

Transcriptional activation by bidirectional RNA polymerase II elongation over a silent promoter

Olivier Leupin^{1*}, Catia Attanasio^{1*}, Samuel Marguerat², Myriam Tapernoux², Stylianos E. Antonarakis¹
& Bernard Conrad^{1,2+}

¹Department of Genetic Medicine & Development, and ²Department of Genetics & Microbiology, University of Geneva Medical School, CMU, Geneva, Switzerland

Transcriptional interference denotes negative *cis* effects between promoters. Here, we show that promoters can also interact positively. Bidirectional RNA polymerase II (Pol II) elongation over the silent human endogenous retrovirus (HERV)-K18 promoter (representative of 2.5×10^3 similar promoters genome-wide) activates transcription. In tandem constructs, an upstream promoter activates HERV-K18 transcription. This is abolished by inversion of the upstream promoter, or by insertion of a poly(A) signal between the promoters; transcription is restored by poly(A) signal mutants. TATA-box mutants in the upstream promoter reduce HERV-K18 transcription. Experiments with the same promoters in a convergent orientation produce similar effects. A small promoter deletion partially restores HERV-K18 activity, consistent with activation resulting from repressor repulsion by the elongating Pol II. Transcriptional elongation over this class of intragenic promoters will generate co-regulated sense–antisense transcripts, or, alternatively initiating transcripts, thus expanding the diversity and complexity of the human transcriptome.

Keywords: transcriptional interference; human LTR transposons; RNA polymerase II; Pol II-regulated transcription; transcriptome diversity

EMBO reports (2005) 6, 956–960. doi:10.1038/sj.embor.7400502

INTRODUCTION

In eukaryotes, RNA polymerase II (Pol II) transcription produces messenger RNAs for more than 25,000 protein-coding genes, and is tightly controlled. Although control was thought to be exerted by proteins, it can be accomplished by RNA (Wassarman, 2004), and even by the elongating Pol II itself. Transcriptional interference—the negative effect of a promoter on an adjacent one *in cis* (Shearwin *et al*, 2005)—has long been recognized (Ward

& Murray, 1979; Adhya & Gottesman, 1982; Proudfoot, 1986; Elledge & Davis, 1989). In tandem promoters, the elongating Pol II emanating from the upstream promoter has been identified as the regulator; Pol II negatively interferes with the binding of an activator at the downstream promoter (Greger *et al*, 1998; Martens *et al*, 2004). Interference in convergent promoters involves colliding RNA polymerase complexes (Prescott & Proudfoot, 2002). A ‘sitting-duck’ model, or increased sensitivity to the collision of Pol II complexes escaping the weak promoter into productive elongation, was put forward (Callen *et al*, 2004). Spacing between promoters is considered important irrespective of promoter orientation. Indeed, the capacity of elongating Pol II complexes to read through DNA-bound protein barriers increases with the number of complexes pushing against the block (Epshtein *et al*, 2003).

A large fraction of genomes, up to one-half in humans, is occupied by transposon-derived interspersed repeats (Smit, 1999). To minimize the adverse effects of transposition and illicit recombination, the activity of these elements is restricted. This is achieved by a variety of means; for instance, inefficient Pol II elongation (Lorincz *et al*, 2004), methylation (Khodosevich *et al*, 2004) and formation of condensed chromatin. In one principal class of transposons, known as the human endogenous retrovirus long terminal repeats (HERV LTRs), the as yet unidentified protein complexes assemble at the promoter and repress Pol II transcription in most lineages except the germ line (Domansky *et al*, 2000). HERV-K18, the model system used here, belongs to this HERV-K (HML-2) subgroup (Barbulescu *et al*, 1999) and represents about 2.5×10^3 similar LTR promoters genomewide (Mager & Medstrand, 2003). HERV-K18 consists of two LTRs flanking intervening genes that encode defective structural and enzymatic activities, and cell-surface glycoproteins (supplementary Fig 1 online). The HERV-K18 element, located in the *CD48* gene on chromosome 1, has been characterized extensively for gene-regulatory features (Stauffer *et al*, 2001; Marguerat *et al*, 2004; Meylan *et al*, 2005).

RESULTS AND DISCUSSION

Pol II activates silent HERV-K18 in tandem promoters

Analysis of episomal reporter constructs comprising the whole HERV-K18 element or the nominal LTR promoter alone showed

¹Department of Genetic Medicine & Development, and ²Department of Genetics & Microbiology, University of Geneva Medical School, CMU, 1 rue Michel Servet, 1211 Geneva, Switzerland

*These authors contributed equally to this work

+Corresponding author. Tel: +41 22 379 57 19; Fax: +41 22 379 57 06;

E-mail: bernard.conrad@medecine.unige.ch

no significant amounts of constitutive reporter gene activity (supplementary Fig 1 online). This reflected the repressed transcriptional state of the endogenous *HERV-K18* element, which was at background levels, as detected by ribonuclease protection analysis (RPA) in all cell lineages except the thymus (Stauffer *et al*, 2001; Meylan *et al*, 2005). For these reasons, we used 5' LTR episomal reporter constructs in all further studies analysing *HERV-K18* promoter activity. When the composite simian virus 40 (SV40)-derived SR α promoter, consisting of one TATA-dependent and a second TATA-less promoter, was placed in tandem upstream of *HERV-K18*, it resulted in strong *HERV-K18* promoter activity (Fig 1A). This could be accounted for by the upstream recruitment at enhancers of factors overriding the repression of *HERV-K18*, or, alternatively, by Pol II itself initiating at the upstream promoter and activating the downstream regulatory region. To differentiate between these possibilities, we performed two experiments. First, we inverted the upstream SR α promoter such that orientation-independent enhancers were still available, but transcription over the *HERV-K18* promoter was completely eliminated. This extinguished *HERV-K18* reporter activity, which suggests that upstream SR α enhancers alone were not able to stimulate the LTR promoter. Second, we introduced a minimal, ~0.3 kb SV40 poly(A) signal (Orozco *et al*, 2002) sufficient for transcriptional termination between the two promoters. This similarly abolished *HERV-K18* reporter activity. Three single point mutations in the poly(A) signal sequence (AATAAAA to AGTACT) fully restored activated *HERV-K18* transcription, indicating that Pol II elongation over the *HERV-K18* promoter was responsible for the effect. The approximately 300 extra nucleotides of the poly(A) signal present in the mutant led to an even stronger signal. This implied a positive effect of spacing between the two promoters on transcription (Callen *et al*, 2004). Consistent with our observation, the likelihood of displacing a DNA-bound protein barrier was shown to increase with an increase in the number of Pol complexes pushing against it (Epshtein *et al*, 2003). The presence of several open reading frames and stop codons upstream of the transcription initiation site made it unlikely that luciferase activity could be derived from transcripts initiating upstream by cap-dependent translation. In conclusion, upstream Pol II transcription elongating over the 5' LTR was sufficient to activate the *HERV-K18* promoter activity.

To extend this idea further, we mutated the TATA box in one of the two known upstream SR α promoters (Fig 1B). Four single point mutations in the TATA box (TATTTA to TGAATT) were sufficient to reduce downstream *HERV-K18* promoter activity. The SV40-derived SR α promoter is known to consist of two overlapping promoters, one containing a TATA box that is active early in the SV40 life cycle, and a second that is turned on only during late replication (Buchman *et al*, 1984). Residual promoter activity of the second, TATA-independent upstream promoter is probably responsible for the residual activity that we measure after mutating the TATA box. Alternatively, TATA box mutations in human promoters reduce Pol II transcription only to some extent. Consistent with this second possibility, disruption of the TATA box in a human disease impairs transcription only partially (Ciotti *et al*, 1998). Therefore, this second set of experiments confirms and extends the idea that Pol II transcription initiating at the upstream SR α promoter and elongating over the LTR activates downstream *HERV-K18* promoter activity.

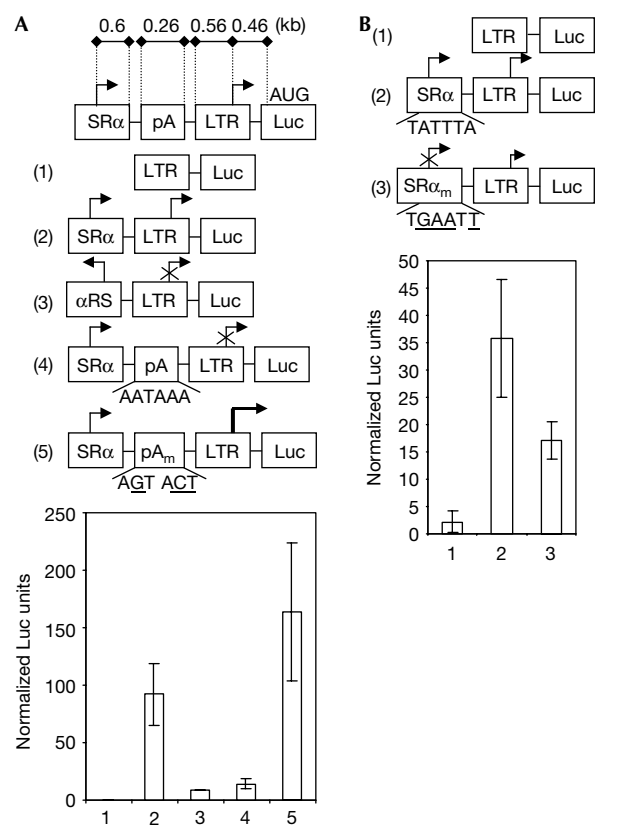


Fig 1 | RNA polymerase II transcription elongating over the tandem human endogenous retrovirus-K18 promoter activates transcription. (A) Physical distances are indicated on the top panel. The promoter constructs used in reporter gene assays with episomal vectors are shown (from top to bottom) as follows: (1) the human endogenous retrovirus (*HERV*)-K18 long terminal repeat (LTR); (2) tandem constructs with a composite simian virus 40 (SV40) promoter (SR α) upstream of the LTR promoter; (3) an LTR construct with the upstream SR α promoter inverted; (4) the SV40 poly(A) (pA) signal separating the two promoters; and (5) a pA signal carrying the mutation AATAAAA to AGTACT. Reporter gene activity in HeLa cells of constructs 1–5 is shown (average \pm 1 s.d. of three independent experiments). Rous sarcoma virus (RSV)-controlled *Renilla* activity was used as an internal control. (B) The following promoter constructs were analysed for reporter activity (from top to bottom): (1) the LTR; (2) tandem constructs with the composite SR α promoter upstream of the LTR; and (3) the TATA-dependent promoter of the two upstream SR α promoters carrying the mutation TATTTA to TGAATT (average \pm 1 s.d. of three independent experiments).

We wished to substantiate with transcription assays that upstream initiation would activate transcription at the *HERV-K18* promoter, and to relate this to activated *HERV-K18* transcription *in vivo*. Transcription was addressed by three technical approaches—RPA, primer extension (PE) and quantitative reverse transcriptase–polymerase chain reaction (qRT–PCR). First, RPA was performed on RNA extracted from cells that were transiently transfected with tandem promoter constructs (Fig 2A). The probe was positioned over the U3–R boundary of the LTR comprising the transcription initiation site such that upstream

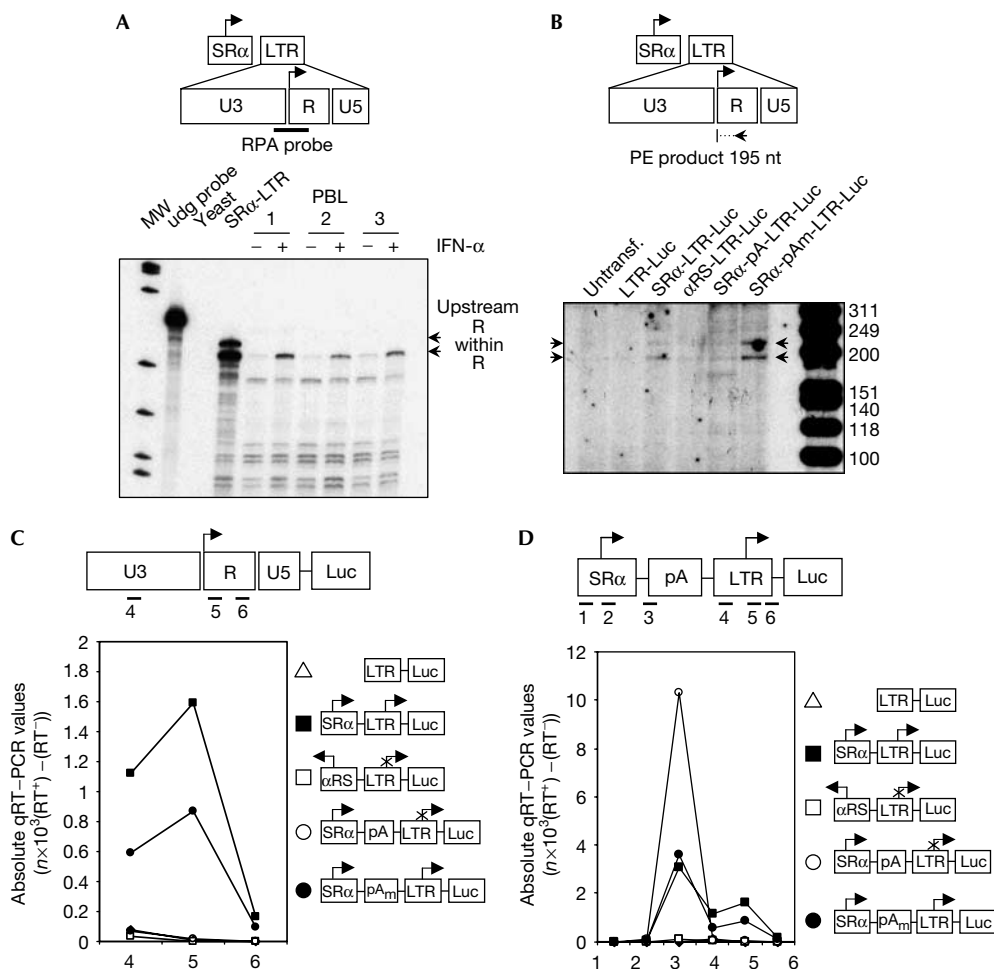


Fig 2 | Consistency between transcription and reporter assays shows that RNA polymerase II is the human endogenous retrovirus-K18 transcriptional activator. (A) The long terminal repeat (LTR) organization is shown, with U3, R and U5 regions. Transcription initiates precisely at the U3–R boundary. A ribonuclease protection assay (RPA) probe was positioned at this boundary. RPA with this probe was performed on RNA from transfected tandem SRα–LTR promoter constructs, and was compared with interferon-α (IFN-α)-activated LTR transcription *in vivo* in primary cells (PBL) of three individuals. Transcriptional initiation upstream of R is derived from SRα (top arrow) and transcription initiating in R is derived from LTR (bottom arrow). Undigested (udg) probes and yeast RNA are shown as controls. (B) RNAs extracted from reporter assays shown in Fig 1A were subjected to primer extension (PE), using a primer located 195 nt from transcriptional initiation of the LTR. Constructs are indicated on the top of the gel. (C,D) Quantitative reverse transcriptase–polymerase chain reaction (qRT–PCR) was performed with RNAs extracted from reporter assays shown in Fig 1A. The absolute values of triplicates for each point, obtained after subtraction of RT[–] from RT⁺ numbers are shown here. The approximate position of the amplification product is indicated on top of the panel, and constructs are shown on the right. (C) Values obtained over the LTR. (D) Values spread over the complete construct length.

transcription could be discriminated from HERV-K18-derived transcription (Fig 2). Both—transcription initiating upstream in the SRα promoter (upper arrow) and transcription emanating from the HERV-K18 LTR—were detected (lower arrow). The biological relevance of the lower band initiating within HERV-K18 is corroborated by findings showing that interferon-α (IFN-α)-activated HERV-K18 transcription *in vivo* in primary cells (PBL) of three independent individuals produces a protected band migrating at exactly the same size as the one resulting from tandem promoter fusions. This experiment validates and extends the results of reporter assays to naturally activated HERV-K18 transcription *in vivo*. Second, PE experiments were carried out on RNA extracted from the luciferase assays shown in Fig 1A. As can

be seen in Fig 2B, a PE product of 195 nt expected for transcripts initiating in the LTR precisely at the U3–R boundary was detected essentially only in conditions that gave rise to consistent reporter gene activity, namely, for the SRα-poly(A) (pA) mutant and the SRα-LTR constructs. These findings were supported by results from a second independent experiment including a further PE probe in the LTR placed 117 nt from initiation (not shown). Third, qRT–PCR corroborated the RPA and PE results (Fig 2C,D). Significant transcription over the LTR was identified exclusively for the two conditions that give rise to strong luciferase activity. As also shown by RPA, the signal intensity of transcription emanating from the LTR is stronger than that detected upstream of the LTR initiation site (compare (4) and (5) in Fig 2C). Conversely, constructs that do

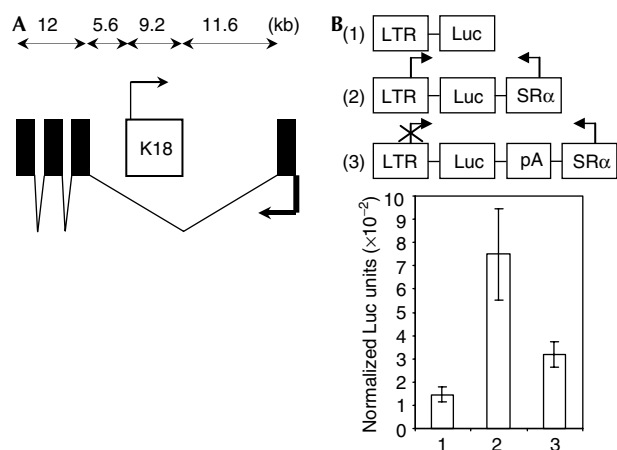


Fig 3 | Human endogenous retrovirus-K18 transcription is also activated in convergent promoter constructs. (A) The organization of the human endogenous retrovirus (*HERV*)-K18/*CD48* locus is depicted. *HERV*-K18 is positioned face to face with the *CD48* promoter at a distance of ~22 kb. (B) The following promoter constructs were analysed for reporter activity (from top to bottom): (1) the long terminal repeat (LTR); (2) constructs with the downstream *SRα* promoter face to face with the LTR at a distance of ~3 kb; (3) the simian virus 40 poly(A) (pA) signal placed to terminate *SRα*-driven convergent transcription (average ± 1 s.d. of three independent experiments).

not produce reporter activity lack detectable transcription over the LTR (Fig 2C,D). A second independent experiment yielding similar results is shown in supplementary Fig 2 online. In conclusion, the perfect consistency of reporter and transcription assays indeed shows that Pol II is the transcriptional activator of *HERV*-K18.

Convergent transcription activates *HERV*-K18

In its genomic setting, *HERV*-K18 is arranged in a convergent orientation with respect to the *CD48* promoter (Fig 3). We therefore examined the same promoter couple used in previous experiments in a convergent orientation. Surprisingly, even face to face, the *SRα* promoter was able to activate *HERV*-K18 transcription (Fig 3). A poly(A) signal between *SRα* and *HERV*-K18 abolished the effect, showing that Pol II elongating over the *HERV*-K18 LTR was required. These experiments show that Pol II, even when elongating transcription opposite to *HERV*-K18, is still able to activate the LTR promoter. The smaller amplitude of LTR activation as compared with the tandem arrangement probably reflects the susceptibility of transcription complexes forming at *HERV*-K18 to Pol II collision, once the putative repressor is dislodged by the first incoming complexes that have initiated at the SV40 promoter (Prescott & Proudfoot, 2002; Callen *et al*, 2004).

HERV-K18 is repressed by unidentified proteins

Previous experiments have shown that as yet unidentified protein complexes assembling at single *HERV*-K LTRs repress transcription (Domansky *et al*, 2000). Although *HERV*-K18 is nominally a member of the same subgroup of *HERV*-K LTRs, we confirmed that it behaved similarly in this respect. Deletion of 70 nt in the 3' end of the U5 moiety of the LTR restored promoter activity significantly, as described for other members of this class of *HERV*-K LTRs (Fig 4; Domansky *et al*, 2000). This effect was weaker, however, than transcriptional activation achieved by Pol II, which

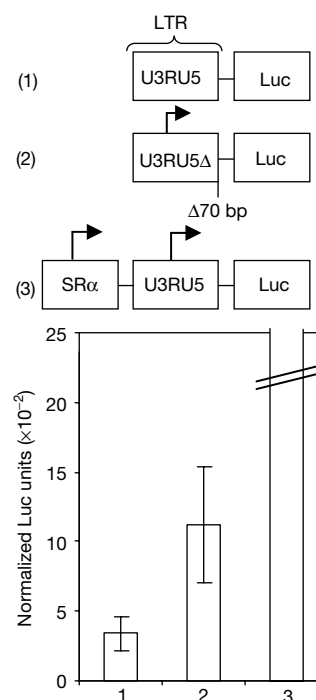


Fig 4 | A small deletion of the human endogenous retrovirus-K18 long terminal repeat (LTR) partially restores transcription. A 70 bp deletion ($\Delta 70$ bp) in the 3' boundary of the LTR (2) was compared with the wild-type LTR (1), and with tandem *SRα*-LTR promoter constructs (3) for reporter gene activity (average ± 1 s.d. of three independent experiments).

could result from concomitant removal of activator binding sites, and/or by repressor binding sites elsewhere on the promoter. Collectively, we conclude that the *HERV*-K18 promoter is actively repressed, and that Pol II complexes elongating bidirectionally over the promoter are responsible for relieving repression. We propose repressor repulsion as the mechanism accounting for transcriptional activation by bidirectional Pol II elongation over the *HERV*-K18 promoter.

HERV-K18 is not silenced post-transcriptionally

Some repressors function by recruitment of cofactors that induce repressive chromatin modifications. Chromatin immunoprecipitation experiments were performed in primary cells in which *HERV*-K18 and *CD48* are differentially expressed (supplementary Fig 3 online). Marks of active chromatin were found over actively transcribing (thymus) *CD48* and *HERV*-K18 regions, but not silent ones (epithelial cells). Post-transcriptional silencing by heterochromatin formation, as reported for other transposons (Lorincz *et al*, 2004; Novina & Sharp, 2004), does not seem to have an important role with respect to *HERV*-K18, because repressive chromatin modifications, such as di- and trimethylated lysine 9 and heterochromatin protein 1, were localized to the silent *CD48* promoter, rather than to *HERV*-K18. In conclusion, repression at the LTR by as yet unidentified negative regulatory factors seems to be the predominant mechanism in place to silence *HERV*-K18 promoter activity *in vivo*.

The presence of about 2.5×10^3 similar LTRs genome-wide may have significant functional consequences because the elongating

Pol II activates transcription at these promoters. Several testable predictions can be made for LTRs in convergent or tandem orientation to cellular promoters. For instance, when Pol II activates transcription from an LTR face to face, double-stranded RNA will be produced (Novina & Sharp, 2004). Such Pol II-regulated antisense transcription may have a regulatory role (Lavorgna et al, 2004). At least 20,000 sense-antisense transcriptional units have been described in humans, with a sizable fraction of noncoding RNAs (Kampa et al, 2004).

For LTRs in tandem with cellular promoters, alternatively initiating transcripts will be generated after Pol II crosses the LTRs. This provides an alternative mechanism to splicing for increasing the coding diversity of individual transcriptional units (Levine & Tjian, 2003). The proof of principle for this has already been provided (van de Lagemaat et al, 2003). Transposable elements, and particularly LTRs, have been identified to a significant extent among the 5' termini of cellular transcripts involved in rapidly evolving functions. Given the considerable differences among species and individuals in nature and the composition of LTR transposons, our findings extend the number of mechanisms that are responsible for transcriptome diversity and complexity in humans.

METHODS

Constructs. Oligonucleotide sequences are indicated in the supplementary information online. The episomal reporter constructs were generated by PCR on the basis of pREP4-Luc and pREP7-Rluc (Liu et al, 2001). The SR α promoter was from pCDL (Takebe et al, 1988) and the SV40 early poly(A) was from pREP4-Luc.

Reporter assays. See the supplementary information online.

Ribonuclease protection analysis. 3T3 mouse fibroblasts were transiently transfected with Fugene (T. Hoffmann-La Roche Ltd, Basel, Switzerland), using the SR α -LTR tandem promoter construct in a pBSK backbone. RPA was performed on total RNA with a probe spanning the U3-R region of the LTR. The probe was generated by PCR (see the supplementary information online). The 259 nt probe was generated with T3 polymerase on *Spe*I-linearized plasmid templates.

Primer extension and quantitative reverse transcriptase-polymerase chain reaction. See the supplementary information online.

Supplementary information is available at *EMBO reports* online (<http://www.emboreports.org>).

ACKNOWLEDGEMENTS

We thank Dr D. Belin, Dr S. Bontron, Dr R. Lyle, Dr W. Reith and Dr M. Strubin for advice and discussion, Dr J. Curran for reading the manuscript, Dr D. Garcin for 4 \times ISRE and Dr K. Zhao for pREP4-Luc and pREP7-Rluc. This work was supported by grants from the Swiss National Science Foundation.

REFERENCES

Adhya S, Gottesman M (1982) Promoter occlusion: transcription through a promoter may inhibit its activity. *Cell* **29**: 939–944

Barbulescu M, Turner G, Seaman MI, Deinard AS, Kidd KK, Lenz J (1999) Many human endogenous retrovirus K (HERV-K) proviruses are unique to humans. *Curr Biol* **9**: 861–868

Buchman AR, Fromm M, Berg P (1984) Complex regulation of simian virus 40 early-region transcription from different overlapping promoters. *Mol Cell Biol* **4**: 1900–1914

Callen BP, Shearwin KE, Egan JB (2004) Transcriptional interference between convergent promoters caused by elongation over the promoter. *Mol Cell* **14**: 647–656

Ciotti M, Chen F, Rubaltelli FF, Owens IS (1998) Coding defect and a TATA box mutation at the bilirubin UDP-glucuronosyltransferase gene cause Crigler-Najjar type I disease. *Biochim Biophys Acta* **1407**: 40–50

Domansky AN, Kopantzev EP, Snezhkov EV, Lebedev YB, Leib-Mosch C, Sverdlov ED (2000) Solitary HERV-K LTRs possess bi-directional promoter activity and contain a negative regulatory element in the U5 region. *FEBS Lett* **472**: 191–195

Elledge SJ, Davis RW (1989) Position and density effects on repression by stationary and mobile DNA-binding proteins. *Genes Dev* **3**: 185–197

Epshtein V, Toulme F, Rahmouni AR, Borukhov S, Nudler E (2003) Transcription through the roadblocks: the role of RNA polymerase cooperation. *EMBO J* **22**: 4719–4727

Greger IH, Demarchi F, Giacca M, Proudfoot NJ (1998) Transcriptional interference perturbs the binding of Sp1 to the HIV-1 promoter. *Nucleic Acids Res* **26**: 1294–1301

Kampa D et al (2004) Novel RNAs identified from an in-depth analysis of the transcriptome of human chromosomes 21 and 22. *Genome Res* **14**: 331–342

Khodosevich K, Lebedev Y, Sverdlov ED (2004) Large-scale determination of the methylation status of retrotransposons in different tissues using a methylation tags approach. *Nucleic Acids Res* **32**: e31

Lavorgna G, Dahary D, Lehner B, Sorek R, Sanderson CM, Casari G (2004) In search of antisense. *Trends Biochem Sci* **29**: 88–94

Levine M, Tjian R (2003) Transcription regulation and animal diversity. *Nature* **424**: 147–151

Liu R, Liu H, Chen X, Kirby M, Brown PO, Zhao K (2001) Regulation of CSF1 promoter by the SWI/SNF-like BAF complex. *Cell* **106**: 309–318

Lorincz MC, Dickerson DR, Schmitt M, Groudine M (2004) Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. *Nat Struct Mol Biol* **11**: 1068–1075

Mager DL, Medstrand P (2003) Retroviral repeat sequences. In *Nature Encyclopedia of the Human Genome* (ed. D. Cooper), Vol 5, pp 57–63. London, UK: Nature Publishing Group, Macmillan Publishers Ltd

Marguerat S, Wang WY, Todd JA, Conrad B (2004) Association of human endogenous retrovirus K-18 polymorphisms with type 1 diabetes. *Diabetes* **53**: 852–854

Martens JA, Laprade L, Winston F (2004) Intergenic transcription is required to repress the *Saccharomyces cerevisiae* SER3 gene. *Nature* **429**: 571–574

Meylan F, De SM, Leclercq G, Plum J, Leupin O, Marguerat S, Conrad B (2005) Negative thymocyte selection to HERV-K18 superantigens in humans. *Blood* **105**: 4377–4382

Novina CD, Sharp PA (2004) The RNAi revolution. *Nature* **430**: 161–164

Orozco IJ, Kim SJ, Martinson HG (2002) The poly(A) signal, without the assistance of any downstream element, directs RNA polymerase II to pause *in vivo* and then to release stochastically from the template. *J Biol Chem* **277**: 42899–42911

Prescott EM, Proudfoot NJ (2002) Transcriptional collision between convergent genes in budding yeast. *Proc Natl Acad Sci USA* **99**: 8796–8801

Proudfoot NJ (1986) Transcriptional interference and termination between duplicated α -globin gene constructs suggests a novel mechanism for gene regulation. *Nature* **322**: 562–565

Shearwin KE, Callen BP, Egan JB (2005) Transcriptional interference—a crash course. *Trends Genet* **21**: 339–345

Smit AF (1999) Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr Opin Genet Dev* **9**: 657–663

Stauffer Y, Marguerat S, Meylan F, Ucla C, Sutkowski N, Huber B, Pelet T, Conrad B (2001) Interferon- α -induced endogenous superantigen: a model linking environment and autoimmunity. *Immunity* **15**: 591–601

Takebe Y, Seiki M, Fujisawa J, Hoy P, Yokota K, Arai K, Yoshida M, Arai N (1988) SR α promoter: an efficient and versatile mammalian cDNA expression system composed of the simian virus 40 early promoter and the R-U5 segment of human T-cell leukemia virus type 1 long terminal repeat. *Mol Cell Biol* **8**: 466–472

van de Lagemaat LN, Landry JR, Mager DL, Medstrand P (2003) Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends Genet* **19**: 530–536

Ward DF, Murray NE (1979) Convergent transcription in bacteriophage λ : interference with gene expression. *J Mol Biol* **133**: 249–266

Wassarman KM (2004) RNA regulators of transcription. *Nat Struct Mol Biol* **11**: 803–804