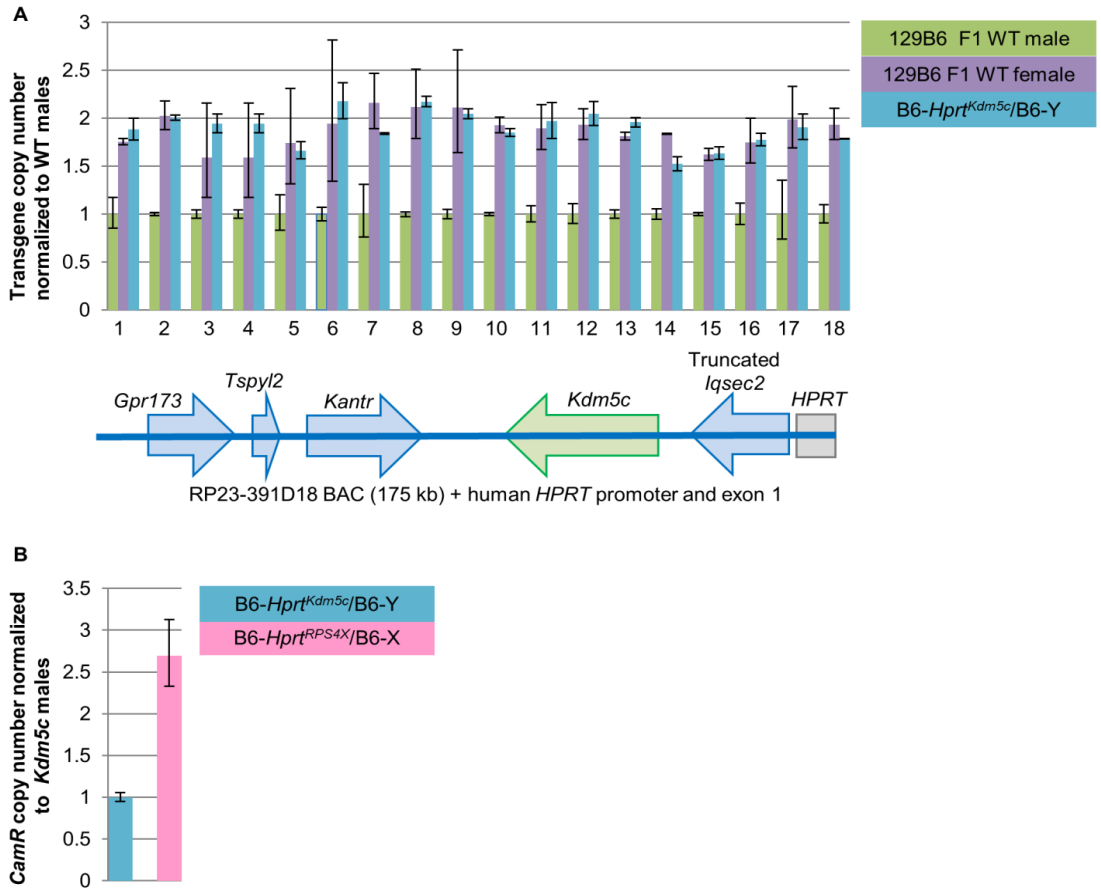
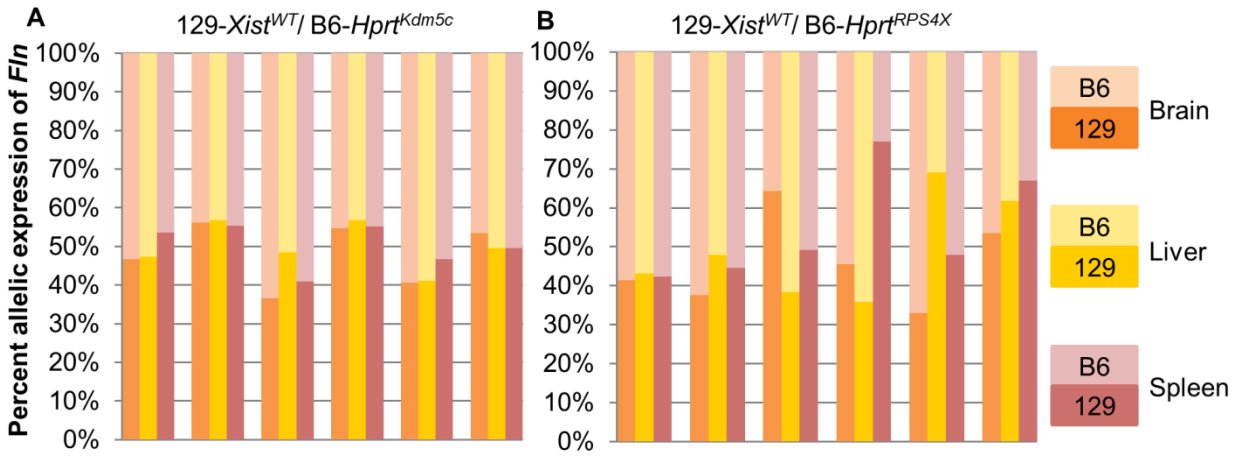
Human *cis*-acting elements regulating escape from X-chromosome inactivation function in mouse

Samantha B Peeters, Andrea J Korecki, Elizabeth M Simpson, and Carolyn J Brown

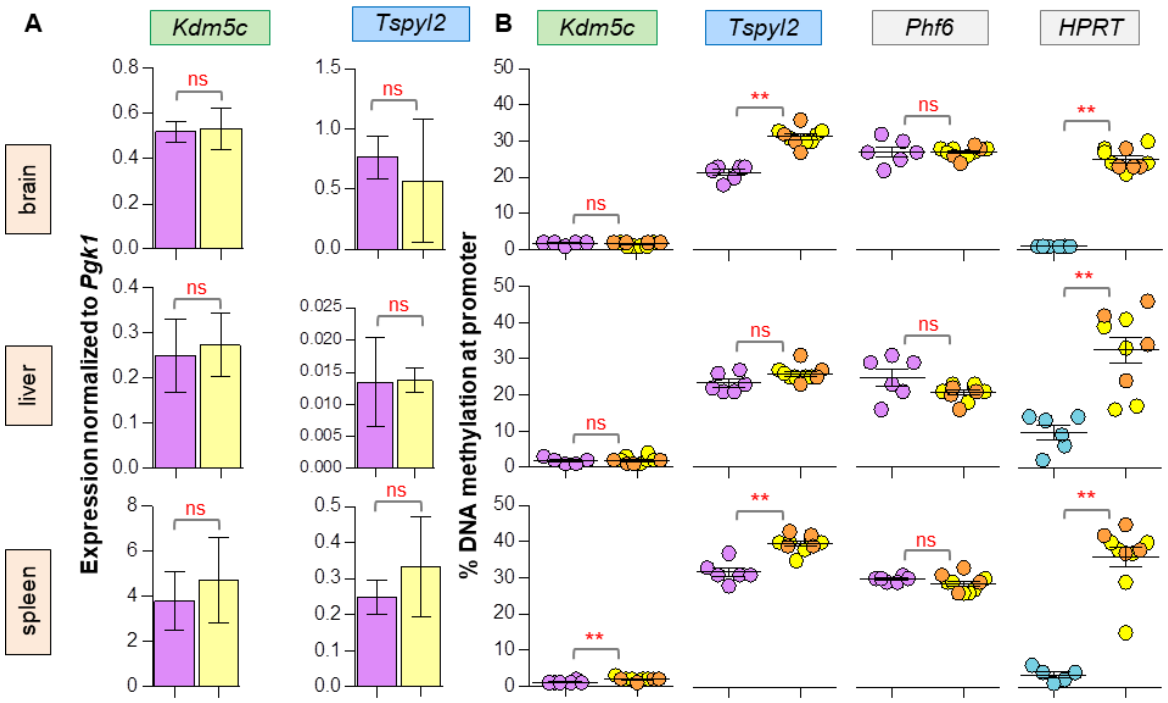
Supplementary Information



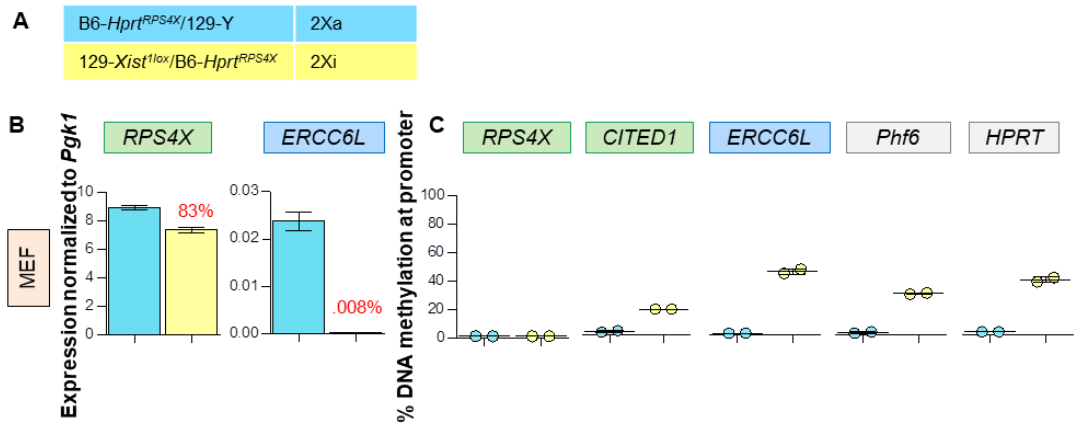
S1 Fig. Copy number analyses suggest single-copy integration of RP23-391D18 and double integration of RP11-1145H7. A) 18 qPCR assays approximately every 10 kb along the *Kdm5c*-containing BAC were generated to test for copy number of the integration. N2 transgenic males with an endogenous and transgenic copy on their Xa were compared to wild-type males with one endogenous copy, and wild-type females with 2 endogenous copies. The N2 transgenic males more closely resembled wild-type females with 2 copies suggesting that indeed only one copy of the transgene had integrated. Each bar is an average of 2 mice per genotype. **B)** N1 females with the *RPS4X* transgene were tested for copy number by comparing the copy number of a gene, *CamR*, in the construct that is common between both the *RPS4X* and *Kdm5c* BACs. As we had determined the *Kdm5c* transgene to be single-copy in A), we normalized 6 *RPS4X* females to 2 *Kdm5c* males which indicated at least 2 copies of the *RPS4X* transgene had been integrated.



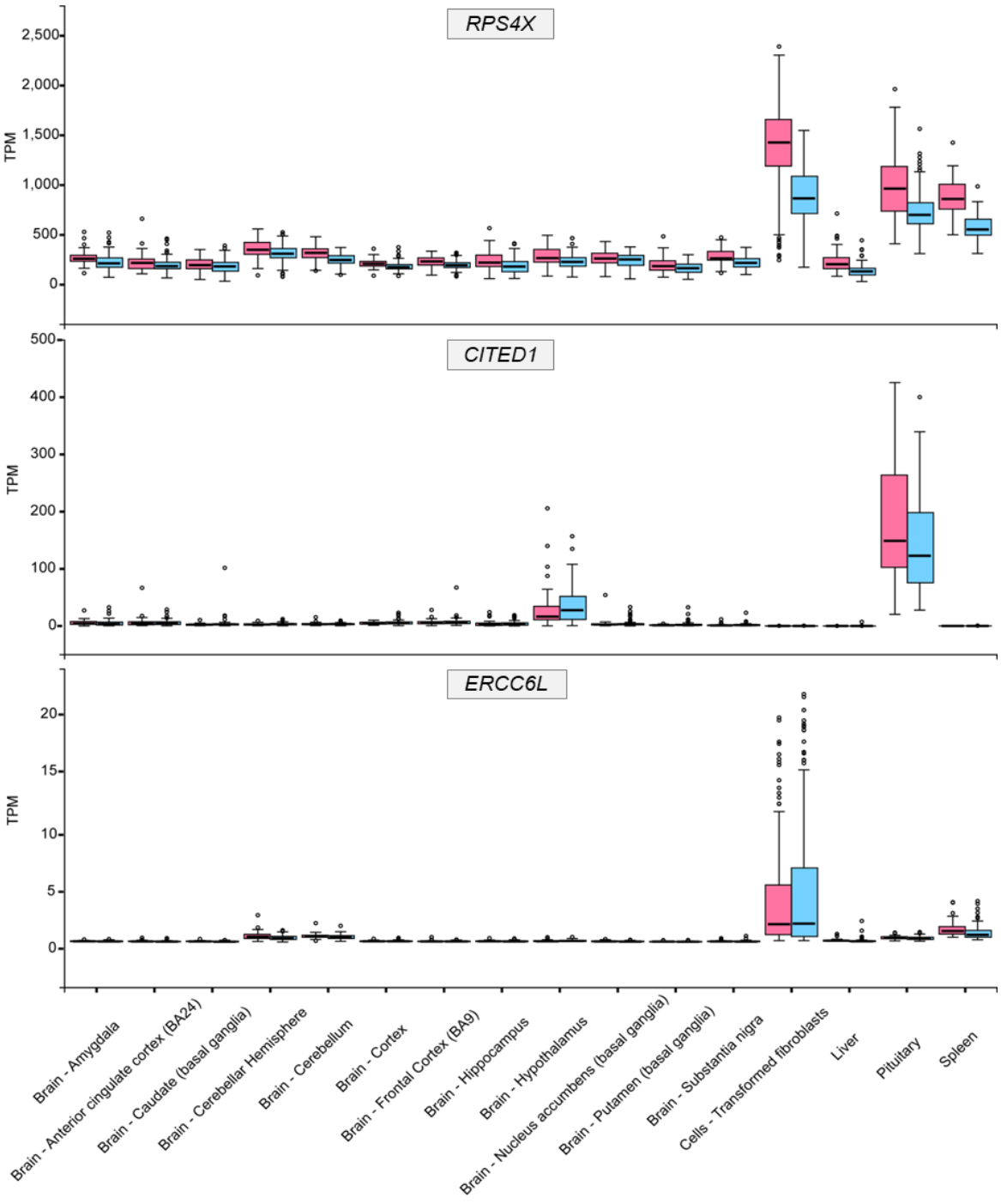
S2 Fig. *Xist*^{WT} females do not have skewed XCI. Percent allelic expression of X-linked gene *Fln* in **A)** six female mice carrying *Kdm5c* transgene and **B)** six female mice carrying *RPS4X* transgene, all with wild-type *Xist* and random XCI. All three tissues are shown for each mouse with the percent expression indicating how often the B6 or 129 X is the Xa.



S3 Fig. Analysis of additional 129-*Xist*^{1lox}/B6-*Hprt*^{Kdm5c} mice supports escape of *Kdm5c* by DNAm. A) Normalized to *Pgk1*, RT-qPCR of *Kdm5c* and *Tspyl2* expression in average of six original wild-type females (purple) compared to average of six original Xi knock-in females (analyzed in Fig 1) plus four additional Xi knock-in females (yellow). **B)** Average DNAm of *Kdm5c* shows promoter hypomethylation in four additional knock-in females (orange) consistent original six Xi knock-in females (yellow) also plotted. *Tspyl2*, *HPRT* and *Phf6* also shown (Mann-Whitney t-test, significance denoted by asterisks; p-value <0.001 ***, 0.001 to 0.01 **, 0.01 to 0.05 *, >0.05 ns). Knock-in females are compared to wild-type females for all assays except *HPRT* (compared to knock-in males, blue) as wild-type females do not carry the human gene.



S4 Fig. *ERCC6L* is subject to XCI in mouse embryonic fibroblasts (MEFs). A) Description of genotypes, n=2 for each. **C)** Normalized to *Pgk1*, RT-qPCR of *RPS4X* and *ERCC6L* expression shows that the transgenes are both active on a male X in MEFs. When on the Xi in MEFs, *RPS4X* escapes at high levels while *ERCC6L* lacks expression in these females indicating it is subject to XCI. Expression from female Xi is shown as percentage of the male X (red text). **D)** Average DNAm of skewed knock-in females shows a hypomethylated *RPS4X* promoter, *CITED1* in the uncallable range, and *ERCC6L, HPRT* and *Phf6* hypermethylated compared to knock-in males. While not statistically significant due to low sample size, the data is similar to other tissues examined and supports *ERCC6L* being subject to XCI in mouse.



S5 Fig. RP11-1145H7 region expression and DNAm in human. *RPS4X* is broadly expressed in all human tissues corresponding to the tissues we examined in mouse. *CITED1* and *ERCC6L* are limited in their expression for the tissues examined. Males shown in blue, females with random XCI in pink (GTEx Portal).

S1 Table. Experimental crosses yield normal sex and genotype ratios. 129-*Xist*^{1lox}/X females were crossed to B6-*Hprt*^{BAC}/Y males to generate F1 129-*Xist*^{1lox}/B6-*Hprt*^{BAC} and F1 129-*Xist*^{WT}/B6-*Hprt*^{BAC} experimental females. Neither male to female ratio nor female *Xist*^{WT} to *Xist*^{1lox} is significantly different from expected (chi-square test, 0.05 significance level) suggesting that the additional expression from the BAC genes in females, and specifically the *Xist*^{WT} mice, did not have a deleterious effect on breeding or significantly skew the ratios of offspring.

***Kdm5c* experimental cross**

	male	female	*Xist*^{WT}	*Xist*^{1lox}
Total mice	38	33	18	15
Chi-square		0.352		0.273
P-value		0.553		0.602

***RPS4X* experimental cross**

	male	female	*Xist*^{WT}	*Xist*^{1lox}
Total mice	71	54	26	25
Chi-square		2.312		0.02
P-value		0.128		0.889

S2 Table. Primer table.

Experiment	Assay	Sequence
BAC retrofit junction PCR	3'retrofit_F1	AAGCCTTCGCGAAAGAAAAT
	3'retrofit_R2	AACCTGGGTAACATGGTGAGA
	5'retrofit_F1	TCCTTAAGCCCCCACTAGGT
	5'retrofit_R1	TGTCCTTTGTTACAGGCCAGA
Hprt integration and BAC backbone vector	5'CamR oEMS4863	TGGCAATGAAAGACGGTGAGC
	5'CamR oEMS4864	GCATCAGCACCTTGTCGCCT
	Hprt Correction oEMS2267	TCAGGCGAACCTCTCGGCTT
	Hprt Correction oEMS2269	TGCTGGACATCCCTACTAACCCA
	Hprt WT oEMS2236	TCCAGGGAATCCATCACAAT
	Hprt WT oEMS2238	CCAGTGCTGACGTTACAAGC
	Hprt Null oEMS2236	TCCAGGGAATCCATCACAAT
	Hprt Null oEMS2240	GGCATCCAGTGCTCTTCACT
RP23-391D18 BAC copy-number qPCR	qRP23-391D18_F1	CTACTTCTGGGCAGGTGGTC
	qRP23-391D18_R1	GGAATGCCACTTCCTAGCCT
	qRP23-391D18_F2	CACCTGTCTTCACTTCCCCA
	qRP23-391D18_R2	CAGCCCATAAGCCAGGTGTA
	qRP23-391D18_F3	CAGACTCCCATGTGCAAGGA
	qRP23-391D18_R3	TGCTACACACATTCCTGACCA
	qRP23-391D18_F4	CGCTCCCTTTGTTGTCAGTG
	qRP23-391D18_R4	ACCGATGCCAATGAGAACCT
	qRP23-391D18_F5	TGGAGGTTAGTTGCCAGCAT
	qRP23-391D18_R5	CTGGAGGAGGAGTTACAGGC
	qRP23-391D18_F6	CTTTCACTGGTCTGGAGCGA
	qRP23-391D18_R6	CGGGGCTCAATAATGGCTTG
	qRP23-391D18_F7	CATAGCCAAACATGCCCCAG
	qRP23-391D18_R7	AGAGAGTACAAAGGCTGGCC
	qRP23-391D18_F8	CAGTCGGAGGAGTCTGTGTT
	qRP23-391D18_R8	CCCCATCCCAACCTGTTACA
	qRP23-391D18_F9	TCACCTCCTTCTTGAGCTCC
	qRP23-391D18_R9	CTGCCACTTTGCTGCTCA
	qRP23-391D18_F10	GTGTGAGTTTCCAGACCTGC
	qRP23-391D18_R10	ACTGTTCCATCATGTCCGCT
	qRP23-391D18_F11	GAGGACCAGGATGGCTTAGAA
	qRP23-391D18_R11	CCGGGAAGTCAGAGTAGAGAAA
	qRP23-391D18_F12	GATCCTGGCCACTTTCCTCA
	qRP23-391D18_R12	TCTGCAGGAAACGACCCAG
	qRP23-391D18_F13	AACTGTGGAGTATGGGGCTG

	qRP23-391D18_R13	GGAAACCGCTGCCAAATTCT
	qRP23-391D18_F14	GAGGAAGGAGCAGGGATGAG
	qRP23-391D18_R14	ACACCTTCCATCTGAACCCC
	qRP23-391D18_F15	AGCTGTACGGTGTTAGTGGT
	qRP23-391D18_R15	CAGATCCAACCTGCCTCTGT
	qRP23-391D18_F16	ACAGCTTCCGTCAGTCCTTT
	qRP23-391D18_R16	GCTAGGGTTCAAGAGGGGAC
	qRP23-391D18_F17	TCAGGGCCTATGTCTGAGGA
	qRP23-391D18_R17	TGTGGCTGAGCTGTCTTCTT
	qRP23-391D18_18F	GGGAGTGGGATACGAAGAGAA
	qRP23-391D18_18R	CCTTACTGTCCCTCCCTGAATA
	qHbb-bs_F (control)	CTGCTCACACAGGATAGAGAGGG
	qHbb-bs_R (control)	GCAAATGTGAGGAGCAACTGATC
	qCamR_F (BAC backbone)	TCCCAATGGCATCGTAAAGAA
	qCamR_R (BAC backbone)	CAGCTGAACGGTCTGGTTATAG
RP11-1145H7 BAC PCR	RP11-1145H7_F1	CAAACTCTTTCTGTAGCTTG
	RP11-1145H7_R1	TGAAGTCTCATTCTGTCATC
	RP11-1145H7_F2	CTTAACAATGGAGTTTGGAG
	RP11-1145H7_R2	CACAGTCTATCTTTGGATTG
	RP11-1145H7_F3	CTGTCCTCTCAACAAGAAC
	RP11-1145H7_R3	ACATCTGTTGTGTCTAATGC
	RP11-1145H7_F4	ATTGGGACGGTATCCAGTAAGA
	RP11-1145H7_R4	TGGGCTTGAGTCCCTGTAAT
	RP11-1145H7_F5	CAGATTGTAATCACTGAACC
	RP11-1145H7_R5	TTAGTATGGGGTTTCACC
	RP11-1145H7_F6	GCCTGTAATCTCAGCTACTC
	RP11-1145H7_R6	TGTCAGCTGCTTTCTATATC
	RP11-1145H7_F7	ATCATCTCTTTCCCTCATC
	RP11-1145H7_R7	TTTAAGACAGGGTTTCACTC
	RP11-1145H7_F8	ATAGAGTGGTGAACAAGATG
	RP11-1145H7_R8	CATCAACCTTCTGAGTAGC
	RP11-1145H7_F9	TCCTCAGACTAGAGAGAAGG
	RP11-1145H7_R9	AGTACATGTGAGATGGATTG
	RP11-1145H7_F10	CTGTAATCCCAGTTACTCAG
	RP11-1145H7_R10	ATAACAAGTGTTGGTGAGC
	RP11-1145H7_F11	CATATGATCCATCTTGGTC
	RP11-1145H7_R11	CTGTAATCTCAGCACTTTG
	RP11-1145H7_F12	TGAGTTAGAATCAAGACCAG
	RP11-1145H7_R12	TGAGTAGCTGAGACTACAGG
	RP11-1145H7_F13	CAAGAATTGGGTCTAGTTG

	RP11-1145H7_R13	TCTTTCCAGTCCTATATTCC
	RP11-1145H7_F14	CTCCTTGGCTAAGTTTATTC
	RP11-1145H7_R14	AACGATCTCTGTCTCAATG
	RP11-1145H7_F15	CATCTTATGAGTTGTGAAGC
	RP11-1145H7_R15	AATTTAACTGGAGAGTGAGG
	RP11-1145H7_F16	GGTCTTTGATATTGCTTGTC
	RP11-1145H7_R16	GTCCTGACTTTCTACTCTGC
	RP11-1145H7_F17	ACTCAAAGGTAGGAGAACTG
	RP11-1145H7_R17	AATCAGCTCTAAAGTGTTCC
DNA methylation PCR and pyrosequencing	DNAmHPRT_F	GGAATTAGGGAGTTTTTTGAATAGG
	DNAmHPRT_R	/5Biosg/CCTACCAATTTACAAACTCACTAAATA
	DNAmHPRT_S	GGGAGGGAAAGGGGT
	DNAmPhf6_F	/5Biosg/GTGGTTTTTTTTATTGTTAGGGATTTT
	DNAmPhf6_R	GAAATATTGGGATGGGGGTTTT
	DNAmPhf6_S	ATAGAGGTTGGYGATTT
	DNAmKdm5c_F	GTAAGGTTGGGAGTTGATGG
	DNAmKdm5c_R	/5Biosg/CCCATATTCTTCCCACACCTACTA
	DNAmKdm5c_S	GTAAGGTTGGGAGTTGA
	DNAmTspyl2_F	TGAGGGGTAGTTAGTTTGATGA
	DNAmTspyl2_R	/5Biosg/CTCAACCCCTACCTTCTCT
	DNAmTspyl2_S	GGGTAGTTAGTTTGATGATT
	DNAmRPS4X_F	ATTAGTAGATGGTAAGAAAGAGTT
	DNAmRPS4X_R	/5Biosg/CCCAACTCAACCCTTTACT
	DNAmRPS4X_S	AGATGGTAAGAAAGAGTTT
	DNAmERRC6L_F	GGGTAGTTTTTTTTATTTATAATGATGGTA
	DNAmERRC6L_R	/5Biosg/TTCAATCCAATTCAAACCATACTACA
	DNAmERRC6L_S	TTTTTTTTATTTATAATGATGGTAT
	DNAmCITED_F	AAGTGGAATTTATTGGGTAAGTT
	DNAmCITED_R	/5Biosg/CCTAACCAATACCCCACTTCTAAAACTATC
	DNAmCITED_S	GTGGAATTTATTGGGTAAGTTTA
SNP PCR and pyrosequencing	Fln_F	/5Biosg/CCAGCTTCCCTAGTCCAAATGC
	Fln_R	TGCATACAGTCAGTGTCAAGTACAAG
	Fln_S	CCTAGAGAGGGCTGAA
Expression RT-qPCR	qPgk1_F (control)	CGTCTGCCGCGCTGTT
	qPgk1_R (control)	AACACCGTGAGGTCGAAAGG
	qKdm5c_F	GCGACTGGGACTTAACTGTAG
	qKdm5c_R	TCCGTTTCTTCCACACCTTAC
	qTspyl2_F	GGATGACAAGGAGAGTGTGAGG
	qTspyl2_R	TCTGGATGAATTTGCGCTTGAG

	qRPS4X_F	CAAGGTCCGAACTGATATAACCTAC
	qRPS4X_R	GGAAATTCTCTCCCGTCTTGTC
	qCITED_F	GGCAGAATCACTCTCTCCTTCT
	qERCC6L_F	GGCTGAGGCTGTGGTTTATT
	qERCC6L_R	CCTGGCTAAGAGAACCTGTATTTC

All PCR and qPCR primers were designed using the PrimerQuest Tool (IDT) and tested for specificity by In-Silico PCR (UCSC). All pyrosequencing primers were designed using PyroMark Assay Design software (Qiagen).