsts17	A disintegrin-like and metallopeptidase with thrombospondin type 1 motif
7	103463658	–	Intron 1	Odz4	Odd Oz/ten-m homolog 4
7	111576512	+	Intron 4	Trim30	Tripartite motif, protein 30

Table S2. Cont.

Chromosome	Integration site*	Orientation†	Position to gene	Gene name	Gene description
7	137533699	–	Intergenic	Ate1	Arginine-tRNA-protein transferase 1 isoform 1
7	5126969	+	Intergenic	Vmn1r56	Vomeronasal 1 receptor 56
7	91751276	–	Intron 1	Fah	Fumarylacetoacetate hydrolase
7	94873398	+	Intron 1	Grm5	Glutamate receptor, metabotropic 5
8	12363789	+	Intergenic	Gm5607	Predicted gene 5607
8	128728068	–	Intergenic	Slc35f3	Solute carrier family 35, member F3
9	6797752	–	Intergenic	Dync2h1	Dynein cytoplasmic 2 heavy chain 1
10	123125763	+	Intron 3	Fam19a2	Family with sequence similarity 19, member A2
10	71760018	–	Intergenic	Zwint	ZW10 interactor
12	7768155	–	Intergenic	Apob	Apolipoprotein B
12	78291911	+	Intergenic	Fut8	Fucosyltransferase 8
13	21329494	+	Intergenic	Trnaa-agc	Transfer RNA alanine
14	3159259	+	Intergenic	1700110I01Rik	RIKEN cDNA 1700110I01 gene
14	5413413	–	Intron 6	Gm3383	Predicted gene 3383
14	6300145	–	Intergenic	Gm3476	Predicted gene 3476
14	30426474	–	Intron 2	Cacna2d3	Calcium channel, voltage-dependent, α2 delta
14	106730687	+	Intergenic	Trim52	Tripartite motif-containing 52
15	62155742	–	Intergenic	H2afy3	H2A histone family, member Y3, noncoding RNA
16	12385854	–	Intergenic	Shisa9	Shisa homolog 9
16	20060381	–	Intergenic	Klhl24	Kelch-like 24
16	58359354	+	Intergenic	4930461C15Rik	LincRNA gene
16	84534838	+	Intergenic	Mir155	microRNA 155
16	89021359	+	Intergenic	Krtap6-1	Keratin-associated protein 6–1
18	29241510	–	Intergenic	4930474G06Rik	RIKEN cDNA 4930474G06 gene
19	24017501	+	Intron 10	Apba1	Amyloid β(A4) precursor protein binding

LincRNA, long intergenic noncoding RNA.

*The coordinate listed for each integration site is the first non-A base following the 3′ end of the insertion or the first base located some distance (typically 1 kb) downstream from the integration site. This is because the polyA/3′ flanking sequence was not reached for a number of iPCR products. The coordinates are derived from build mm9, 37.1 of the mouse reference genome sequence.

†The orientation is the direction of tet-*ORFeus* transcription relative to the reference sequence.

Table S3. PCR cycling conditions for genotyping assays

Number of cycles	ORF2	GT3-1.2	HPRT	CMVrtTA
1×	95 °C, 2 min	95 2 min	95 °C, 2 min	
35×	95 °C, 30 s	95 °C, 30 s	95 °C, 30 s	
	60 °C, 30 s	55 °C, 30 s	57.5 °C, 30 s	
	72 °C, 30 s	72 °C, 30 s	72 °C, 30 s	
1×	72 °C, 10 min	72 °C, 10 min	72 °C, 10 min	
	4 ∞	4 ∞	4 ∞	
1×				94 °C, 2 min
30×				94 °C, 30 s
				57 °C, 30 s
				72 °C, 30 s
1×				72 °C, 10 min
				4 ∞

Amplicon sizes: ORF2 = 307 bp; GT3-1.2 = 369 bp (unspliced), 232 bp (spliced); HPRT = 340 bp; and CMVrtTA = 400 bp. min = minutes, s = seconds.

Table S4. PCR reaction conditions for Tet-*ORFeus* genotyping (in μL)

Components	1×
ORF2, GT3-1.2, HPRT	
H_2O	17.35
10× buffer	2.5
dNTP mix	2.0
10 μM forward primer	1.0
10 μM reverse primer	1.0
Ex Taq polymerase	0.15
DNA template	1.0
Total	25
CMVrtTA	
H_2O	17
DMSO	1.5
10× buffer	2.5
dNTP Mix	2.5
100 μM forward primer	0.25
100 μM reverse primer	0.25
Ex Taq polymerase	0.15
DNA template	1.0
Total	25.15

Amplicon sizes: ORF2 = 307 bp; GT3-1.2 = 369 bp (unspliced), 232 bp (spliced); HPRT = 340 bp; and CMVrtTA = 400 bp.

Table S5. Tet-*ORFeus* genotyping primers

Primer name	Primer sequence	bp
smORF2 forward primer	AAGGAGGAAGTGAAGATCAGCCTGT	307
smORF2 reverse primer	TCCTTGATCTCCTTCTTCAGGCTCT	
GT3-1.2 forward primer	CTAGTCCTGCAGGCCAAAAT	369 (unspliced)
GT3-1.2 reverse primer	TCTGGGGACCATCTGTTCTT	232 (spliced)
HPRT control forward primer	CTTTCTTGTCACTCCCACTTTTCC	340
HPRT control reverse primer	CACATCCTTCATTCAGGTGTCACT	
CMVrtTA forward primer	GTGAAGTGGGTCCGCGTACAG	400
CMVrtTA reverse primer	GTACTCGTCAATTCCAAGGGCATCG	

Table S6. Cycling conditions for iPCR

Number of cycles	iPCR
1×	94 °C, 2 min
10×	94 °C, 15 s
	70 (−0.5 °C/cycle), 30 s
	72 °C, 60 s
25×	94 °C, 15 s
	65 °C, 30 s
	72 °C, 60 s
1×	72 °C, 7 min
	4 ∞

Table S7. Reaction conditions for iPCR (in μL)

Components	1×
H_2O	34.95
10x Buffer	5.0
dNTP Mix	5.0
10 μM Forward primer	1.9
10 μM Reverse primer	1.9
Ex Taq HS polymerase	0.25
DNA (from conc. ligation rxn)	1.0
Total	50

Conc., concentrated; rxn, reaction.