Transcriptional regulation of three genes, C4, DAP3 and PKIB by antisense-oriented intronic endogenous retroviruses

F. Bröcker[1,2], J. Heinrich[1], R. Horton[2], M.R. Hoehe[2], H. Lehrach[2], K. Moelling[1,2]

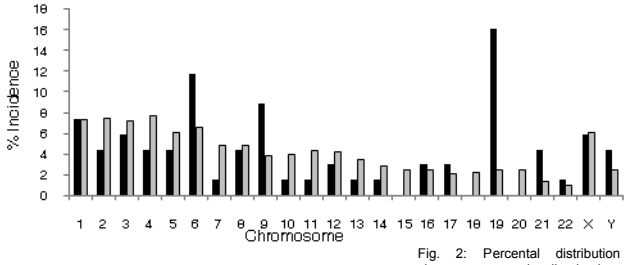
[1]University of Zuerich, Research Group Viruses and Cancer, Switzerland; [2]Max-Planck-Institut für molekulare Genetik, Department Lehrach, Berlin; email: moelling@imm.uzh.ch

INTRODUCTION

Human endogenous retroviruses (HERVs) are viruses that once infected ancestral germ line cells and became fixed in the genome, now making up approximately 8% of the human genome. A HERV located in intron 9 of the C4 gene on 6p21.33, designated HERV-K(C4), was shown to possess promoter activity in its 3' LTR in opposite direction to the encompassing gene, thus generating an antisense transcript with a suggested role in negatively regulating C4 gene expression. It was also demonstrated that LTR promoter activity is inhibited by interferon(IFN)-γ-stimulation in a dose-dependent way [1]. There are two C4 gene variants, a long form containing the HERV-K(C4) and a short form lacking the insertion. The C4 gene is part of the hypervariable RCCX region within the MHC and shows an extensive copy number variation [2]. The repeated block in the RCCX has a length of about 26.5kb or 32.8kb in the presence of the short or long form of C4 respectively. Variation between copies of the HERV-K(C4) is low [3]. Performing BLAT searches against the 2009 human genome reference assembly we identified two HERV-K(C4)-related endogenous retroviruses located in antisense orientation within intron 1 of the DAP3 gene on 1q22 and within intron 2 of the PKIB gene on 6q22.31, respectively. DAP3 is an IFN-γ inducible positive regulator of apoptosis [4] and is over-expressed in numerous cancers. PKIB is an inhibitor of protein kinase C and G (PKC/PKG) and has recently been shown to be upregulated in prostate cancer [5].

We are conducting luciferase-reporter assays to test for promoter activity of HERV LTRs and assess the effect of cytokine stimulation. Furthermore, PCR-based techniques are used to detect putative antisense transcripts. Their effect on the expression level of the respective genes will be assessed by over-expression and strand-specific siRNA approaches. These data will allow us to elucidate the role of the presented HERVs in the regulation of their respective genes.

Distribution of HERV-K(C4) elements



$$\%\,\text{Incidence} = \frac{n(\text{elements on individual chromosome})}{n(\text{elements on all chromosomes})} \times 100\%$$

Fig. 2: Percental distribution on chromosomes visualized by the incidence of HERV elements on individual chromosomes represented by bars. Black: HERV-K(C4)-related elements identified by performing BLAT searches against the UCSC 2009 human genome reference assembly GRCh37 (http://genome.ucsc.edu), n = 68. Grey: all HERV elements according to Kim et al. [6], n = 38054. Note that, when compared with the the distribution of all HERVs, HERV-K(C4)-related elements are clearly overrepresented on chromosomes 6, 9, 19, 21 and Y, while being underrepresented on chromosomes 2-5, 7, 10-15, 18 and 20. This bias indicates that in the evolutionary past HERV-K(C4) proviruses were either preferentially formed on particular chromosomes, or preferentially deleted on others.

Models of antisense regulation

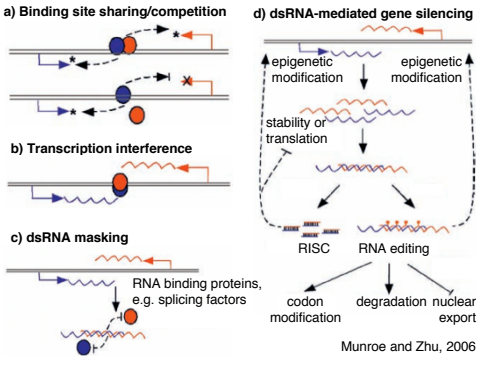


Fig. 4: Schematic representation proposed for antisense-regulation by overlapping sense-antisense transcripts. (a) Juxtaposed promoters share or compete with transcription factors. (b) Concurrent transcription on opposite DNA strands induces topological contraints. (c) Sense- and antisense-transcripts form dsRNA, so that RNA binding proteins cannot bind. (d) dsRNA is a putative substrate for both the RNA induced silencing complex (RISC) and the RNA editing machinery, that are able to induce post-transcriptional inhibition as well as epigenetic modifications.

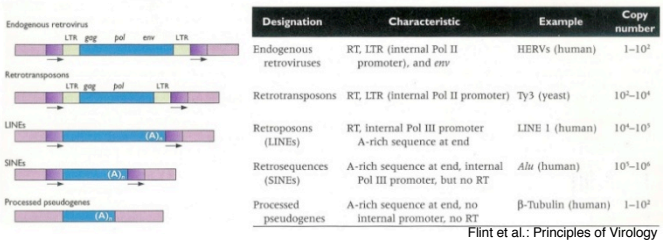
Promoter activites of HERV-K(C4) LTRs

Platzhalter für Balkendiagramm mit Promoteraktivitäten der HERV-K(C4) LTRs (Luciferase-Aktivitäten), Ergebnisse am Montag

6:

References: **[1] Mack et al.; *Immunogenetics* 56, 321-332 (2004)**
[2] Yu; *Exp. Clin. Immunogen.* 15, 213-230 (1998)
[3] Horton et al.; *Immunogenetics* 60, 1-18 (2008)
[4] Kissil et al.; *J. Biol. Chem.* 270, 27932-27936 (1995)
[5] Chung et al.; *Oncogene* 28, 2849-2859 (2009)
[6] Kim et al.; *Mol. Cells* 18, 87-93 (2004)

Retroviral elements and retrotransposons in the human genome

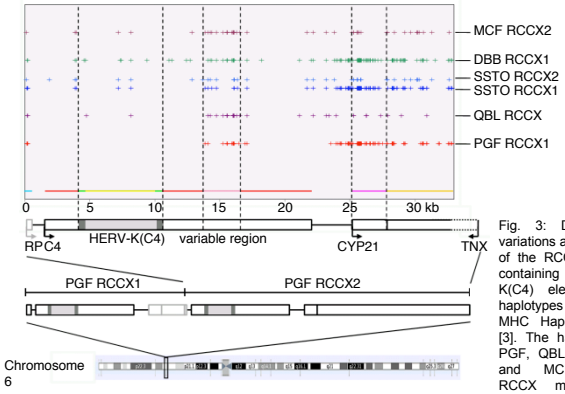


Designation	Characteristic	Example	Copy number
Endogenous retroviruses	RT, LTR (internal Pol II promoter), and *env*	HERVs (human)	1–10²
Retrotransposons	RT, LTR (internal Pol II promoter)	Ty3 (yeast)	10²–10⁴
Retroposons (LINEs)	RT, internal Pol III promoter A-rich sequence at end	LINE 1 (human)	10⁴–10⁵
Retrosequences (SINEs)	A-rich sequence at end, internal Pol III promoter, but no RT	*Alu* (human)	10⁵–10⁶
Processed pseudogenes	A-rich sequence at end, no internal promoter, no RT	β-Tubulin (human)	1–10²

Flint et al.: Principles of Virology

Fig. 1: Different types of retroviral elements and retrotransposons in eukaryotic genomes. They make up approximately 50% of the human genome, human endogenous retroviruses (HERVs) about 8%. HERVs are retroviruses that once infected ancestral germ line cells and became fixed in the genome. Their LTRs (long terminal repeats) contain strong promoters for the expression the retroviral genes gag, pol and env.

Distribution of variations in the MHC haplotypes RCCX modules



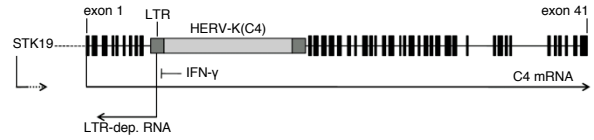
RCCX modules

RP (STK19): Ser/Thr nuclear protein kinase
C4: Complement component 4
CYP21: Steroid-21 hydroxylase
TNX: Tenascin X (ECM protein)

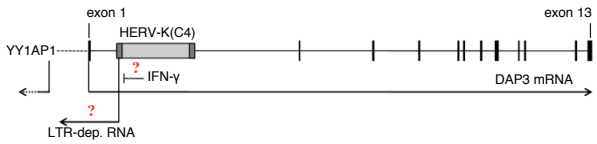
Fig. 3: Distribution of variations amongst copies of the RCCX duplication containing the HERV-K(C4) element in the haplotypes studied in the MHC Haplotype Project [3]. The haplotypes are: PGF, QBL, SSTO, DBB and MCF. Indicated RCCX modules were aligned against the second copy in the PGF haplotype (PGF RCCX2). Single base mutations are indicated by crosses. Although there is considerable variation in the regions differentiating the C4 genes (variable region) and the CYP21 genes and pseudogenes, the HERV-K(C4) appears relatively well conserved. Pseudogenes are indicated by grey outlines.

Chromosomal context of C4, DAP3 and PKIB genes

C4 on chromosome 6p21.33 (20623 bp)



DAP3 on chromosome 1q22 (49443 bp)



PKIB on chromosome 6q22.31 (254455 bp)

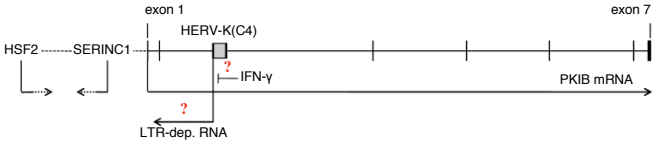


Fig. 5: Schematic overview and chromosomal locations of the analyzed genes DAP3, C4 and PKIB and their respective intronic HERV-K(C4) insertions. Intron/exon structure (exons: black bars) and HERV elements (dark grey, LTRs; light grey, open reading frame) are drawn to scale according to the 2009 human genome reference assembly GRCh37 (http://genome.ucsc.edu). HERV-K(C4) locations and orientations as annotated by the RepeatMasker database (http://www.repeatmasker.org). Upstream genes that might be regulated by LTR-dependent transcripts are indicated. Abbrev.: DAP3, death-associated protein 3; PKIB, protein kinase inhibitor beta; YY1AP1, YY1 (a zinc finger transcription factor) associated protein 1; HSF2, heat shock transcription factor 2; SERINC1, serine incorporator 1.

CONCLUSION

…