Regulation of alternative polyadenylation by genomic imprinting

Citation for published version:
Wood, A, Schulz, R, Woodfine, K, Koltowska, K, Beechey, CV, Peters, J, Bourc'his, D & Oakey, RJ 2008, 'Regulation of alternative polyadenylation by genomic imprinting' Genes & Development, vol 22, no. 9, pp. 1141-6. DOI: 10.1101/gad.473408

Digital Object Identifier (DOI):
10.1101/gad.473408

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Genes & Development

Publisher Rights Statement:
Copyright © 2008, Cold Spring Harbor Laboratory Press

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Regulation of alternative polyadenylation by genomic imprinting

Andrew J. Wood,1,4 Reiner Schulz,1 Kathryn Woodfine,1 Katarzyna Koltowska,1 Colin V. Beechey,2 Jo Peters,2 Deborah Bourc’his,3 and Rebecca J. Oakey1,5

1Department of Medical and Molecular Genetics, King’s College London, Guy’s Hospital, London SE1 9RT, United Kingdom; 2Medical Research Council Mammalian Genetics Unit, Harwell OX11 0RD, United Kingdom; 3Institut National de la Santé et de la Recherche Médicale (INSERM) U741, Institut Jacques Monod, 75251 Paris, Cedex 05, France

Maternally and paternally derived alleles can utilize different promoters, but allele-specific differences in co-transcriptional processes have not been reported. We show that alternative polyadenylation sites at a novel murine imprinted gene [H13] are utilized in an allele-specific manner. A differentially methylated CpG island separates polyA sites utilized on maternal and paternal alleles, and contains an internal promoter. Two genetic systems show that alleles lacking methylation generate truncated H13 transcripts that undergo internal polyadenylation. On methylated alleles, the internal promoter is inactive and elongation proceeds to downstream polyadenylation sites. This demonstrates that epigenetic modifications can influence utilization of alternative polyadenylation sites.

Supplemental material is available at http://www.genesdev.org.

Received January 28, 2008; revised version accepted March 10, 2008.

Transcription by RNA polymerase II [Pol II] requires multiple linked steps, including the assembly of an initiation complex, promoter release, elongation followed by splicing, polyadenylation [polyA], and dissociation of the polymerase complex from the template DNA. Each step provides an opportunity for the cell to regulate gene expression, either by changing the abundance or availability of mRNAs or by generating variant gene products. Epigenetic modifications to DNA and histones can influence transcription at the initiation stage by altering the accessibility of promoter sequences to initiation complex components [Kass et al. 1997a,b]. Following initiation, regions of heterochromatin can subtly impede the progress of an elongating polymerase complex, reducing transcriptional output [Lorincz et al. 2004]. However, it is not clear whether epigenetic modifications downstream from a promoter can lead to variant gene products through alternative polyadenylation or splicing.

Polyadenylation, the addition of multiple adeny1 residues to the 3′ end of a newly synthesized transcript, is required for nuclear export [Huang and Carmichael 1996; Jacobson and Peltz 1996]. This occurs following cleavage of the newly synthesized pre-mRNA, typically 15–30 nucleotides downstream from a conserved hexamer motif [usually AUUAAA or AUUAGA] [Proudfoot and Brownlee 1976; Zhao et al. 1999]. A large proportion of human genes utilize more than one polyadenylation [polyA] site [Tian et al. 2005], making alternative polyadenylation a major source of transcriptional diversity. The mechanisms governing alternative polyA site selection in mammalian cells have been studied in detail at only a small number of loci; notably the immunoglobulin heavy chain [Takagaki and Manley 1998] and calcitonin/CGRP genes [Lou et al. 1998]. In both cases, polyA site selection is cell type-specific and is regulated by changes in the concentration of diffusible RNA processing factors [Lou et al. 1998; Takagaki and Manley 1998].

Imprinted genes are differentially expressed on maternally and paternally derived alleles and are estimated to comprise 1%–2% of the mammalian transcriptome [Brown 2006; Wood and Oakey 2006]. These genes are typically clustered into regions of the genome containing multiple imprinted transcripts, which share common cis-regulatory elements. Such elements undergo differential CpG methylation when passing through the male and female germline. Allele-specific markings are maintained during embryonic development and can influence transcription of adjacent genes via cis-acting regulatory RNAs such as Air [Sleutels et al. 2002] or insulator factors such as CTCF [Hark et al. 2000].

Imprinted loci represent an attractive model for studying epigenetic mechanisms of transcriptional control, as both active and silent alleles are present in the same cellular environment. In the presence of an identical complement of trans-regulatory factors, allelic differences in transcription must be the result of epigenetic regulation. While the allele-specific use of alternative promoters has been well documented at imprinted loci [Blagitko et al. 2000; Holmes et al. 2003], allele-specific differences in cotranscriptional events such as splicing and polyadenylation have not been reported previously.

We undertook a computational screen for imprinted genes, identifying a novel imprinted domain on mouse Chromosome 2 containing at least two protein-coding genes: H13 and Mcts2 [Wood et al. 2007]. Mcts2 is a duplicate of the X-linked Mcts1 gene that originated by retrotransposition of an mRNA from the X-linked locus into the fourth intron of H13 [Wood et al. 2007]. Unlike most retrotransposed gene copies, Mcts2 has maintained the capacity to encode a protein and shares 94% amino acid identity with the X-linked paralog >65 million years after the gene duplication event. The promoter of Mcts2 is associated with a CpG island that acquires DNA methylation in the female germline, and the retrogene is expressed primarily from paternally derived alleles in the developing embryo [Wood et al. 2007].

We identify five polyA sites for H13 and show that at least three are utilized in an allele-specific manner. This is likely to depend on allele-specific DNA methylation.
at the internal Mcts2 promoter, which is established during maternal gametogenesis and serves to promote functional H13 products from maternally derived alleles in the embryo. These findings provide strong evidence that epigenetic modifications can influence alternative polyadenylation, and pave the way for future studies to determine precisely how this influence can be exerted.

Results and Discussion

Alternative polyadenylation at the H13 locus

H13 encodes signal peptide peptidase, an intramembrane aspartic protease responsible for the dislocation of signal peptides from the endoplasmic reticulum membrane (Weihofen et al. 2002). The ORF is spread over 12 exons (Fig. 1A) and encodes a protein consisting of seven transmembrane helices (Fig. 1B). EST evidence suggests a number of transcript variants. Northern blots (Fig. 1C) and 5' and 3' RACE were performed in material extracted from whole brain tissues of neonatal mice, revealing one promoter and five distinct polyA sites (Fig. 1A). Although alternative H13 promoters cannot be absolutely excluded, all transcript species identified by Northern analysis (Fig. 1C) can be accounted for by different 3' RACE products. Three polyA sites are clustered within a single block of 3' untranslated region (UTR) situated 20 kb downstream from the monoexonic Mcts2 gene [H13a–c]. The alternatively skipped penultimate exon of the 2-kb H13a transcript contains a premature stop codon that removes a C-terminal KKXX membrane retention signal from full-length signal peptide peptidase proteins, altering subcellular localization patterns (Urny et al. 2006). Two internal polyA sites lie 0.5 kb (H13d) and 3 kb (H13e) upstream of the Mcts2 transcriptional start site (Fig. 1A). Biochemical characterization of the protein product of H13 has previously demonstrated that the catalytic residues essential for signal peptide peptidase activity lie within the region of the ORF downstream from the H13d and H13e polyA sites (Weihofen et al. 2002). Hence, only transcripts extending beyond Mcts2 and the CpG island [H13a–c] can encode functional signal peptide peptidase proteins (Fig. 1B). H13d

---

**Figure 1.** Gene structure and CpG methylation at the murine H13/Mcts2 locus. [A] Transcript map of H13 isoforms, determined by 5' and 3' RACE, Northern, and GenBank EST evidence. The bar indicates nucleotide position on mouse Chromosome 2 (Mm. build 37). Representative GenBank accession numbers are as follows: H13a 1.8 kb, AJ345032; H13a 2.0 kb, AK049101; H13b, NM_010376; H13c, F830002G13; H13d, 4021402J09; H13e, AK004690; Mcts2, NM_025543. Exons are black rectangles and splice patterns are indicated. Vertical arrows indicate polyA sites, and horizontal arrows show the direction of transcription. The greater abundance of H13a prevented the simultaneous detection of H13a, H13b, and H13c in a single 3' RACE reaction. Consequently, we did not determine whether the alternatively skipped penultimate exon is also included in H13b or H13c transcripts. The positions of Northern probes for C are shown at the bottom, as are the three CpG islands at which methylation was investigated in D. [B] Predicted structure of proteins translated from H13 transcript isoforms. Cylinders represent transmembrane (TM) helices. Residues critical for signal peptide peptidase activity are indicated within transmembrane helices 4 and 5. Distinct C termini for H13a–c protein isoforms arise from the inclusion or exclusion of a penultimate exon (see the text). Membrane topology for the larger protein was determined by Weihofen et al. (2002). The proteins putatively translated from H13d and H13e have not been described previously; hence, membrane topology is not shown. [C] Northerns performed with neonatal brain mRNA identify transcripts utilizing all five polyA sites identified by 3' RACE. [D] DNA methylation at three CpG islands: the H13 promoter, the Mcts2 promoter, and the H13 downstream. Horizontal lines denote individual DNA strands on which filled circles represent methylated CpG dinucleotides; open circles are unmethylated CpGs. Genomic DNA was extracted from whole brain of neonatal mice carrying uniparental partial duplications for distal Chromosome 2 [MatDp(dist2) and PatDp(dist2)].
and H13e have the potential to encode a protein of unknown function, consisting of the three N-terminal transmembrane helices of signal peptide peptidease with 51 unique C-terminal residues [Fig. 1B].

Allele-specific CpG methylation

Three CpG islands lie in the region: one at the intronic Mcts2 promoter, another at the H13 promoter, and a third separating the H13b and H13c polyA sites [Fig. 1A]. To investigate the allele-specific methylation profile of these three regions, we sequenced bisulphite-modified genomic DNA from whole neonatal brains of mice carrying uniparental partial disomies for distal Chromosome 2 [MatDp(dist2) and PatDp(dist2) methods]. At the Mcts2 promoter, methylation is present on maternally derived alleles but absent on paternally derived alleles [Fig. 1D]. We have shown previously that this region undergoes methylation during oogenesis but is unmethylated in sperm [Wood et al. 2007], suggesting that the somatic methylation pattern is inherited from the gametes. The CpG island at the H13 promoter is predominantly unmethylated on both alleles in neonatal brain [Fig. 1D], as is the CpG island downstream from the H13c polyadenylation site. It is not presently clear whether transcription initiates from the 3′ CpG island.

Allele-specific polyadenylation in intersubspecies hybrid mice

The differentially methylated region (DMR) at the Mcts2 promoter lies ~200 base pairs (bp) downstream from the H13d polyA site [Fig. 2A], and ~3.5 kb downstream from H13e [Supplemental Fig. 1]. The proximity of these two sites to the DMR raised the possibility that they could be differentially utilized on maternally and paternally derived alleles. To address this, 3′ RACE was performed in neonatal brain material from intersubspecies hybrid mice. Following gel purification, 3′ RACE products were sequenced over regions containing SNPs between the parental strains, revealing the relative utilization of each polyA site by each parental allele [Fig. 2B,C, Supplemental Fig. 1]. Transcripts terminating upstream of the Mcts2 DMR [H13d and H13e] originate preferentially from paternally derived alleles [Fig. 2C, Supplemental Fig. 1]. In contrast, the H13a polyA site downstream from the DMR is utilized predominantly on maternally derived alleles [Fig. 2C]. This is observed in cDNA prepared from whole neonatal brains and from embryonic stem (ES) cells [Fig. 2C]. RT–PCR amplification of exons 1–3 yields products that originate from both parental alleles [Fig. 2D], consistent with biallelic transcriptional initiation from the unmethylated H13 promoter [Fig. 1C]. The allele-specific utilization of polyA sites within a single cell type indicates that polyadenylation can be regulated by epigenetic mechanisms.

Disrupted polyadenylation following uniparental inheritance

To confirm these findings in an independent genetic system, polyadenylation at H13 was studied in mice carrying uniparental partial disomies for distal Chromosome 2. This allowed the expression of H13 transcripts to be compared in mice that inherit two copies of the H13 locus from their mothers [MatDp(dist2)] versus mice that inherit two paternal copies of this genomic region [PatDp(dist2)]. Northern blots were generated from total brain RNA and hybridized with a probe to the common 5′ exons [Fig. 1A, probe #3]. This confirmed the preferential maternal expression of transcripts utilizing the H13a and H13c polyA signals, but H13b transcripts were not detectable on this genetic background [C3H × 101[F1]] [Fig. 3A]. Transcripts utilizing both the H13d and H13e polyA signals are present at greater levels following pa-
Collectively, these data indicate that polyA sites situated 0.2 kb \(H13d\) and 3 kb \(H13e\) upstream of the \(Mcts2\) DMR are utilized primarily on unmethylated paternal alleles. Maternal alleles preferentially generate \(H13\) transcripts that elongate through the methylated and transcriptionally inactive \(Mcts2\) promoter to the three polyA sites located \(-20\) kb downstream \(H13a\–c\).

Alternative polyadenylation is controlled by maternal germline methylation

To further investigate the role of allele-specific methylation in this phenomenon, embryos lacking maternally derived methylation imprints were generated by crossing \(Dnmt3l^{-/-}\) mothers with \(cast\) fathers (Bourc’his et al. 2001). Heterozygous offspring lack allele-specific DNA methylation at the \(Mcts2\) promoter (Fig. 4A), and this is associated with activation of the normally silent maternal allele (Fig. 4B). Qualitative 3′ RACE for the \(H13\) gene demonstrates that loss of intronic methylation is associated with increased utilization of the \(H13d\) polyA site in \(Dnmt3l^{-/-}\) embryos (Fig. 4C). This is entirely consistent with \(PatDp\(\text{dist2}\) individuals, which also exhibit biallelic hypomethylation at this region (Fig. 3; data not shown). Despite up-regulation of \(H13d\) relative to \(H13a\) maternally derived methylation imprints were generated by crossing \(Dnmt3l^{-/-}\) mothers with \(cast\) fathers (Bourc’his et al. 2001). Heterozygous offspring lack allele-specific DNA methylation at the \(Mcts2\) promoter (Fig. 4A), and this is associated with activation of the normally silent maternal allele (Fig. 4B). Qualitative 3′ RACE for the \(H13\) gene demonstrates that loss of intronic methylation is associated with increased utilization of the \(H13d\) polyA site in \(Dnmt3l^{-/-}\) embryos (Fig. 4C). This is entirely consistent with \(PatDp\(\text{dist2}\) individuals, which also exhibit biallelic hypomethylation at this region (Fig. 3; data not shown). Despite up-regulation of \(H13d\) relative to \(H13a\)
in the Dnmt3l-/- embryos, both products were generated at levels sufficient to assess their allele-specific expression pattern by sequencing. The H13a and H13d polyA sites show reciprocal allele-specific biases in wild-type embryonic day 8.5 (E8.5) embryos but are utilized at approximately equal levels on each allele in Dnmt3l-/- individuals [Fig. 4D]. Thus, alternative polyadenylation in the embryo is controlled by DNA methylation marks that are established during oogenesis. Furthermore, biallelic hypomethylation in Dnmt3l-/- embryos and PatDp(dist2) neonates decreases the output of functional signal peptide peptidase transcripts (H13a–c) [Figs. 4C, 3C].

In this report, we identified multiple polyA sites for the murine signal peptide peptidase gene [H13] and showed that at least three are utilized in an allele-specific manner. Previous in vivo studies of alternative polyadenylation mechanisms in mammals have focused on tissue-specific variation, which can be explained on the basis of tissue-specific differences in the concentration of trans-regulatory factors. Such explanations are insufficient to account for polyA site selection at the H13 locus, as maternally and paternally derived alleles in ES cells are potentially exposed to the same array of diffusible molecules. Hence, allele-specific epigenetic differences must play a role. Consistent with this conclusion, a differentially methylated CpG island and promoter were identified in the region separating maternally and paternally utilized polyA sites. Hypomethylation within this region correlates with the use of polyA sites 0.5 and 4 kb upstream, whereas hypermethylated alleles preferentially utilize sites between 20 and 25 kb downstream. Several possible explanations could account for these observations, three of which are outlined in Figure 5.

On paternally derived alleles, silencing of H13a–c is not complete. Transcriptional output was approximately sevenfold lower relative to maternal alleles [Fig. 3C], but still readily detectable by Northern blot [Fig. 3A]. In a previous report, we examined the allele-specific expression of H13a–c (exons 3–12) using qualitative methodology and found some degree of expression from the paternal allele in whole embryo and placental tissues [Wood et al. 2007]. It will be important to determine whether these observations reflect cell type-specific imprinting or low level expression from the paternal allele across all cell types of these complex tissues. Silencing of H13d on maternal alleles is more absolute, showing 30-fold lower expression levels than paternal alleles in neonatal brain.

In summary, we identified a novel mechanism of imprinted gene regulation acting at the level of alternative polyadenylation. The imprinted H13 locus has the potential to provide novel insights into this important transcriptional phenomenon. Further work is now required to elucidate the precise mechanism by which identical DNA sequences can generate alternatively polyadenylated transcripts within a single cellular environment.

Materials and methods

Tissue sources

Mice with uniparental partial disomies (or uniparental duplications) were generated by intercrossing heterozygotes for the Tj2-911H reciprocal translocation [Cattanach and Beechey 1997]. Mice carrying maternal duplications for distal Chromosome 2 [MatDp(dist2)] also carry paternal duplications for distal Chromosome 9, and vice versa. Maternal imprint-free embryos were generated as described previously [Bourc'his et al. 2001]. RNA from the hybrid ES cell line used in Figure 2 was a gift from Wolf Reik.

DNA methylation

RNA and DNA were prepared using the RNaseasy and DNeasy kits (Qia-GEN) primed with oligo dT. Three microliters of total RNA were used for a 50-µL final RT reaction volume, of which 2 µL were used per 10 µL of QPCR reaction. Expression of H13 isoforms was measured relative to...
β‐actin, which did not change significantly between MatDp(dist2) and PatDp(dist2) samples measured by microarray (Supplemental Fig. 2). Sample values were determined using the relative expression ratio mathematical model [Pfaffl 2001], with Ct values transformed into expression ratios according to the formula

\[ \text{Ratio} = 2^{-\Delta \Delta C_t} \]

where targetbaseline denotes the mean of the Ct values measured for the target gene in the baseline sample, controlbaseline is the equivalent value for the endogenous control gene, and targetsample and controlsample are the mean Ct values determined in the nonbaseline sample.

Acknowledgments

We thank Woli Reik and En Li for hybrid ES cell RNA. We thank Nick Proudfoot for useful discussions. This work was supported by the Wellcome Trust (R.J.O.), the Generation Trust (A.J.W.), the Biotechnology and Biological Sciences Research Council (R.J.O.), and The European Molecular Biology Organization (R.S.).

References

Barlow, D.P. 1995. Gametic imprinting in mammals. Science 270: 1610–1613.

Blagitko, N., Mengenthaler, S., Schulz, U., Wollmann, H.A., Craigen, W., Eggermann, T., Ropers, H.H., and Kalscheuer, V.M. 2000. Human GRR10 is imprinted and expressed from the paternal and maternal allele in a highly tissue- and isoform-specific fashion. Hum. Mol. Genet. 9: 1587–1595.

Bourc’his, D., Xu, G.-L., Lin, C.-S., Rollman, B., and Bestor, T.H. 2001. Dnmt3L and the establishment of maternal genomic imprints. Science 294: 2536–2539.

Cattanach, B.M. and Beechey, C.V. 1997. Genomic imprinting in the mouse: Possible final analysis. In Genomic imprinting: Frontiers in molecular biology [eds. W. Reik and A. Surani], pp. 118–145. IRL Press, Oxford, New York, Tokyo.

Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature 405: 486–489.

Holmes, R., Williamson, C., Peters, J., Denny, P., Wells, C., Group, R.G., and Members, G. 2003. A comprehensive transcript map of the mouse Gnai imprinted complex. Genome Res. 13: 1410–1415.

Huang, Y. and Carmichael, G.G. 1996. Role of polyadenylation in nucleocytoplasmic transport of mRNA. Mol. Cell. Biol. 16: 1534–1542.

Jacobson, A. and Peltz, S.W. 1996. Interrelationships of the pathways of mRNA decay and translation in eukaryotic cells. Annu. Rev. Biochem. 65: 693–739.

Kass, S.U., Landsberger, N., and Wolfe, A.P. 1997a. DNA methylation directs a time‐dependent repression of transcription initiation. Curr. Biol. 7: 157–165.

Kass, S.U., Pruss, D., and Wolfe, A.P. 1997b. How does DNA methylation repress transcription? Trends Genet. 13: 444–449.

Lorincz, M.C., Dickerson, D.R., Schmitt, M., and Groudine, M. 2004. Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. Nat. Struct. Mol. Biol. 11: 1068–1075.

Lou, H., Neugebauer, K.M., Gagel, R.F., and Berget, S.M. 1998. Regulation of alternative polyadenylation by U1 snRNPs and SRp20. Mol. Cell. Biol. 18: 4977–4985.

Pfaffl, M.W. 2001. A new mathematical model for relative quantification in real‐time RT‐PCR. Nucleic Acids Res. 29: e45. doi: 10.1093/nar/29.9.e45.

Proudfoot, N.J. and Brownlee, G.G. 1976. 3’ non‐coding region sequences in eukaryotic messenger RNA. Nature 263: 211–214.

Shearwin, K.E., Callen, B.P., and Egan, J.B. 2005. Transcriptional interference—A crash course. Trends Genet. 21: 339–345.

Sleutels, F., Zwart, R., and Barlow, D.P. 2002. The non‐coding Air RNA is required for silencing autosomal imprinted genes. Nature 415: 810–813.

Takagaki, Y. and Manley, J.L. 1998. Levels of polyadenylation factor CstF64 control IgM heavy chain mRNA accumulation and other events associated with B cell differentiation. Mol. Cell. Biol. 2: 761–771.