Published online 3 November 2011

Nucleic Acids Research, 2012, Vol. 40, No. 4 ***1523–1535***
doi:10.1093/nar/gkr871

Tissue-specific alternative polyadenylation at the imprinted gene *Mest* regulates allelic usage at *Copg2*

Julia L. MacIsaac[1,2], Aaron B. Bogutz[1,2], A. Sorana Morrissy[2,3] and Louis Lefebvre[1,2,*]

[1]Molecular Epigenetics Group, Life Sciences Institute, [2]Department of Medical Genetics, The University of British Columbia, Vancouver, Canada, V6T 1Z3, and [3]Genome Sciences Centre, British Columbia Cancer Agency, Vancouver, Canada, V5Z 1L3

Received April 30, 2011; Revised September 27, 2011; Accepted September 29, 2011

ABSTRACT

The gene *Mest* (also known as *Peg1*) is regulated by genomic imprinting in the mouse and only the paternal allele is active for transcription. *MEST* is similarly imprinted in humans, where it is a candidate for the growth retardation Silver-Russell syndrome. The MEST protein belongs to an ancient family of hydrolases but its function is still unknown. It is highly conserved in vertebrates although imprinted expression is only observed in marsupials and eutherians, thus a recent evolutionary event. Here we describe the identification of new imprinted RNA products at the *Mest* locus, longer variants of the RNA, called *MestXL,* transcribed >10 kb into the downstream antisense gene *Copg2.* During development *MestXL* is produced exclusively in the developing central nervous system (CNS) by alternative polyadenylation. *Copg2* is biallelically expressed in the embryo except in *MestXL*-expressing tissues, where we observed preferential expression from the maternal allele. To analyze the function of the *MestXL* transcripts in *Copg2* regulation, we studied the effects of a targeted allele at *Mest* introducing a truncation in the mRNA. We show that both the formation of the *MestXL* isoforms and the allelic bias at *Copg2* are lost in the CNS of mutants embryos. Our results propose a new mechanism to regulate allelic usage in the mammalian genome, via tissue-specific alternative polyadenylation and transcriptional interference in sense–antisense pairs at imprinted loci.

INTRODUCTION

The proximal end of mouse chromosome 6 contains an imprinted domain associated with a growth retardation phenotype in maternal uniparental disomy (mUPD) mice (1,2). This ~220-kb domain contains the paternally expressed gene *Mest* (also known as *Peg1*) (3) and the maternally expressed gene *Klf14* (4). The promoter CpG island of *Mest* is differentially methylated in somatic cells (5). This DNA methylation mark is directly inherited from oocytes (6) constituting the only known differentially methylated region (DMR) and gametic imprint in the entire domain. Whereas DNA methylation at *Mest* is associated with silencing of the maternal allele (7), expression of the maternal *Klf14* was found to require maternally inherited methylation (4). A targeted mutation at *Mest*, the $Mest^{KO}$ allele (official name $Mest^{tm1Lef}$) established that *Mest*-deficient mice exhibit two main phenotypes: intrauterine growth restriction and an abnormal maternal behavior of mutant females (7). *Mest* codes for a highly conserved enzyme of the α/β-hydrolase fold superfamily, although the substrate and biological function of MEST are still unknown (3,8). This evolutionary conservation contrasts with the imprinting of the locus, which has only been observed in marsupials and placental mammals (9,10). It is plausible that selective pressures for imprinting at the locus followed acquisition of novel therian-specific functions of the locus. In accordance with this proposition, *Mest* is highly expressed in extra-embryonic mesoderm where it may play important roles in the establishment or function of the placental vascular network (11).

We considered the possibility that imprinted transcription at *Mest* might be involved in alternative functions, independent of its coding potential. For instance, the *Mest* locus contains the microRNA *Mir335*, produced from intron 2, which is implicated in the suppression of human breast cancer metastasis (12). Another striking feature of the genomic organization of *Mest* is its overlap with the neighboring antisense gene *Copg2* (13). The imprinted status of *Copg2* itself is unclear in both mouse and human. In human, where the ortholog *MEST* is also paternally expressed in some tissues (14),

*To whom correspondence should be addressed. Tel: +1 604 822 5310; Fax: +1 604 822 5348; Email: louis.lefebvre@ubc.ca

the allelic usage at *COPG2* is disputed in the literature with a report of paternal expression (15) and another of biallelic expression and no imprinting (16). Similarly in the mouse conflicting results have been published. *Copg2* was shown to be preferentially expressed from the maternal allele only in particular intra-subspecific hybrids (13) but not in interspecific hybrids involving *Mus spretus* (17). Whereas *Mest* is transcribed from the (+) strand of the mouse genome, *Copg2* is produced from the (–) strand, and their respective final coding exons overlap such that the two transcriptional units are convergent.

In order to resolve the imprinting status of *Copg2* and its relationship with imprinting at the convergent *Mest* locus, we have analyzed the relationship between the two neighboring transcripts. Here we document the formation of larger *Mest* isoforms, called *MestXL,* which are transcribed several kilobases into *Copg2*. During development, the synthesis of *MestXL* is regulated by tissue-specific alternative polyadenylation in neural tissues, where these large transcripts interfere with the production of *Copg2* from the paternal allele. Consequently, allelic usage at *Copg2* is regulated secondarily to imprinting at *Mest* in a tissue-specific manner. Moreover, we show that this type of regulation affects only *Copg2* in the domain as *MestXL* does not regulate allelic usage at the downstream *Klf14* gene. Our results suggest a new model for the acquisition of imprinted expression, via tissue-specific transcriptional interference (TI).

MATERIALS AND METHODS

Mice and genotyping

CD-1 mice are from the Animal Care Centre and the congenic mouse line with Chr 6 *Mus musculus castaneus* SNPs on the 129S1 background was derived in our laboratory (129S1cCAST6/Lef). The *Mest* KO mouse line (MGI: 2181803, official name $Mest^{tm1Lef}$) is maintained on a CD-1 background. All animal experimentation followed the guidelines from the Canadian Council on Animal Care (CCAC) under UBC animal care license numbers A03-0289 and A03-0292.

Heterozygous $Mest^{+/KO}$ males were mated to CD-1 or 129S1cCAST6 female mice. For timed mating, the day of the vaginal plug is E0.5. Females were sacrificed at E14.5 and yolk-sac samples were taken for PCR genotyping and scored for lacZ expression from the $Mest^{KO}$ allele by X-gal staining as described (7,18). Genotyping of the $Mest^{KO}$ allele was performed with primers 702A and IRES1. Genotyping of the wild-type (WT) *Mest* allele was performed with primers 702A and 702C. Primer sequences are given in Supplementary Table S1. Note that for all genotypes the maternally inherited allele is always presented first.

SAGE analysis

Publicly available LongSAGE (19) and SAGELite (20) libraries were generated as part of the Mouse Atlas Project and the Mammalian Organogenesis—Regulation by Gene Expression Networks (MORGEN) Project [www.mouseatlas.org;(21)]. Tag sequences mapping to the sense and antisense strand of the *Mest*/*Copg2* region were identified using in-house Perl scripts.

DIG-labeled RNA probe preparation

Selected mouse cDNA PCR fragments for *Mest*, *Copg2* and *MestXL* were TA-cloned into the pGEM-T vector (Promega). Primer reactions and sequences are given in Supplementary Table S1. Approximately 25 μg of plasmid DNA was linearized with the appropriate restriction enzymes and purified by phenol–chloroform extraction. Approximately 1 μg was transcribed by T7 or SP6 RNA polymerase (40 units, Roche) in the presence of digoxigenin (DIG)-labeled UTP (Roche) to generate sense and antisense DIG-labeled probes. The *in vitro* transcription was carried out at 37°C for 2 h, incubated with DNase for 15 min to remove the plasmid DNA, and purified by precipitation with 4 M LiCl.

RNA purification and qRT–PCR analysis

Conceptuses from crosses between female CAST (or 129S1cCAST6/Lef) or CD-1 mice and male $Mest^{+/KO}$ heterozygotes were obtained at E14.5. For additional experiments, certain tissues such as the developing mid and hind brain (MHB), spinal cord (SC), heart and liver, were further dissected. Embryonic total RNA was purified using Trizol (Invitrogen) according to manufacturer's directions and DNase-treated (RQ1, Promega) for at least 1 h to remove contaminating gDNA. Anywhere from 500 ng to 1 μg of RNA was reverse-transcribed with SSII (Invitrogen), using random (N15) according to SSII protocol. Q–PCR was performed on a Step-One Plus Real time PCR system (Applied Biosystems) using Eva Green (Biotium) as the dye. An amount of 1 μl of cDNA was used in a 25 μl reaction as follows: 35–40 cycles of 95°C for 30 s, 56–62°C for 30 s, 72°C for 30 s, and 79–89°C 10 s with a plate read depending on the primer reaction, followed by a melt curve. Primer sequences are given in Supplementary Table S1. Ct values of three biological replicates, obtained by the LinReg PCR program [(22); www.linregpcr.nl], were averaged and used to calculate relative amounts of transcripts, normalized to *Ppia*.

Northern blot analysis

For the analysis of WT embryonic and adult tissues, northern hybridization was performed on commercial blots (Clontech) containing 2 μg of poly(A)+ RNA from mouse tissues (Cat. # 7762-1) and mouse embryos (Cat. # 7763-1) using a ^{32}P-labeled *Mest* ORF clone (exons 1–12) as a probe. The blots were hybridized at 68° overnight using Clontech quickhyb buffer and washed at maximum recommended stringency (2× SSC/0.1% SDS, 30 min; 55° 0.1× SSC, 0.1% SDS 30 min).

Dynabeads (Invitrogen) were used to obtain polyadenylated RNA from total RNA from E14.5 conceptuses from crosses between female CD-1 mice and male $Mest^{+/KO}$ heterozygous mice. The RNA was separated by electrophoresis on a 1% agarose, 2.2 M formaldehyde denaturing gel. Neutral transfer (20× SSC) was performed

onto nylon membranes, followed by UV-light crosslinking. Northern blot hybridization was carried out with DIG-UTP labeled RNA probes, with an overnight incubation at 65°C in DIG hybridization buffer (50% formamide, 5× SSC, 2% blocking reagent (Roche), 0.1% sarkosyl, 0.5% SDS and 200 μg/ml denatured salmon sperm DNA). Membranes were washed twice with 1× SSC. 0.5% SDS, 0.1% sarkosyl for 5–10 min each, twice with 0.1× SSC, 0.5% SDS, 0.1% sarkosyl for 30 min each, and then twice with DIG-buffer (0.1 M Maleic acid pH 7.5, 0.15 M NaCl, 0.1% Tween-20) for 5 min each to prepare for the DIG antibody. The membranes were then incubated at room temperature for 45–60 min in blocking solution [DIG-buffer with 1% blocking reagent (Roche)], and 45–60 min with DIG antibody (Roche) (1:10 000 dilution) in blocking solution, then followed by two washes with DIG-buffer for 15 min. After equilibrating in 100 mM Tris pH9.5, 100 mM NaCl, chemiluminescent detection with CDP-star reagent (Amersham Biosciences) was performed for 5 min before exposing to film. Exposure times ranged from 30 s to overnight. For one blot we used an alpha ^{32}P-CTP-labeled DNA probe (*Mest* exon 12) generated through linear amplification (20 cycles of 95°C for 30 s, 60°C for 30 s, and 72°C for 1.5 min) from a reverse primer, P1 E12 NO R1, in *Mest* exon 12 to generate an antisense probe. Hybridization was carried out in Church's buffer (350 mM Na2HPO4, 150 mM NaH2PO4, 10 mM EDTA, 7% SDS) overnight at 65°C, washed the following morning twice with low stringency wash (1× SSC, 0.1% SDS) for 10 min at 37°C, and twice with high stringency wash (0.1× SSC, 0.1% SDS) at 65°C for 10 min each. The membrane was then exposed to a phosphor screen overnight and imaged with the Storm Phospho-imager autoradiography system.

Nuclear enrichment analysis

Brain tissue was dissected from three adult CD-1 female mice and snapped frozen with dry-ice. After thawing, the brain tissue was immediately put into lysis solution (40 mM sodium citrate, 1% Triton X-100) and homogenized. The suspensions were transferred into new tubes by straining them with cell-strainer caps (Falcon) to obtain pure suspensions. Centrifugation was carried out at 1000*g* for 5 min and the cytoplasmic supernatant was transferred to another tube. The nuclei pellet was also transferred to a new tube and both the cytoplasmic and nuclear RNA were prepared using Trizol (Invitrogen) by manufacturer's instructions. The RNA was reverse-transcribed as described above and analyzed by several qRT–PCR reactions (see Supplementary Table S1 for primers). Ct values of triplicate samples, obtained by the LinReg PCR program [(22); www.linregpcr.nl], were averaged and used to calculate relative amounts of transcripts, normalized to an exogenous human *XIST* control. The *XIST* control RNA was generated by *in vitro* transcription and 800 fg were introduced into each RNA sample before the RT.

In situ hybridization

Conceptuses from crosses between female CD-1 mice and male $Mest^{+/KO}$ heterozygotes were obtained at E14.5. Embryos were fixed in 4% paraformaldehyde (PFA) overnight, treated in 30% sucrose the following night, and embedded and frozen in OCT (Tissue-Tek). They were sectioned at 12 μm on a Leica cryostat (model CM3050 S). To prepare the sections for hybridization, they were fixed with 4% PFA, washed in 1× PBS, treated with proteinase K, fixed again in 4% PFA, and put through an acetylation step (TEA/AA). The slides were then hybridized (50% formamide, 5× SSC, 5× Denhardts, 0.25 mg/ml yeast tRNA and 0.5 mg/ml single stranded DNA) overnight at 65°C with their respective probe. The following day the slides were washed in different concentrations of SSC to remove non-specific probe, put through an RNase A step, the DIG antibody (1:2000 dilution) reaction, and finally the color reaction step (usually overnight) using NBT/BCIP (Roche) as the substrate. To mount, the slides were prepared by washes with PBS, fixed for 2 h (3.7% formaldehyde, 1× MEM buffer), counterstained with nuclear fast red, dehydrated with varying amounts of ethanol concentrations, and mounted with Entellan.

Allele-specific analysis

The allele-specific *Copg2* hot-stop PCR analysis was based on a polymorphic BstUI site present in the *M. m. castaneus* RT–PCR product (exon 3 SNP rs36827909). The 250-bp PCR products digested with BstUI yield bands of 200 and 50 bp for the CAST allele. Random-primed cDNA was generated from dissected E14.5 embryonic organ RNA and cDNA amplification was carried out with primer reaction *Copg2* ex1-5 for 35 cycles (T_m = 58°C). An amount of 5 μl of the PCR product was transferred into another master mix containing gamma ^{32}P-CTP to go through one final PCR cycle. The hot PCR products were digested with BstUI (New England Biolabs) and analyzed on an 8% polyacrylamide gel. The bands were detected using the Storm Phospho-imager autoradiography system and the band densities (maternal and paternal alleles) were analyzed with the Image J software (W. Rasband, National Institute of Mental Health, Bethesda, Maryland, USA.).

As well, *Copg2* ex1-5 and *Copg2* ex17–19 PCR products from F1 cDNA were directly sequenced. The *Copg2* ex17–19 reaction contains two additional SNPs (exon 19; rs49583711 and rs47417386). The phred program was used to read the sequence trace data and assign quality values to the bases (23), normalized to F1 gDNA.

RESULTS

Longer variants of *Mest*: the *MestXL* isoforms

The current annotation of the *Mest* locus presents a complex production of different isoforms of the mRNA. First, the expressed sequence tag (EST) data suggest variability in promoter usage, with three minor isoforms (#2–4) transcribed from upstream promoters, relative to the main isoform 1, which initiates at the differentially methylated

CpG island defining exon 1 (Supplementary Figure S1A and B) (5). Second, by mapping SAGE tags from the Mouse Atlas of Gene Expression (21) for the *Mest* mRNA, we also identified variability in the size of the 3′-UTR of exon 12 leading to the production of three isoforms (A to C; Supplementary Figure S1C). The majority of *Mest* transcripts (74%) are processed by transcriptional termination and polyadenylation at the most downstream termination signal within exon 12, yielding a terminal exon of 1406 bp.

The production of these isoforms in adult tissues and during development was assessed by northern blotting using a random-primed dsDNA probe against the *Mest* coding region (Figure 1). Our results show that the main ~2.5 kb isoforms of *Mest* (isoforms A, band IV) are broadly expressed in the samples analyzed (see Supplementary Figure S1D for molecular weights of isoforms). Although this analysis cannot resolve the isoforms 1 to 4 generated by alternative promoter usage, most ESTs support formation of isoform 1 in adult and embryonic tissues (96.6% of all ESTs), and ESTs for isoforms 2 and 3 suggest that these promoters are egg-specific. The results also identified lower molecular weight RNA species consistent with the structures of the isoforms B (band V) and C (band VI) produced by alternative polyadenylation within exon 12 and supported by the LongSAGE data (Supplementary Figure S1C). Surprisingly, we also detected polyadenylated transcripts larger than the fully spliced *Mest* mRNA. In adult tissues, the mature *Mest* mRNA of 2.5 kb is detected at high levels in heart and brain. The larger molecular weight species (bands I to III) were mostly detected in brain, where they represent close to 30% of the transcripts (Figure 1A). These larger transcripts represent variants >10.5 kb (band I) and ~6.5 kb of transcribed sequences (bands II and III). During development, the main *Mest* transcript of 2.5 kb is very abundant at all stages analyzed (Figure 1B, band IV). In addition to the shorter isoforms with alternative polyadenylation within exon 12 (band VI, ~1.7 kb; band V, ~2.3 kb) embryos also show the two higher molecular forms at ~6.5 kb (bands II and III) at all stages analyzed.

The annotation of the *Mest-Copg2* loci on the Ensembl browser (NCBIM37) presents evidence for two clusters of ESTs, transcribed from the (+) strand of chromosome 6, like *Mest*, and extending from the end of exon 12 of *Mest* all the way into intron 20 of the neighboring antisense gene *Copg2* (Supplementary Figure S2). Imprinted transcription from the paternal allele has previously been described for these two EST clusters (13). The transcripts from intron 20 of *Copg2* were referred to as *Mit1*, encompassing EST clones covering >5 kb of intron 20 sequences. This region of the current mouse genome sequence (Build 37/mm9) still contains a gap within intron 20 of *Copg2*, the size of which is set arbitrarily and not supported by physical data (24). Spliced *Mest* transcripts with exon 12 variants extending up to this gap would make transcripts of a predicted size of ~6.8 kb. Those extending beyond the gap, assuming a negligible size for this gap and absence of splicing, and up to exon 20 of *Copg2* would span ~13 kb (Supplementary Figure S2). We refer to these long transcripts as the *MestXL* isoforms.

To obtain further experimental evidence for these longer variants, we analyzed data from SAGE tags libraries from the Mouse Atlas of Gene Expression Project, which provide the location and abundance of ESTs located close to terminal or internal poly(A) stretches (21). We identified six main SAGE tags expressed from the (+) strand of the genome in the region of overlap between *Mest* and *Copg2* (Figure 2A). The first four tags (tags 1 to 4) map within the exon 12 of *Mest* and represent expression of the short isoforms of *Mest,* with tag 4 marking the main site of polyadenylation. They are expressed at low levels in adult tissues (notably in brain, pituitary and adrenal glands) but show high levels of expressions during development (Figure 2B). Their expression patterns in

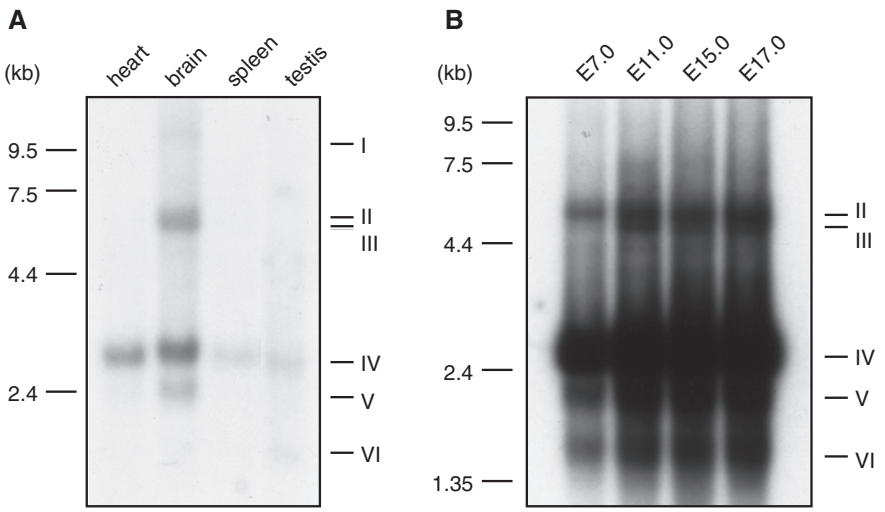

Figure 1. Northern blot analysis of *Mest* expression. A dsDNA probe of the entire *Mest* open reading frame (spanning exons 1–12) was hybridized on mouse northern blots containing 2 μg of poly(A)+ mRNA per lane. (**A**) Adult tissues. (**B**) RNA samples from whole mouse embryos at the indicated developmental stages. The different *Mest* isoforms detected were labeled I to VI. Band IV corresponds to the main isoform A of *Mest* (Supplementary Figure S1).

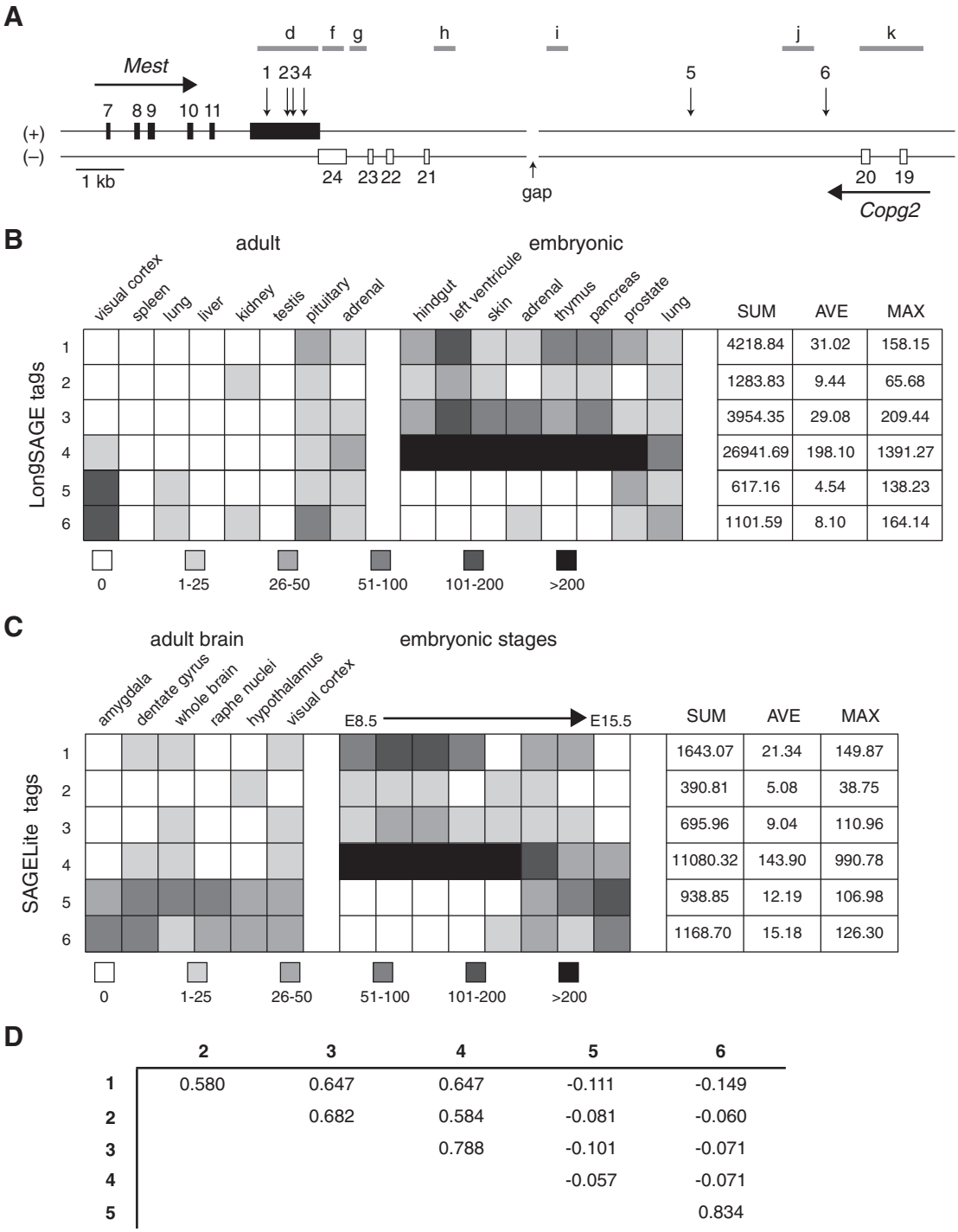

Figure 2. Analysis of SAGE tags for the *Mest* and *MestXL* transcripts. (**A**) Schematic representation of the *Mest-Copg2* overlap region, drawn to scale. The position of the gap in the genome sequence is indicated (below) but omitted in the estimation of the size of intron 20 of *Copg2*. Above the (+) strand, on which terminal *Mest* exons are represented by black rectangles, six SAGE tags are positioned (arrows). Tags 1 to 4 map to exon 12 and represent the short isoforms of *Mest*. Tags 5 and 6, located between the gap and exon 20 of *Copg2* represent the larger *MestXL* isoforms. Probes used in northern and ISH analyses are labeled d to k (gray lines). (**B**) Gene expression profiles for SAGE tags 1 to 6 in adult and embryonic LongSAGE (un-amplified) libraries from the given tissues, presenting the tag counts according to the shading scale defined below. Statistics for each tag is from 136 different LongSAGE libraries. (**C**) Expression levels of tags 1–6 in adult and embryonic SAGELite (amplified) libraries from the given adult tissues. For the embryos, whole head RNA was analyzed from E8.5 to E11.5, and whole brain from E12.5 to 16.5. Statistics for each tag is from 77 different libraries. (**D**) Pearson's correlation coefficient for the six SAGE tags in the 136 LongSAGE libraries.

136 different adult and embryonic LongSAGE libraries (19) show high levels of correlation (Figure 2D), suggesting that the ratios of isoforms A to C are rather constant across tissues and do not represent tissue-specific isoforms.

The SAGE tags 5 and 6 map within intron 20 of *Copg2*, beyond the gap in the mouse genome sequence, relative to exon 12 of *Mest* (Figure 2A), and within the *Mit1* cluster of ESTs (Supplementary Figure S2). We refer to these tags as representative of longer isoforms of *Mest,* the *MestXL* variants. Expression of tags 5 and 6 show the highest degree of correlation amongst the tags analyzed in the 136 LongSAGE libraries (Figure 2D). They show moderate to low levels of expression in adult and embryonic tissues expressing the short *Mest* isoforms, but they are also sometimes expressed at higher levels than tags 1 to 4 (for instance in adult visual cortex and lung, Figure 2B). The analysis of amplified libraries (SAGELite (20)) showed that transcripts terminating at tags 5 and 6 are often more abundant than the short isoforms in several regions of the adult brain (Figure 2C). During embryogenesis we see a gradual shift from *Mest* to *MestXL* polyadenylation in the developing brain, the latter of which is first detected at developmental stage ~E13 (Figure 2C).

We also analyzed all the ESTs from the (+) strand at *Mest* (from UniGene entry Mm.335639) and within the two EST clusters on either side of the gap (Supplementary Tables S2 and S3). Whereas only 35% of *Mest* ESTs are from libraries generated from central nervous system (CNS) tissues, most ESTs from the 5′- and 3′-clusters are from the CNS (77 and 84%, respectively; Supplementary Figure S3). Together with the northern blot analysis, these results suggest the existence of longer isoforms of *Mest*, and that the production of *MestXL* transcripts is under developmental and tissue-specific regulation.

Loss of *MestXL* transcripts from the mutant $Mest^{KO}$ allele

The targeted allele $Mest^{tm1Lef}$ (or $Mest^{KO}$) is based on a promoter-trap construct in which exons 3 to 8 and part of exon 9 of *Mest* were replaced by a splice–acceptor–IRES–βgeo–pA cassette (Figure 3A) (7). Transcription from this mutant allele is imprinted and expression of the βgeo reporter faithfully recapitulates *Mest* expression in paternal heterozygotes (7). Based on the structure of the mutant allele, transcriptional termination is predicted to occur upstream of the remaining exon 9 sequences and to produce truncated *Mest* transcripts. To provide support for this prediction we performed northern blots on poly(A)+ RNA from WT and $Mest^{+/KO}$ E14.5 embryos and adult brain tissue (Figure 3B and Supplementary Figure S4). Since only the paternal allele of *Mest* is active for transcription, all *Mest* transcripts are from the mutant KO allele in $Mest^{+/KO}$ heterozygotes (7). We used strand-specific RNA probes hybridizing to different regions of *Mest,* to the mutant allele (*neo* probe), and to three different positions along the *MestXL* isoforms (BB692936 EST, BY247745 EST, and *Mit1*). Probe locations are presented in Figures 2A and 3A.

Within the *Mest* locus, probes against exons 3–8 (probe b) and exon 12 (probe d) detected the *Mest* and *MestXL* isoforms in WT E14.5 head RNA (Figure 3B). These blots are over-exposed so that we could observe the weaker higher molecular weight signals. Since these exons are deleted in the $Mest^{KO}$ allele, the absence of signal with the exons 3–8 probe on RNA from $Mest^{+/KO}$ embryos confirms that only the paternal allele is active for all the isoforms detected in the WT embryos. Premature transcriptional termination at the inserted poly(A) cassette of the $Mest^{KO}$ allele is confirmed by the absence of signal with the exon 12 probe on RNA from $Mest^{+/KO}$ embryos as well as the detection of the expected 5-kb mutant RNA using a probe against the *neo* marker. We did not obtain a good signal for the mutant transcript using a small exons 1–2 probe on embryonic RNA although the 5-kb mutant mRNA was detected with this probe on adult brain RNA (Supplementary Figure S4), as well as with the *neo* probe. Furthermore, qRT–PCR analysis of the same region showed that the *Mest* KO transcripts are expressed at ~50% of WT levels (Supplementary Figure S5) and rapid amplification of cDNA end (5′-RACE) experiments confirmed that the mutant transcript indeed initiates at exon 1 of *Mest* (Supplementary Figure S6). It is therefore possible that the formation of stable mutant transcripts including exons 1 and 2 is affected by abnormal processing of the intron 2 *Mir335* during development.

We also considered the possibility that the higher molecular weight isoforms might represent the pre-mRNA (10.5 kb) or partially spliced forms of *Mest*. To address this, we used a genomic probe against intron 11 (probe c). The overexposed signal presents a smear that ranges from 2.8 to 7 kb with no prominent bands detected (Figure 3B). The absence of the higher molecular weight bands detected with *Mest* cDNA probes suggests that they represent processed *Mest* isoforms and not unspliced primary transcripts. The absence of signal on the mutant RNA sample provides further evidence for the absence of read-through from the inserted polyA cassette, a result which is also supported with qRT–PCR analysis (Supplementary Figure S5).

To analyze the formation of the *MestXL* isoforms, we used three probes (h-j) from intron 20 of *Copg2* (Figure 2A). These probes, which span >7.5 kb of genomic sequences, detect RNA bands of ~4.8 kb, ~6.0 kb, as well as higher molecular weight forms (Figure 3B). Notably, the probe against the EST BB692936, located at the end of *Copg2* intron 20 and 5′ of the gap (probe h), detects the higher molecular weight RNA isoforms corresponding to bands I-III seen with the *Mest* cDNA probes. More importantly, all the isoforms detected with these intron 20 probes are absent in $Mest^{+/KO}$ embryos, in which transcription terminates at the inserted polyA signal (Figure 3B).

Our northern blot analysis shows that *Mest* and *MestXL* transcripts are imprinted and expressed only from the paternal allele during development. In $Mest^{+/KO}$ embryos, *Mest* transcription terminates at the inserted polyA cassette of the mutant allele and production of the *MestXL* isoforms is concomitantly eliminated.

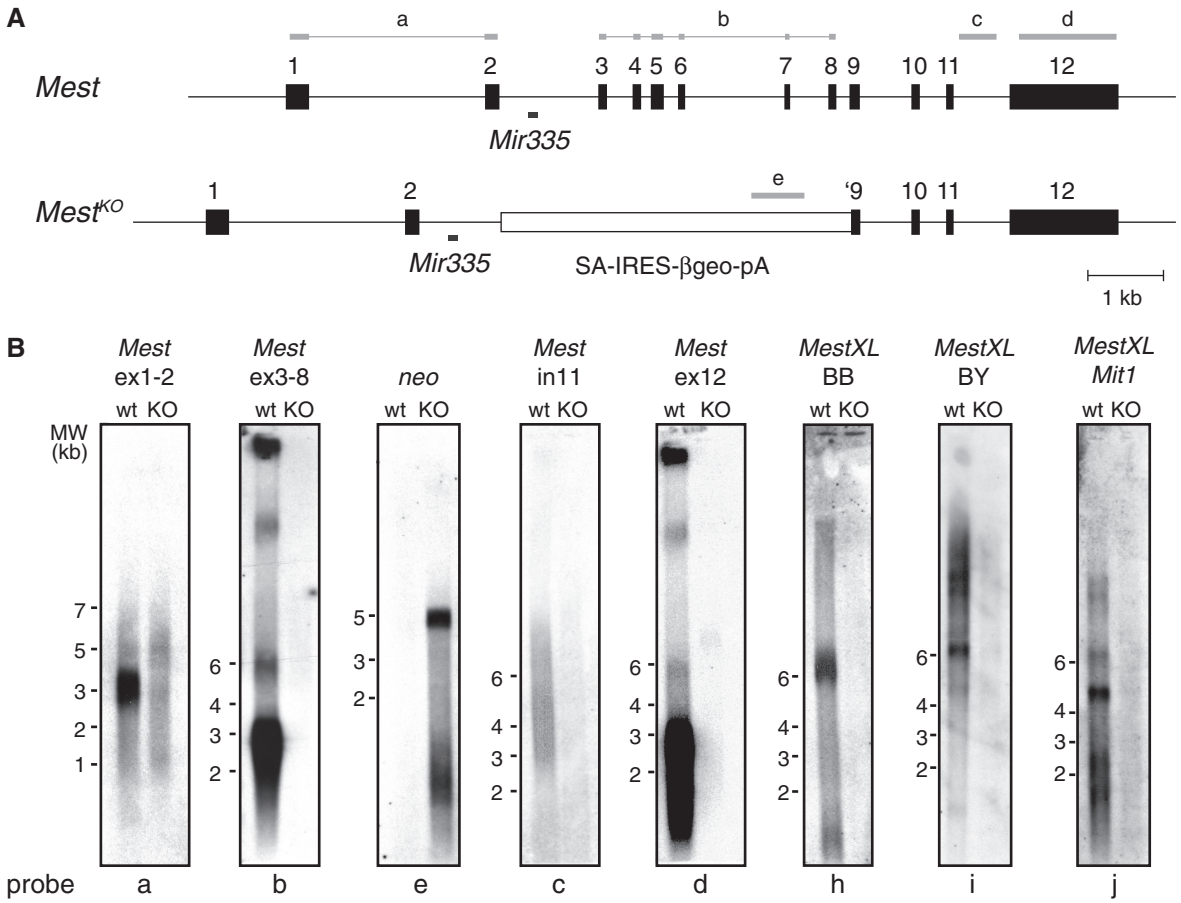

Figure 3. Loss of *MestXL* expression from the mutant $Mest^{KO}$ allele. (**A**) Schematic diagrams of the *Mest* and $Mest^{KO}$ alleles. The exons are represented by black rectangles. The $Mest^{KO}$ allele consists of a splice acceptor-IRES-lacZ/neo fusion-poly(A) cassette (white rectangle) replacing *Mest* exons 3–8, and part of exon 9. The positions of the RNA probes **a** to **e** are show above the alleles (gray lines). The intronic *Mir335* is shown below each allele (black line). (**B**) Strand-specific northern blot analyses of *Mest*, $Mest^{KO}$, and *MestXL* expression. Various DIG-labeled RNA probes were hybridized to mouse northern blots containing 1.5 µg of poly(A)$^+$ RNA purified from E14.5 WT and $Mest^{+/KO}$ embryo heads. See Figures 2A and 3A for probe positions, identified by the letters given below the bots. BB and BY are (+) strand ESTs (BB692936 and BY247745, probes h and i, Figure 2A) mapping 5′ and 3′ of the gap, respectively. The position of molecular weights marker bands in kb (MW) are indicated on the left sides of each blot.

MestXL transcripts are produced in the developing nervous system

To look at the developmental expression pattern of *MestXL* transcripts in more detail, we performed *in situ* hybridization experiments on sagittal sections of E14.5 embryos (Figure 4), at a stage when both *Mest* and *MestXL* SAGE tags are detected in the developing brain (Figure 2C). First, we used antisense RNA probes to look at the expression of *Mest* in WT and $Mest^{+/KO}$ embryos. We also analyzed expression of the *neo* marker present in the *β-geo* cassette of the $Mest^{KO}$ allele. Our results confirm the widespread mesodermal expression of *Mest* during development (3,25) as well as the previously described expression in neuronal tissues (7,8). In the E14.5 spinal cord, immunohistochemistry experiments show that both *Mest* and *Mest* KO are broadly expressed in NCAM1-positive neurons (Supplementary Figure S7). The ISH signal is seen for all the antisense *Mest* cDNA probes assayed, but not for the intron 11 probe (Figure 4A). For the exon 1–2 probe (a), we detected weaker expression for the mutant allele than for the WT transcript, which is consistent with our qRT–PCR data (Supplementary Figure S5). Together these results show that mRNA levels from the promoter-trap $Mest^{KO}$ allele are lower than those detected in WT embryos as a consequence of the insertion of the *β-geo* cassette. Our results also show absence of transcription from the WT maternal allele in $Mest^{+/KO}$ heterozygotes, using the exon 3 to 8 probe, which hybridizes to coding sequences deleted in the $Mest^{KO}$ allele. We evaluated the amount of read-through beyond the inserted polyadenylation signal of the $Mest^{KO}$ allele by using an exon 12 probe. In agreement with our northern blot and qRT–PCR results, very little transcription was detected downstream of the $Mest^{KO}$ poly(A) site.

The expression pattern of *MestXL* was studied using RNA ISH probes complementary to the (+) or (–) strand that hybridize downstream of the *Mest* isoform 1A to ESTs annotated in the genome browsers. In addition, we generated one probe for which there is no

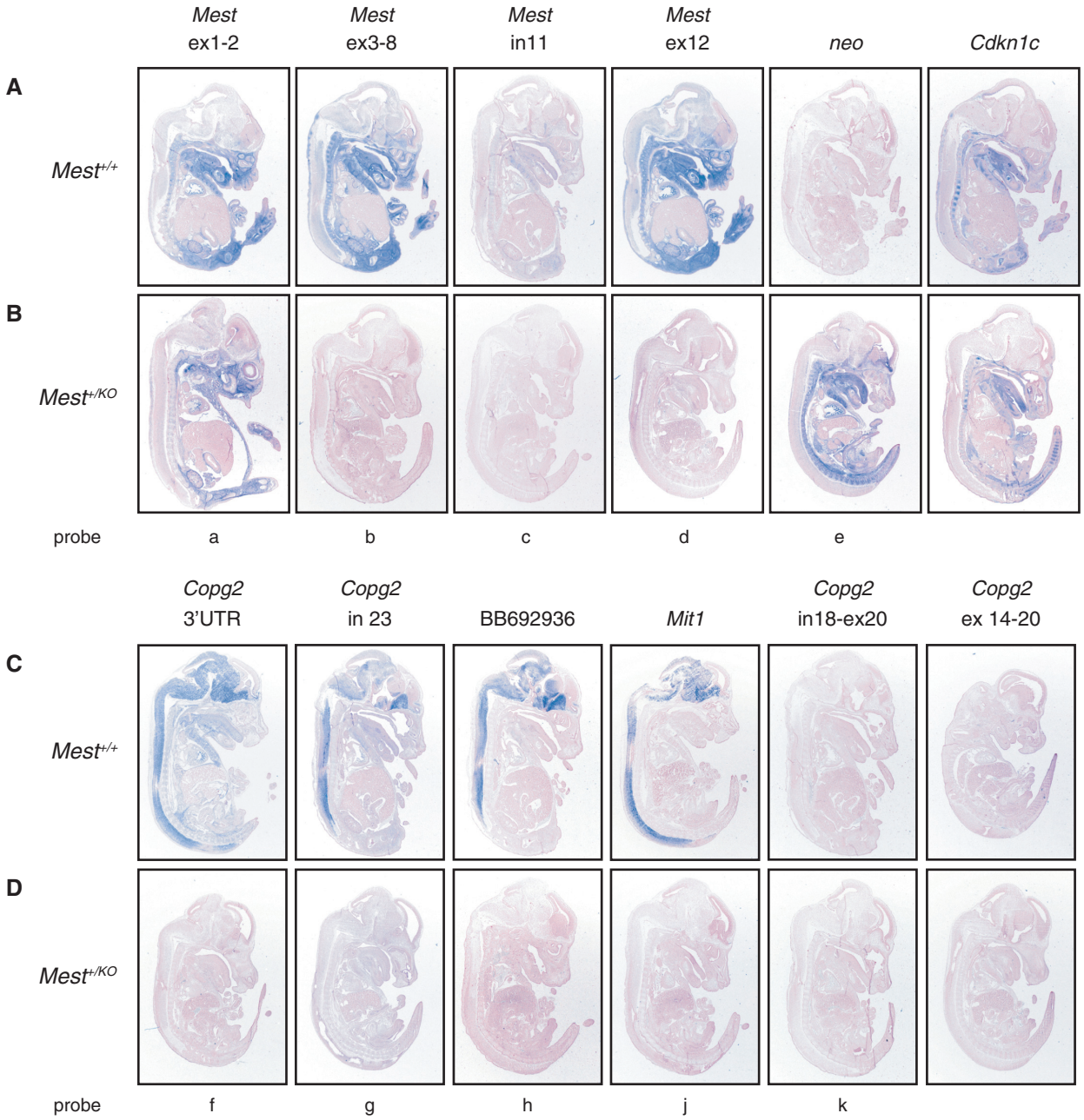

Figure 4. *MestXL* is expressed in the developing central nervous system of WT but not $Mest^{+/KO}$ embryos at E14.5. ISH was performed on 12 µm sagittal cryosections of WT ($Mest^{+/+}$) (**A** and **C**) and $Mest^{+/KO}$ (**B** and **D**) E14.5 embryos with various DIG-labeled RNA probes. The probe locations are indicated in Figures 2A and 3A using the letter code presented below each panel. All probes hybridize to the (+) strand. Blue staining indicates gene expression and nuclear fast red was used as a counterstain. An antisense probe against the imprinted gene *Cdkn1c* was used as a control.

supporting EST data, the *Copg2* intron 23 probe (probe g, Figure 2A). The BB692936 EST and *Mit1*/*AF217545* are located in intron 20 of *Copg2* and transcribed from the (+) strand. The four probes designed to detect *MestXL* show a localized expression signal in the developing nervous system (Figure 4C, probes f to j). More specifically, *MestXL* appears to be expressed in most parts of the developing mid/hind brain and in the spinal cord, but not in the developing forebrain. This CNS-specific transcription is absent in $Mest^{+/KO}$ embryos, confirming imprinting of *MestXL* and suggesting that it represent longer *Mest* transcripts, produced by tissue-specific alternative polyadenylation (Figure 4D). A probe designed to detect transcription downstream of *Mit1* (*Copg2* in18-ex20, probe k) shows that transcription of *MestXL* terminates in intron 20, which is supported by SAGE tag 6 (Figures 2A and 4D).

The reciprocal probe partners detecting transcription from the (–) strand, either the *Copg2* mRNA or primary mRNA, were analyzed as well. For all probes assayed,

minimal expression was observed for *Copg2* mRNA or primary mRNA transcripts in both WT and $Mest^{+/KO}$ heterozygous embryos with the exception of the *Copg2* exons 14–20 and 3′-UTR cDNA probes (Supplementary Figure S8). As observed for *MestXL, Copg2* is mainly expressed in the developing neural tissues.

Together with our northern data, the expression analysis by ISH suggests that *MestXL* is imprinted, produced exclusively in the developing and adult CNS, and lost in $Mest^{+/KO}$ mutants. These conclusions are also supported by extensive RT–PCR experiments detecting brain-specific transcripts in the *Mest-Copg2* region only in WT samples (Supplementary Figure S9). Our analysis of expression profiling by high-throughput sequencing (RNA-seq) also revealed transcription extending from exon 12 of *Mest* to exon 20 of *Copg2*, specifically in developing and adult neuronal tissues (Supplementary Figure S10).

Allelic usage at *Copg2*, but not *Klf14*, is regulated by *MestXL*

The relationship between *Mest*/*MestXL* and *Copg2* is of particular interest due to their 3′-end overlap, which is conserved in humans (13,16). *Copg2* was reported to be predominantly maternally expressed in a tissue-specific manner, specifically in tissues where we document the formation of *MestXL* transcripts (13). To investigate whether the production of *MestXL* via alternative polyadenylation at *Mest* has an effect on allelic-usage at *Copg2* we examined allele-specific expression of *Copg2* in $Mest^{+/KO}$ heterozygous animals, in which production of *MestXL* is abolished. Having established that the *MestXL* transcripts are mainly detected in the CNS in the developing mouse embryo, we investigated the parental allele contributions and expression of *Copg2* in a tissue specific manner.

We first investigated the imprinting pattern of *Copg2* in four E14.5 WT and $Mest^{+/KO}$ tissues: two where *MestXL* is expressed, the MHB and spinal cord (SC), and two where *MestXL* is not expressed; the heart and liver. To determine the maternal and paternal allele contributions from *Copg2* we used hot-stop PCR/RFLP analysis (Figure 5A). Labeling the PCR products only in the final cycle is advantageous since this circumvents the formation of un-digestible heteroduplexes (26). Our analysis relied on a SNP in exon 3 of *Copg2* (rs36827909) which creates a BstUI site unique to the *M. m. castaneus* allele. By measuring the band densities of the maternal (CAST) and paternal (CD-1) alleles we obtained quantitative measurements and calculated maternal to paternal (M/P) allele ratios for WT and $Mest^{+/KO}$ samples purified from F1 hybrid embryos. In WT samples, the MHB shows the highest *Copg2* M/P ratio as the maternal allele is expressed 2.3 times more than the paternal allele (Figure 5B). When *MestXL* expression is lost in the $Mest^{+/KO}$ MHB, this allelic bias is lost and the M/P ratio plummets to 0.86. A similar situation occurs in the SC but it is not as striking; the WT and $Mest^{+/KO}$ M/P ratios are 1.40 and 0.92, respectively. In tissues where *MestXL* is not expressed, the heart and liver, the M/P ratios basically stay the same between WT and $Mest^{+/KO}$ samples. Here the WT and $Mest^{+/KO}$ ratios are 1.38 and 1.49 for the heart and 0.78 and 0.71 for the liver.

We also directly sequenced the exon 3 SNP and two additional SNPs in exon 19 (rs49583711 and rs47417386). The quantification for this method involves using the phred software (23) to calculate the area under the called base (SNP) tracing. After normalizing the read-outs to the F1 genomic DNA tracing, the M/P ratios were calculated by averaging the reads for each SNP in 4 WT and 6 $Mest^{+/KO}$ samples from each E14.5 tissue. The results showed similar patterns to the hot-stop PCR analysis and suggest that allelic usage at *Copg2* is regulated by *MestXL* (Figure 5C).

Considering that the paternal allele of *Copg2* is aberrantly expressed in $Mest^{C/KO}$ embryonic neural tissues, our next question was whether this leads to an overall increase in *Copg2* expression levels, as suggested by the allele-specific data. qRT–PCR analysis of *Copg2* levels was performed on cDNA from four $Mest^{C/+}$ (WT) and six $Mest^{C/KO}$ E14.5 embryos from two F1 (CAST6/CD-1) litters. The results showed striking differences between WT and mutant samples only in tissues where *MestXL* is expressed (Figure 5D). In the mutant MHB, *Copg2* is expressed 1.4× higher than in the WT ($P = 0.012$), a significant difference between the two genotypes. In the $Mest^{C/KO}$ SC, *Copg2* is expressed 1.2× higher than in the WT, a difference that approaches significance ($P = 0.09$). On the other hand, in the heart and liver, *Copg2* expression levels remain similar between the $Mest^{C/+}$ and $Mest^{C/KO}$ samples ($P = 0.81$ and 0.62, respectively). Our results show that *Copg2* expression exhibits tissue-specific allele-specific bias which are regulated by the formation of *MestXL*.

We also analyzed whether *MestXL* is involved in regulating *Klf14*, located ~60-kb downstream from the *Copg2* promoter and transcribed from the (–) strand (4). A previous report demonstrated that *Klf14* itself does not harbor any of the known imprinting hallmarks such as a DMR or differential histone modifications that would mediate its imprinting, but is regulated by maternal methylation (4). A mechanism whereby the maternally methylated *Mest* DMR might regulate this imprinting is through a possible long-range effect by paternally expressed *MestXL* in *cis*. To investigate this possibility we again took advantage of our $Mest^{C/KO}$ line that contains a truncated *MestXL*. Allele-specific analysis and qRT–PCR on $Mest^{C/KO}$ E14.5 tissues show that *MestXL* is not implicated in the regulation of imprinting at *Klf14* (Supplementary Figure S7). This result is in agreement with our observation that *MestXL* transcription terminates before exon 20 of *Copg2*, therefore supporting a local *cis* effect. In support of this model, we additionally found that *MestXL* is not localized to the nucleus, a characteristic feature of regulatory non-coding RNAs such as *Air*, *Kcnq1ot1* and *Xist* that are thought to mediate long-range regulation of epigenetic silencing (Figure 5E).

DISCUSSION

We show here that transcription at the imprinted *Mest* locus is regulated by tissue-specific alternative

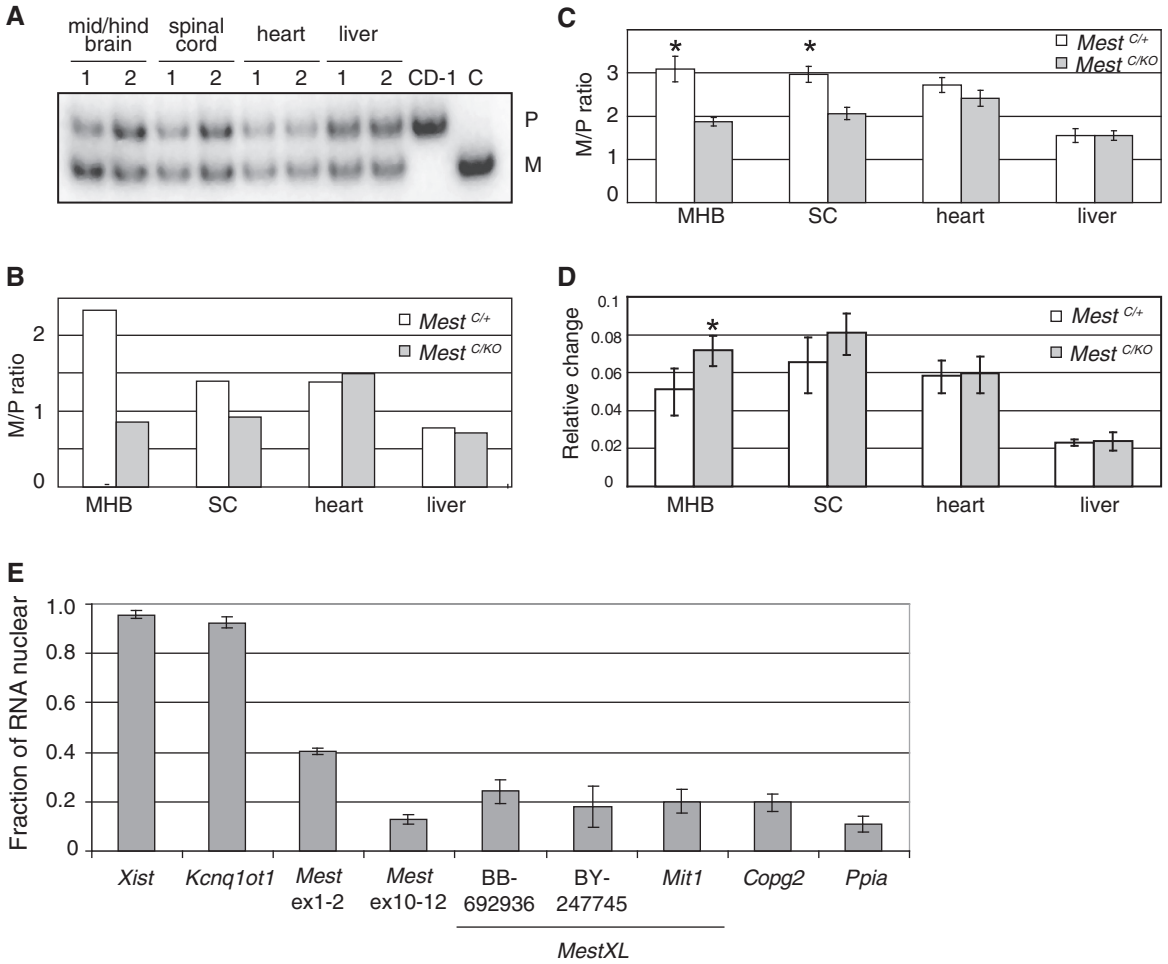

Figure 5. Allelic usage at *Copg2* is regulated by *MestXL*, which is not nuclear localized. (**A**) Hot-stop PCR/RFLP analysis on various E14.5 F1 $Mest^{C/+}$ (1) and $Mest^{C/KO}$ (2) tissues reveals allelic parental differences for *Copg2* in tissues that express *MestXL*. ^{32}P-labeled *Copg2* RT–PCR products were digested with BstUI, which cuts only the maternal CAST (C) allele, and ran out on an 8% polyacrylamide gel. Controls on paternal (CD-1) and maternal (C) RNA show complete digestion of the maternal CAST allele with BstUI. P:paternal allele; M:maternal allele. (**B**) Maternal and paternal band densities were calculated using Image J analysis software and an M/P ratio was generated for each sample. (**C**) Direct sequencing of three *Copg2* SNPs for RNA isolated from four different tissues from 4 $Mest^{C/+}$ and 6 $Mest^{C/KO}$ E14.5 embryos expressed as maternal to paternal (M/P) allele ratios. The phred program was used to read the sequence trace data and assign quality values to the bases (24) and the numbers were generated after normalizing to F1 genomic DNA. Differences between $Mest^{C/+}$ and $Mest^{C/KO}$ for each E14.5 tissue were evaluated using the two-tailed unpaired Student *t*-test (*P*-values) and those differences that were statistically significant are marked with an asterisk ($P < 0.0001$ in MHB and $P = 0.0004$ in SC). Error bars indicate the standard error of the mean. (**D**) qRT–PCR analysis of *Copg2* expression levels in various E14.5 tissues (MHB—mid/hind brain, SC—spinal cord, H—heart, L—liver). *Copg2* mRNA levels were normalized to *Ppia* and the error bars indicate the standard deviation from four $Mest^{C/+}$ and six $Mest^{C/KO}$ biological replicates for each tissue analyzed. *($P = 0.012$). (**E**) Nuclear and cytoplasmic RNA fractions from female adult brain tissue were separated using a sodium citrate/Triton X-100 protocol and analyzed by qRT–PCR. mRNA levels were normalized to an exogenous *XIST* control that was generated by *in vitro* transcription and added to the nuclear and cytoplasmic RNA samples at the RT step. The error bars indicate the standard deviation from biological triplicates. A value of 1 indicates 100% nuclear enrichment and the values were calculated by dividing the nuclear E^dCt values by the total (nuclear E^dCT + cytoplasmic E^dCt) values. *Xist* and *Kcnq1ot1* are enriched in the nuclear RNA fraction but not *Mest*, *MestXL*, or *Copg2*.

polyadenylation, resulting in the production of longer variants, the *MestXL* transcripts. These CNS-specific transcripts extend far into the 3′-end of the convergent *Copg2* locus and regulate the allelic usage at *Copg2* in these specific tissues. The mechanism of this tissue-specific alternative polyadenylation at *Mest* remains to be elucidated but is certainly not unique to *Mest*. Over half of mammalian genes contain multiple polyadenylation sites that lead to variants with different 3′-UTRs or coding regions (27). Using SAGE data, several mRNAs expressed in neuronal tissues were shown to have progressively longer 3′-UTRs throughout embryonic and postnatal developmental stages (28), a characteristic shared by *MestXL*. Examples of neuronal-specific larger variants of imprinted genes have also been documented. *Dlk1* was found to produce a transcript in brain that is 2.7 kb larger than the main transcript (29). Longer 3′-UTRs for genes expressed in the brain have been involved in translational control/repression and differential RNA localization (30–32). Whether *MestXL* is involved in the

translational repression of *Mest*, besides its role in regulating allelic usage at *Copg2* remains to be determined.

The *Mest* locus consists of a variety of alternatively spliced transcripts coupled with alternative promoter usage. Together with our findings that it is also under the regulation of tissue-specific alternative polyadenylation, the *Mest* locus appears to be more complex than previously thought. We showed that the two higher molecular weight isoforms detected on our northern blots include the *Mest* ORF or part of it. This suggests that *MestXL* consists of the mature *Mest* mRNA with an extended terminal exon, a conclusion supported by our results using downstream probes proximal to the genomic gap. Using additional probes detecting transcripts from the (+) strand distal to the gap, we observed not only higher molecular bands that appear to be >10 kb, but also prominent bands at sizes such as 4.8 and 6.5 kb suggesting these transcripts are either differentially processed isoforms or cleavage products. This type of regulation has been observed for many genes (33). Interestingly, the *Mest* and *MestXL* probes seem to share the 6.5-kb band perhaps suggesting they could all be part of the same transcript. Given that a 4.8-kb band is not detected from the *Mest* probes, it is likely that this transcript does not include the *Mest* mRNA transcript and is comprised of sequences mostly on the distal side of the gap. The highest molecular weight transcripts themselves that we detect could also be processed transcripts if the gap is very large. The gap might complicate and impede our understanding of how the *MestXL* transcripts are regulated and processed throughout this region, but our results provide evidence that their expression relies on transcription from the *Mest* locus and are under the regulation of alternative polyadenylation.

The allele-specific expression of *Copg2* shows great variability in different genetic backgrounds of F1 hybrid mouse strains. Three different kinds of hybrids have been used to study the imprinting at *Copg2* and all have given different results. In an intraspecific hybrid between C57BL/6 and *M. m. molossinus*, *Copg2* is more exclusively maternally expressed in the developing embryo and adult brain tissue (13). However, in an interspecific F1 hybrid between C57BL/6 and *M. spretus*, *Copg2* was biallelically expressed in all adult tissues analyzed including in the brain (17). For our studies, we used a different intraspecific F1 hybrid cross between CD-1 and *M. m. castaneus* and found that *Copg2* is preferentially maternally expressed in a tissue specific manner. Our results suggest that *Copg2* transcription itself is not primarily imprinted. Although not studied in great detail, and despite the claims of the authors, *Copg2* does not appear to have a promoter-proximal DMR that would mediate its imprinting (17). Instead we propose that *Copg2* transcription is secondarily imprinted due to the production of larger *Mest* variants, *MestXL*, that are under the regulation of alternative polyadenylation (Figure 6). As a consequence, these variants would interfere and regulate the allelic usage at *Copg2*. The variability observed in different hybrid strains suggests that whatever mechanism is controlling the imprinting at *Copg2* is possibly rapidly evolving and, considering our results, perhaps the variation is at the level of alternative polyadenylation from the *Mest* locus.

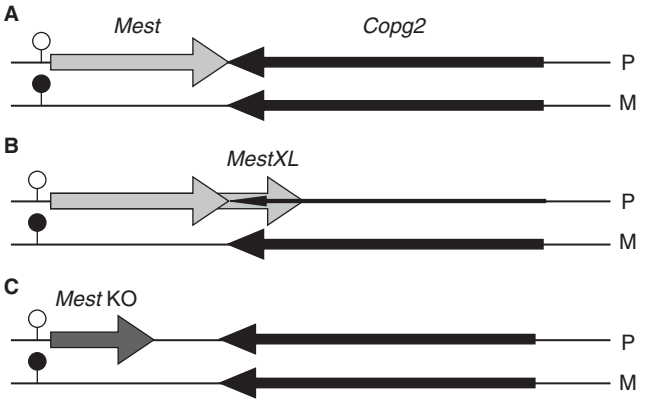

Figure 6. Model for regulation of allelic bias at *Copg2* by *MestXL*. On the basis of the results presented in this study, we propose the following model for the regulation of allelic bias at *Copg2*. Transcription from the *Mest* promoter is directly regulated by a germ line DNA methylation imprint inherited from oocytes (black lollipop on the maternal allele, M), such that only the paternal allele (P) is active for transcription. (**A**) In tissues where *Mest* transcription terminates within exon 12, such as embryonic liver, *Copg2* is biallelically expressed (M/P = 1). (**B**) In the CNS, alternative polyadenylation leads to the formation of the *MestXL* transcripts which extend all the way to the beginning of intron 20 of *Copg2*, in addition to the normal *Mest* transcripts. We propose that *MestXL* exerts a *cis* negative effect on the production of the *Copg2* mRNA from the paternal allele (reduced black arrow), via TI. As a consequence, stable *Copg2* transcripts are preferentially derived from the maternal allele (M/P = 2.3). (**C**) In $Mest^{+/KO}$ embryos, *MestXL* transcripts are truncated by the inserted polyadenylation sequence of the mutant allele and *Copg2* is biallelically expressed in the CNS.

Our results support a model in which allelic usage at *Copg2* is regulated by the production of larger variants at the *Mest* locus. In tissues that express *MestXL*, *Copg2* is preferentially maternally expressed due to the TI on the paternal chromosome. This imprinting is lost in $Mest^{+/KO}$ tissues where *MestXL* is truncated due to the introduced poly(A) signal in the $Mest^{KO}$ allele (Figure 6). The effect is an overall increase in *Copg2* expression levels which may in fact contribute to the phenotype described for $Mest^{+/KO}$ mice. In tissues where *Mest* is basally expressed such as in the developing liver, *Copg2* is biallelically expressed because there is no competing TI. Previous microarray expression data from $Dnmt3l^{-/+}$ E8.5 embryos, which do not inherit proper methylation marks from the maternal germline, provides further evidence for this model (34). Since the maternal *Mest* promoter is not methylated in these embryos, *Mest* is upregulated 1.5 times presumably due to biallelic expression. A probe located downstream of *Mest* that maps to BB692936, a *MestXL* EST studied in our experiments, shows a 1.8-fold upregulation which is most likely because the maternal allele is transcribed as well. Most importantly, *Copg2* is downregulated 2-fold in the mutants. We propose that it is the biallelic expression of *MestXL* which interferes with the transcription of *Copg2* in these mutant embryos.

The mechanism of this TI is unknown but current models propose two possibilities. The 'collision' model suggests that collisions between converging sense and

antisense elongation complexes in *cis* can lead to premature termination of the transcriptional progress of one or both complexes in yeast (35). Elongation of the regulatory *MestXL* Pol II complex could dislodge the target *Copg2* Pol II complex terminating its transcription. Alternatively, the overlapping double-stranded 3′-UTRs could lead to adenosine–inosine editing of the paternal *Copg2* 3′-UTR which might target it for degradation (36,37). Perhaps the termination complex of *MestXL* interferes with changes in chromatin structure needed for transcription of *Copg2,* an idea previously proposed (38,39).

We demonstrated that *MestXL* regulates allelic usage at *Copg2* but not at *Klf14*. Since *Klf14* is exclusively expressed from the maternal allele, does not carry a germline imprint, and is silenced in absence of maternal methylation imprints, it was previous proposed that *Klf14* might be under long-range regulation via the *Mest* promoter DMR (4). Since loss of *MestXL* does not abrogate imprinting at *Klf14,* our results indicate that the regulation of imprinting within this domain involves a combination of different mechanisms.

The *Mest* imprinting domain shares homology with human chromosome 7q32 and there is supporting data that the human *MEST* locus produces *MESTXL* variants. Yamasaki *et al*. identified an alternative 3′-UTR for *MEST* that is at least 2.7 kb larger than the main 3′-UTR in the four fetal tissues analyzed (not neural specific) and they additionally demonstrated that *COPG2* intronic transcript 1 (CIT1), an antisense transcript located in intron 20 of COPG2, was paternally expressed (16). The mouse and human sequences for intron 20 of *Copg2* are 80% identical, although both genomic sequences are incomplete and contain a gap. This *MEST*/*COPG2* region has been associated with growth retardation and Silver-Russell syndrome (SRS) in maternal uniparental disomies (mUPD) for this chromosome 7 segment, where both copies are inherited from the mother. Nearly 10% of SRS cases are associated with mUPD for this region presumably due to the loss of paternally expressed genes (40,41). As a consequence, *MEST* was designated as a key candidate locus for the etiology of SRS in these cases yet extensive sequencing and methylation studies have failed to provide any evidence and the SRS etiology associated with UPD at 7q32 remains ambiguous and unknown (42,43). Since *MestXL* is paternally expressed and its loss affects *Copg2* expression, this raises the possibility that increased *COPG2* levels might be implicated in the etiology of SRS.

To conclude, this is the first report to demonstrate that the *Mest* locus produces larger transcripts, *MestXL*, that include the previously reported transcript *Mit1*. Furthermore, we document for the first time that they have a role in regulating the allelic usage at *Copg2* and propose a model for allelic interactions between neighboring transcripts as a new mechanism for the regulation of imprinted expression in mammals. The analysis presented in this study further shows that growth retarded $Mest^{+/KO}$ mutant mice (7) are not only deficient for the MEST protein, but also lack the *MestXL* transcripts, a conclusion which might have important implications for the role of this region in SRS in humans.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary methods, Supplementary Tables 1–3, Supplementary Figures 1–11, Supplementary references [7,22,44–48].

ACKNOWLEDGEMENTS

The authors would like to thank Rosemary Oh-McGinnis for help with the ISH protocol and Sam Aparicio for the multiple tissue northern blots.

FUNDING

Canadian Institutes of Health Research [MOP-82863 to L.L.]; Canada Research Chair to L.L. Funding for open access charge: Canadian Institutes of Health Research [MOP-82863 to L.L.].

Conflict of interest statement. None declared.

REFERENCES

1. Beechey,C.V. (2000) *Peg1*/*Mest* locates distal to the currently defined imprinting region on mouse proximal chromosome 6 and identifies a new imprinting region affecting growth. *Cytogenet. Cell Genet*., **90**, 309–314.
2. Beechey,C.V. (2004) A reassessment of imprinting regions and phenotypes on mouse chromosome 6: Nap1l5 locates within the currently defined sub-proximal imprinting region. *Cytogenet. Genome Res*., **107**, 108–114.
3. Kaneko-Ishino,T., Kuroiwa,Y., Miyoshi,N., Kohda,T., Suzuki,R., Yokoyama,M., Viville,S., Barton,S.C., Ishino,F. and Surani,M.A. (1995) *Peg1*/*Mest* imprinted gene on chromosome 6 identified by cDNA subtraction hybridization. *Nat. Genet*., **11**, 52–59.
4. Parker-Katiraee,L., Carson,A., Yamada,T., Arnaud,P., Feil,R., Abu-Amero,S., Moore,G., Kaneda,M., Perry,G., Stone,A. *et al.* (2007) Identification of the imprinted *KLF14* transcription factor undergoing human-specific accelerated evolution. *PLoS Genet*., **3**, e65.
5. Lefebvre,L., Viville,S., Barton,S.C., Ishino,F. and Surani,M.A. (1997) Genomic structure and parent-of-origin-specific methylation of *Peg1*. *Hum. Mol. Genet*., **6**, 1907–1915.
6. Lucifero,D., Mertineit,C., Clarke,H.J., Bestor,T. and Trasler,J.M. (2002) Methylation dynamics of imprinted genes in mouse germ cells. *Genomics*, **79**, 530–538.
7. Lefebvre,L., Viville,S., Barton,S.C., Ishino,F., Keverne,E.B. and Surani,M.A. (1998) Abnormal maternal behaviour and growth retardation associated with loss of the imprinted gene *Mest*. *Nat. Genet*., **20**, 163–169.
8. Sansom,S.N., Hébert,J.M., Thammongkol,U., Smith,J., Nisbet,G., Surani,M., McConnell,S.K. and Livesey,F.J. (2005) Genomic characterisation of a Fgf-regulated gradient-based neocortical protomap. *Development*, **132**, 3947–3961.
9. Hahn,Y., Yang,S.K. and Chung,J.H. (2005) Structure and expression of the zebrafish mest gene, an ortholog of mammalian imprinted gene *PEG1*/*MEST*. *Biochim. Biophys. Acta*, **1731**, 125–132.
10. Suzuki,S., Renfree,M., Pask,A., Shaw,G., Kobayashi,S., Kohda,T., Kaneko-Ishino,T. and Ishino,F. (2005) Genomic imprinting of *IGF2, p57(KIP2)* and *PEG1*/*MEST* in a marsupial, the tammar wallaby. *Mech. Dev*., **122**, 213–222.
11. Mayer,W., Hemberger,M., Frank,H.G., Grümmer,R., Winterhager,E., Kaufmann,P. and Fundele,R. (2000) Expression of the imprinted genes *MEST*/*Mest* in human and murine placenta suggests a role in angiogenesis. *Dev. Dyn*., **217**, 1–10.
12. Tavazoie,S., Alarcón,C., Oskarsson,T., Padua,D., Wang,Q., Bos,P., Gerald,W. and Massagué,J. (2008) Endogenous human

microRNAs that suppress breast cancer metastasis. *Nature*, **451**, 147–152.
13. Lee,Y.J., Park,C.W., Hahn,Y., Park,J., Lee,J., Yun,J.H., Hyun,B. and Chung,J.H. (2000) *Mit1/Lb9* and *Copg2*, new members of mouse imprinted genes closely linked to *Peg1/Mest*. *FEBS Lett*., **472**, 230–234.
14. Riesewijk,A.M., Hu,L., Schulz,U., Tariverdian,G., Höglund,P., Kere,J., Ropers,H.H. and Kalscheuer,V.M. (1997) Monoallelic expression of human *PEG1/MEST* is paralleled by parent-specific methylation in fetuses. *Genomics*, **42**, 236–244.
15. Blagitko,N., Schulz,U., Schinzel,A.A., Ropers,H.H. and Kalscheuer,V.M. (1999) *gamma2-COP*, a novel imprinted gene on chromosome 7q32, defines a new imprinting cluster in the human genome. *Hum. Mol. Genet*., **8**, 2387–2396.
16. Yamasaki,K., Hayashida,S., Miura,K., Masuzaki,H., Ishimaru,T., Niikawa,N. and Kishino,T. (2000) The novel gene, *gamma2-COP* (COPG2), in the 7q32 imprinted domain escapes genomic imprinting. *Genomics*, **68**, 330–335.
17. Yun,J., Park,C.W., Lee,Y.J. and Chung,J.H. (2003) Allele-specific methylation at the promoter-associated CpG island of mouse *Copg2*. *Mamm. Genome*, **14**, 376–382.
18. Tada,M., Tada,T., Lefebvre,L., Barton,S.C. and Surani,M.A. (1997) Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells. *EMBO J*., **16**, 6510–6520.
19. Saha,S., Sparks,A.B., Rago,C., Akmaev,V., Wang,C.J., Vogelstein,B., Kinzler,K.W. and Velculescu,V.E. (2002) Using the transcriptome to annotate the genome. *Nat. Biotechnol*., **20**, 508–512.
20. Peters,D.G., Kassam,A.B., Yonas,H., O'Hare,E.H., Ferrell,R.E. and Brufsky,A.M. (1999) Comprehensive transcript analysis in small quantities of mRNA by SAGE-lite. *Nucleic Acids Res*., **27**, e39.
21. Siddiqui,A.S., Khattra,J., Delaney,A.D., Zhao,Y., Astell,C., Asano,J., Babakaiff,R., Barber,S., Beland,J., Bohacec,S. *et al*. (2005) A mouse atlas of gene expression: large-scale digital gene-expression profiles from precisely defined developing C57BL/6J mouse tissues and cells. *Proc. Natl Acad. Sci. USA*, **102**, 18485–18490.
22. Ramakers,C., Ruijter,J.M., Deprez,R.H. and Moorman,A.F. (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci. Lett*., **339**, 62–66.
23. Ewing,B., Hillier,L., Wendl,M.C. and Green,P. (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res*., **8**, 175–185.
24. Church,D.M., Goodstadt,L., Hillier,L., Zody,M., Goldstein,S., She,X., Bult,C., Agarwala,R., Cherry,J.L., DiCuccio,M. *et al*. (2009) Lineage-specific biology revealed by a finished genome assembly of the mouse. *Plos Biol*., **7**, e1000112.
25. Sado,T., Nakajima,N., Tada,M. and Takagi,N. (1993) A novel mesoderm-specific cDNA isolated from a mouse embryonal carcinoma cell line. *Dev. Growth Different*., **35**, 551–560.
26. Uejima,H., Lee,M.P., Cui,H. and Feinberg,A. (2000) Hot-stop PCR: a simple and general assay for linear quantitation of allele ratios. *Nat. Genet*., **25**, 375–376.
27. Tian,B., Hu,J., Zhang,H. and Lutz,C.S. (2005) A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res*., **33**, 201–212.
28. Ji,Z., Lee,J.Y., Pan,Z., Jiang,B. and Tian,B. (2009) Progressive lengthening of 3′ untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. *Proc. Natl Acad. Sci. USA*, **106**, 7028–7033.
29. Hagan,J.P., O'Neill,B.L., Stewart,C., Kozlov,S. and Croce,C.M. (2009) At least ten genes define the imprinted *Dlk1-Dio3* cluster on mouse chromosome 12qF1. *PLoS One*, **4**, e4352.
30. An,J.J., Gharami,K., Liao,G.Y., Woo,N.H., Lau,A.G., Vanevski,F., Torre,E.R., Jones,K.R., Feng,Y., Lu,B. *et al*. (2008) Distinct role of long 3′ UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. *Cell*, **134**, 175–187.
31. Irier,H.A., Shaw,R., Lau,A., Feng,Y. and Dingledine,R. (2009) Translational regulation of GluR2 mRNAs in rat hippocampus by alternative 3′ untranslated regions. *J. Neurochem*., **109**, 584–594.
32. Yu,M., Sha,H., Gao,Y., Zeng,H., Zhu,M. and Gao,X. (2006) Alternative 3′ UTR polyadenylation of *Bzw1* transcripts display differential translation efficiency and tissue-specific expression. *Biochem. Biophys. Res. Commun*., **345**, 479–485.
33. Edwalds-Gilbert,G., Veraldi,K.L. and Milcarek,C. (1997) Alternative poly(A) site selection in complex transcription units: means to an end? *Nucleic Acids Res*., **25**, 2547–2561.
34. Schulz,R., Woodfine,K., Menheniott,T.R., Bourc'his,D., Bestor,T. and Oakey,R. (2008) WAMIDEX: a web atlas of murine genomic imprinting and differential expression. *Epigenetics*, **3**, 89–96.
35. Prescott,E.M. and Proudfoot,N.J. (2002) Transcriptional collision between convergent genes in budding yeast. *Proc. Natl Acad. Sci. USA*, **99**, 8796–8801.
36. Scadden,A.D. and Smith,C.W. (2001) Specific cleavage of hyper-edited dsRNAs. *EMBO J*., **20**, 4243–4252.
37. Yang,W., Chendrimada,T.P., Wang,Q., Higuchi,M., Seeburg,P.H., Shiekhattar,R. and Nishikura,K. (2006) Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nat. Struct. Mol. Biol*., **13**, 13–21.
38. Mazo,A., Hodgson,J.W., Petruk,S., Sedkov,Y. and Brock,H. (2007) Transcriptional interference: an unexpected layer of complexity in gene regulation. *J. Cell Sci*., **120**, 2755–2761.
39. Proudfoot,N. (2004) New perspectives on connecting messenger RNA 3′ end formation to transcription. *Curr. Opin. Cell Biol*., **16**, 272–278.
40. Kotzot,D., Schmitt,S., Bernasconi,F., Robinson,W.P., Lurie,I.W., Ilyina,H., Méhes,K., Hamel,B.C., Otten,B.J. and Hergersberg,M. (1995) Uniparental disomy 7 in Silver-Russell syndrome and primordial growth retardation. *Hum. Mol. Genet*., **4**, 583–587.
41. Preece,M.A., Price,S.M., Davies,V., Clough,L., Stanier,P., Trembath,R.C. and Moore,G.E. (1997) Maternal uniparental disomy 7 in Silver-Russell syndrome *J. Med. Genet*., **34**, 6–9.
42. Riesewijk,A.M., Blagitko,N., Schinzel,A.A., Hu,L., Schulz,U., Hamel,B.C., Ropers,H.H. and Kalscheuer,V.M. (1998) Evidence against a major role of *PEG1/MEST* in Silver-Russell syndrome. *Eur. J. Hum. Genet*., **6**, 114–120.
43. Schöherr,N., Jäger,S., Ranke,M.B., Wollmann,H.A., Binder,G. and Eggermann,T. (2008) No evidence for isolated imprinting mutations in the *PEG1/MEST* locus in Silver-Russell patients. *Eur. J. Med. Genet*., **51**, 322–324.
44. Gregg,C., Zhang,J., Weissbourd,B., Luo,S., Schroth,G.P., Haig,D. and Dulac,C. (2010) High-resolution analysis of parent-of-origin allelic expression in the mouse brain. *Science*, **329**, 643–648.
45. Trapnell,C., Pachter,L. and Salzberg,S.L. (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*, **25**, 1105–1111.
46. Robinson,J.T., Thorvaldsdóttir,H., Winckler,W., Guttman,M., Lander,E., Getz,G. and Mesirov,J. (2011) Integrative genomics viewer. *Nat. Biotechnol*., **29**, 24–26.
47. Mamo,S., Gal,A.B., Bodo,S. and Dinnyes,A. (2007) Quantitative evaluation and selection of reference genes in mouse oocytes and embryos cultured in vivo and in vitro. *BMC Dev. Biol*., **7**, 14.
48. Umlauf,D., Goto,Y., Cao,R., Cerqueira,F., Wagschal,A., Zhang,Y. and Feil,R. (2004) Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat. Genet*., **36**, 1296–1300.