Supplementary Material

The functional consequences of alternative promoter use in mammalian genomes

Ramana V Davuluri[1], Yutaka Suzuki[2], Sumio Sugano[2], Christoph Plass[1] and Tim H.-M. Huang[1]

[1]Human Cancer Genetics Program, Comprehensive Cancer Center, Department of Molecular Virology, Immunology, and Medical Genetics, The Ohio State University, Columbus, Ohio, 43210, USA
[2]Department of Medical Genome Sciences, Graduate School of Frontier Sciences, The University of Tokyo: 301 LS Bldg, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8562, Japan
Corresponding author: Davuluri, R (ramana.davuluri@osumc.edu).

Supplementary Table 1. List of human and mouse alternative promoters obtained from published literature

Gene Symbol	TSS location of the promoter (strand)	Summary of expression pattern
IGF2	P1 – chr11:2127509 (-) P2 – chr11:2116880 (-) P3 – chr11:2115178 (-) P4 – chr11:2111580 (-)	The human insulin-like growth factor-II (IGF2) gene, which is transcribed from four promoters, P1-P4, is imprinted in fetal liver but biallelic expression occurs in adult liver. Fetal liver uses primarily promoters P3 and P4, however, adult liver transcribes IGF2 from promoter P1. It was reported that in liver and chondrocytes, IGF2 transcripts from promoter P1 were always derived from both parental alleles, whereas transcripts from promoters P2, P3 and P4 were always from one parental allele[1], demonstrating imprinting and a lack of imprinting can both occur within a single gene in a single tissue.
HTR3B	P1 – chr11:113280799 (+) P2 – chr11:113285165 (+)	The HTR3B gene codes for the subunit B of the serotonin receptor type 3. Two alternative promoters control the expression of different HTR3B transcripts in the peripheral and central nervous system. The transcription start site of P1 has been found to control the gene expression in the brain and the transcription start sites of promoter P2 has been observed in the intestine [2].
SLC7A7	P1 – chr14:22358798(-) P2 – chr14:22355516(-)	The human SLC7A7 gene is mainly expressed at the basolateral membrane of the polarised epithelial cells in the renal tubules and the small intestine. The alternative promoter usage is the mechanism for SLC7A7 gene differentially expressed in brain (P1), and other tissues in small intestine and kidney (P2) [3]
HRH1	P1 – chr3:11153799(+) P2 – chr3:11171214(+) P3 – chr3:11242364(+)	Three separate promoters lead to human histamine H1 receptor (HRH1) gene differentially expressed in primary cultured human airway smooth muscle (HASM) cells (P1), primary cultured human bronchial epithelial cells and bronchial epithelial cell line [BEAS2B]), and other tissues (brain) known to express histamine H1 receptors (P1, and P3) [4].
RUNX1	P1 – Chr21:36298835 (-) P2 – Chr21:35343501 (-) P3 – Chr21:35182857 (-)	Alternatives promoters transcribe many mRNA isoforms that are differentially expressed [5]. P1 transcribes two isoforms (b and c), leading to the production of two distinct proteins with a variety of biological functions. The switching of the promoters controls the expression of RUNX1 during embryonic hematopoiesis [6]. Disruption of P2 by the 12;21 chromosomal translocation results

		in the most common subtype of childhood acute lymphoblastic leukaemia [7]
RUNX2	P1 – chr6:45404032 (+) P2 – chr6:45497892 (+)	Predominantly expressed in bone, colon, heart, tonsil, head/neck, lung, and ovary. Expression was also reported in B and T cells of various developmental stages [8]. P1 transcribes two isoforms (a and b), leading to the production of two distinct proteins with a variety of biological functions. The isoform (c) transcribed by P2 encodes a protein with a shorter and distinct N-terminus when it is compared to isoform a. Aberrant expression was also reported in some tumors (e.g. adenocarcinoma, colon tumor).
RUNX3	P1 – chr1:25164088 (-) P2 – chr1:25129357 (-)	RUNX3, predominantly expressed in hematopoietic cells, is a tumor suppressor gene that is frequently deleted or transcriptionally silenced in cancer. Multiple transcript variants driven by two distinct promoters encoding different isoforms have been found for this gene [8].
RGS4	P1 – chr1:161305775 (+) P2 – chr1:161305559 (+) P3 – chr1:161305189 (+) P4 – chr1:161308319 (+)	RGS4 gene expression in the human brain is spatially and temporally regulated in dorsolateral prefrontal and visual cortex, through differential transcription of five different isoforms from four alternative promoters [9].
FMO1	P1 – chr1:169484234 (+) P2 – chr1:169493501 (+) P3 – chr1:169493720 (+)	Use of three alternative promoters regulate the species-dependent tissue-specific transcription of FMO1 in human and mouse [10]. In humans expression of the FMO1 gene is silenced postnatally in liver, but not in kidney. The transcription of the gene in fetal human liver is exclusively from the P1 promoter, whereas in extra-hepatic tissues of both species, P2 and P3 are active. [10]
PDE4B	P1 – chr1:66030781(+) P2 – chr1:66031281 (+) P3 – chr1:66230978 (+) P4 – chr1:66390581 (+) P5 – chr1:66568999 (+) P6 – chr1:66592649 (+) P7 – chr1:66598697 (+)	The four PDE4 (cAMP-specific phosphodiesterase-4) genes, targets of several potential selective therapeutic inhibitors, generate several distinct protein-coding isoforms through the use of alternative promoters and 5′-coding exons. PDE4B, linked to schizophrenia in humans, transcribes many isoforms driven by at least seven distinct promoters [11].
PRL	P1 – chr6:22411061 (-) P2 – chr6:22405709 (-)	The pituitary hormone prolactin (PRL), best known for its role in the regulation of lactation, is transcribed by two different promoters that regulate pituitary versus extrapituitary expression of prolactin in primates [12].
TEX101	P1– chr19:48584603 (+) P2– chr19:48587253 (+) P3– chr19:48610875 (+)	TEX101 transcribes three major isoforms regulated by distinct alternative promoters and usage of three 5′-untranslated first exons [13].
GR	P1- chr5:142795270 (-) P2 - chr5:142764238 (-) P3 - chr5:142763805 (-) P4 - chr5:142763447 (-) P5 - chr5:142763347 (-) P6 - chr5:142762495 (-) P7 - chr5:142760610 (-)	Alternative first exons each under the control of specific transcription factors control both the tissue specific *GR* expression and are involved in the tissue specific *GR* transcriptional response to stimulation [14].
BBOX1	P1 - chr11:27019085 (+) P2 - chr11:27033423 (+) P3 - chr11:27033501 (+)	The transcription initiation of the human BBOX1 gene might occur at 3 different exons, and that the expression level of each type of transcript is organ-specific [15].
KLK11	P1- chr19:56223102 (-) P2- chr19:56222697 (-) P3- chr19:56221682 (-)	Tissue-specific use of multiple promoters regulates the expression and intracellular trafficking of KLK11/hippostasin isoforms [16].
IKBKG	P1- chrX:153429034 (+) P2- chrX:153429256 (+)	Two alternative first exons, one is housekeeping required for proper expression and the other is active in cells of hepatic origin at a tissue-specific site[17].
ART3	P1- chr4:77222712 (-) P2- chr4:77252962 (+)	ART3 expression in human macrophages, testis, semen, tonsil, heart and skeletal muscle appears to be governed by a combination of differential splicing and tissue-preferential use of two alternative promoters [18].
CDC2	P1- chr10:62208255 (+) P2- chr10:62208142 (+)	In humans two transcripts exist for *CDC2*, one including and one excluding the untranslated first exon, that both result in the same protein [19]

PTHrP	P1- chr12:28016283 (-) P2- chr12:28014261 (-) P3- chr12:28013694 (-)	Three alternative promoters have been found in this gene. P3-initiated transcripts were detectable in most tumors, whereas transcripts initiated by either P1 or P2 were present in only a subset of tumors [20]
NAT1	P1- chr8:18112973 (+) P3- chr8:18111794 (+)	Most mRNAs of NAT1 gene originate at a promoter, P1, an alternative NAT1 promoter designated P3, to be most active in specific tissues, including kidney, liver, lung, and trachea. [21]
CD36	P1- chr7:79836727(+) P2- chr7:80069359(+) P3- chr7:80105793(+) P4- chr7:80113471(+) P5- chr7:80113614(+)	CD36 gene has 5 alternative first exons. The alternative transcripts are all expressed in more than one human tissue and their expression patterns vary highly in skeletal muscle, heart, liver, adipose tissue, placenta, spinal cord, cerebrum and monocytes. [22]
AFP	P1- chr4:74515619(+) P2- chr4:74520697(+)	Like the traditional AFP promoter (P1), the alternative promoter (P2) is active in the yolk sac and fetal liver and contributes to early expression of the AFP gene.[23]
AQP4	P1- chr18: 22699814(-) P2- chr18: 22696673(-)	The aquaporin-4 (AQP4) gene encodes two proteins isoforms. Both protein isoforms are expressed in brain, whereas mainly the smaller isoform is found in other tissues. However differential transcriptional regulation and tissue-specific factors regulate their relative expression by using alternative promoters. [24]
Ppp1r3b (mouse)	P1 – chr8:36438765 (+) P2 – chr8:36439705 (+)	Ppp1r3b utilizes two alternative promoters and non-coding first exons, which produce at least three alternatively spliced transcripts that encode identical proteins. All three transcripts are uniformly expressed in the liver, heart, and fetal lung; but uses distal and proximal promoters to differentially express in developing mouse airways [25].
MITF	P1-chr3:69871323 Isoform-A P2-chr3:69895652 Isoform-C P3-chr3:69998132 Isoform-B P4-chr3:70010946 Isoform-H P5-chr3:70068443 Isoform-M	MITF gene consists of 4 widely spaced multiple promoters, which generate not only the diversity in the transcriptional regulation of these promoters but also the structurally different isoforms. The 5′-flanking regions of these isoform-specific exons are termed A, H, B, and M promoters, respectively. Among these promoters, the M promoter has received particular attention, because it is functional only in melanocyte-lineage cells and is upregulated by Wnt signaling via the functional LEF-1-binding site. In contrast to MITF-M, other MITF isoforms are widely expressed in many cell types [26]
AC133		Transcription of *AC133* (human stem cell surface protein gene) isoforms is controlled by 5 different alternative promoters in a tissue-dependent manner, where exon 1A-containing *AC133* transcript was specifically expressed in human CD34+ cord blood cells [27].
Bcor (mouse)	P1-chrX:11737481(-) P2-chrX:11657180(-) P3-chrX:11656994(-)	Each promoter appears to be used at similar levels in all tissues tested (Ovary, eye, spleen, blood, testis, lung, kidney, adipose, small intestine, heart, liver, muscle, stomach, brain), with the exception of whole blood, which appears to preferentially use promoter 3 relative to other tissues[28].
Wnk1 (mouse)	P1-chr6:119874356–119987797(-) P2-chr6:119874356–119877279(-)	P1 is expressed mostly in heart, muscle, and brain. P2 is produced mostly in kidney[29]
Ntrk2 (mouse)	P1-chr13:58909194–59231328(+) P2-chr13:58907957–59231328(+)	The mouse neurotrophin receptor trkB gene is transcribed from two different promoters[30]
Bcl2l1 (mouse)	P1-chr2:152655877(-) P2-chr2:152656528(-) P3-chr2:152657612(-) P4-chr2:152658447(-) P5-chr2:152659138(-)	P1 and P2 are active in all tissues analyzed (uterus, spleen, heart, liver), whereas the other three promoter show tissue-specific activities. P3 is active in spleen, liver, and kidney, P4 is active in uterus and spleen, and P5 is active in spleen, liver, brain, and thymus[31].
Olfm3 (mouse)	P1-chr3:114607281 (-) P2-chr3:114783883 (-)	In the mouse brain and retina, only P1 is actively used. P2 is expressed in the combined tissues of the eye angle (trabecular mesh work, iris, and ciliary body), although not strongly[32].
Mtap1a (mouse)	P1-chr2:121115336(+) P2-chr2:121121199(+)	P1 and P2 transcripts are expressed abundantly in brain. However, expression of P2 transcript was not restricted in a cell- or tissue-specific manner[33].

Abcg2 (mouse)	P1-chr6:58546566 P2-chr6:58590440 P3-chr6:58600446	*ABCG2,* highly expressed in hematopoietic stem cells, encodes a transmembrane transporter associated with multidrug resistance in various cancer cells. The expression of Abcg2 during hematopoiesis is transcriptionally regulated by alternative use of three first exons and promoters in a developmental stage-specific manner[34].

Column 1 provides the standard gene symbol, column 2 provides the TSS location of the promoter, and column 3 provides a summary of gene expression driven by alternative promoter usage. The numbering of promoters is from 5′ farthest to closest of TSS, such that P1 is the most distal promoter and P4 is the closest for *IGF2.* (Human genome coordinates – March 2006 assembly (NCBI Build 36.1); Mouse genome coordinates – July 2007 assembly (NCBI Build 37))

References

1 Vu, T.H. and Hoffman, A.R. (1994) Promoter-specific imprinting of the human insulin-like growth factor-II gene. *Nature* 371, 714–717

2 Tzvetkov, M.V. *et al.* (2007) Tissue-specific alternative promoters of the serotonin receptor gene HTR3B in human brain and intestine. *Gene* 386, 52–62

3 Puomila, K. *et al.* (2007) Two alternative promoters regulate the expression of lysinuric protein intolerance gene SLC7A7. *Mol. Genet. Metab.* 90, 298–306

4 Swan, C. *et al.* (2006) Alternative promoter use and splice variation in the human histamine H1 receptor gene. *Am. J. Respir. Cell Mol. Biol.* 35, 118–126

5 Levanon, D. *et al.* (1996) A large variety of alternatively spliced and differentially expressed mRNAs are encoded by the human acute myeloid leukemia gene AML1. *DNA Cell Biol.* 15, 175–185

6 Pozner, A. *et al.* (2007) Developmentally regulated promoter-switch transcriptionally controls Runx1 function during embryonic hematopoiesis. *BMC Dev. Biol.* 7, 84

7 Pui, C.H. *et al.* (2001) Childhood acute lymphoblastic leukaemia–current status and future perspectives. *Lancet Oncol.* 2, 597–607

8 Okumura, A.J. *et al.* (2007) Expression of AML/Runx and ETO/MTG family members during hematopoietic differentiation of embryonic stem cells. *Exp. Hematol.* 35, 978–988

9 Ding, L. *et al.* (2007) Full length cloning and expression analysis of splice variants of regulator of G-protein signaling RGS4 in human and murine brain. *Gene*, 401, 46–60

10 Shephard, E.A. *et al.* (2007) Alternative promoters and repetitive DNA elements define the species-dependent tissue-specific expression of the FMO1 gene of human and mouse. *Biochem. J.*, 406, 491–499

11 Cheung, Y.F. *et al.* (2007) PDE4B5, a novel, super-short, brain-specific cAMP phosphodiesterase-4 variant whose isoform-specifying N-terminal region is identical to that of cAMP phosphodiesterase-4D6 (PDE4D6). *J. Pharmacol. Exp. Ther.* 322, 600–609

12 Gerlo, S. *et al.* (2006) Prolactin in man: a tale of two promoters. *Bioessays* 28, 1051–1055

13 Tsukamoto, H. *et al.* (2007) Genomic organization and structure of the 5′-flanking region of the TEX101 gene: alternative promoter usage and splicing generate transcript variants with distinct 5′-untranslated region. *Mol. Reprod. Dev.* 74, 154–162

14 Turner, J.D. *et al.* (2006) Tissue specific glucocorticoid receptor expression, a role for alternative first exon usage? *Biochem. Pharmacol.* 72, 1529–1537

15 Rigault, C. *et al.* (2006) Genomic structure, alternative maturation and tissue expression of the human BBOX1 gene. *Biochim. Biophys. Acta* 1761, 1469–1481

16 Mitsui, S. *et al.* (2006) Multiple promoters regulate tissue-specific alternative splicing of the human kallikrein gene, KLK11/hippostasin. *FEBS J.* 273, 3678–3686

17 Fusco, F. *et al.* (2006) Multiple regulatory regions and tissue-specific transcription initiation mediate the expression of NEMO/IKKgamma gene. *Gene* 383, 99–107

18 Friedrich, M. *et al.* (2006) Genomic organization and expression of the human mono-ADP-ribosyltransferase ART3 gene. *Biochim. Biophys. Acta* 1759, 270–280

19 Veerla, S. and Hoglund, M. (2006) Analysis of promoter regions of co-expressed genes identified by microarray analysis. *BMC Bioinformatics* 7, 384

20 Richard, V. *et al.* (2003) Quantitative evaluation of alternative promoter usage and 3′ splice variants for parathyroid hormone-related protein by real-time reverse transcription-PCR. *Clin. Chem.* 49, 1398–1402

21 Husain, A. *et al.* (2007) Functional analysis of the human N-acetyltransferase 1 major promoter: quantitation of tissue expression and identification of critical sequence elements. *Drug Metab. Dispos.* 35, 1649–1656

22 Andersen, M. *et al.* (2006) Alternative promoter usage of the membrane glycoprotein CD36. *BMC Mol. Biol.* 7, 8

23 Scohy, S. *et al.* (2000) Identification of an enhancer and an alternative promoter in the first intron of the alpha-fetoprotein gene. *Nucleic Acids Res.* 28, 3743–3751

24 Umenishi, F. and Verkman, A.S. (1998) Isolation and functional analysis of alternative promoters in the human aquaporin-4 water channel gene. *Genomics* 50, 373–377

25 Niimi, T. *et al.* (2006) Identification and expression of alternative splice variants of the mouse Ppp1r3b gene in lung epithelial cells. *Biochem. Biophys. Res. Commun.* 349, 588–596

26 Shibahara, S. *et al.* (2001) Microphthalmia-associated transcription factor (MITF): multiplicity in structure, function, and regulation. *J. Investig. Dermatol. Symp. Proc.* 6, 99–104

27 Shmelkov, S.V. *et al.* (2004) Alternative promoters regulate transcription of the gene that encodes stem cell surface protein AC133. *Blood* 103, 2055–2061

28 Wamstad, J.A. and Bardwell, V.J. (2007) Characterization of Bcor expression in mouse development. *Gene Expr. Patterns* 7, 550–557

29 Delaloy, C. *et al.* (2003) Multiple promoters in the WNK1 gene: one controls expression of a kidney-specific kinase-defective isoform. *Mol. Cell. Biol.* 23, 9208–9221
30 Barettino, D. *et al.* (1999) The mouse neurotrophin receptor trkB gene is transcribed from two different promoters. *Biochim. Biophys. Acta* 1446, 24–34
31 Pecci, A. *et al.* (2001) Promoter choice influences alternative splicing and determines the balance of isoforms expressed from the mouse bcl-X gene. *J. Biol. Chem.* 276, 21062–21069
32 Grinchuk, O. *et al.* (2005) The Optimedin gene is a downstream target of Pax6. *J. Biol. Chem.* 280, 35228–35237
33 Nakayama, A. *et al.* (2001) Characterization of two promoters that regulate alternative transcripts in the microtubule-associated protein (MAP) 1A gene. *Biochim. Biophys. Acta* 1518, 260–266
34 Zong, Y. *et al.* (2006) Expression of mouse Abcg2 mRNA during hematopoiesis is regulated by alternative use of multiple leader exons and promoters. *J. Biol. Chem.* 281, 29625–29632