OPEN ACCESS Freely available online

ISSN: 2153-0602
Journal of
Data Mining in
Genomics & Proteomics

Perspective

Endogenous Retrovirus Gene Transcription Inhibition in Human Cancers

Swieton Liane*

Department of Genetics, Medical University of Bialystok, Bialystok, Poland

ABOUT THE STUDY

Up to 8% of the human genome is made up of Endogenous Retroviruses (ERVs), which are the remains of long-gone retroviral infections. Many proviruses from the Human Endogenous Retrovirus Type K (HERV-K) Human Endogenous Retrovirus K (HML-2) family, the only lineage still proliferating in the genome after the human-chimpanzee split, have intact open reading frames, even though these elements are mainly fragmented and dormant. Some of these proviruses encode accessory genes called Np9 and rec that interact with oncogenic pathways. Numerous studies have demonstrated that the temporary production of ERVs causes abnormal self-renewal and unchecked proliferation in both stem cells and malignancies.

The Cancer Genome Atlas (TCGA) is a useful tool for cancer research due to the abundance of high-quality genomic and transcriptome Illumina sequence data accessible from a variety of distinct tumour types. RNA-seq data can be used to infer the expression of certain repeating sequences, like genes encoded by HERV-K, although there is currently no standardized computational approach for doing so (HML-2). This paper introduces a novel and highly specialized pipeline that can detect and measure transcription of the provirus-encoded proteins Np9 and rec, which have a striking degree of sequence similarity and are transcribed at extremely low levels. Using our unique methods, we demonstrate that, in contrast to healthy tissues, np9 and rec are overexpressed in breast cancer, germ cell tumours, cutaneous melanoma, lymphoma, ovarian cancer, and prostate cancer. We further demonstrate that np9 and rec are uniquely expressed in human pre-implantation embryos at the 8 and 16-cell stages. Using our unique methods, we demonstrate that, in contrast to healthy tissues, Np9 and rec are overexpressed in breast cancer, germ cell tumours, cutaneous melanoma, lymphoma, ovarian cancer, and prostate cancer. We further demonstrate that np9 and rec are uniquely expressed in human pre-implantation embryos at the 8 and 16-cell stages.

Endogenous retroviruses (ERVs) are ancient retroviral infections that account for 5%-8% of the human genome. Because of mutations or deletions, these ancient elements are mostly inactive or epigenetically repressed. Many proviruses in the HERV-K (HML-2) (shortened to HK2 here) family, the only lineage still proliferating in the genome after the human-chimp split, have intact open reading frames, some of which encode for accessory proteins called Np9 and Rec, which have been found to interact with cancer-related cellular pathways. A 292 base pair deletion in type I proviruses results in a different splice donor, producing an Np9 transcript instead of rec, which is encoded by type II proviruses. Rec is a functional homolog of the accessory proteins Rev and Rex, which are encoded by Human Immunodeficiency Virus type 1 (HIV-1) and Human T-Lymphotropic Virus type 1 (HTLV-1). By transporting viral mRNA from the nucleus into the cytosol, Rev/Rex/Rec regulates viral gene expression. Rec protein has been shown to disrupt germ cell development and cause carcinoma in mice. It is unknown what role Np9 plays in virus replication, if any, but it does appear to be oncogenic.

Rec is a functional homolog of the accessory proteins Rev and Rex, which are encoded by HIV-1 and HTLV-1. By transporting viral mRNA from the nucleus into the cytosol, Rev/Rex/Rec regulates viral gene expression. Rec protein has been shown to disrupt germ cell development and cause carcinoma in mice. It is unknown what role Np9 plays in virus replication, if any, but it does appear to be oncogenic. Np9 protein can activate the Extracellular Signal-Regulated Kinase (ERK), Ak Strain Transforming (AKT) and Neurogenic Locus Notch Homolog Protein 1 (Notch1) pathways, and up regulate β-catenin, promoting survival and growth of leukemia stem/progenitor cells. Np9 protein is also preferentially expressed in transformed cells and is known to interact with LNX, an important part of the Notch signal transduction pathway, implicated in regulating breast and prostate cancer proliferation. Its expression significantly promotes the growth of leukemia cells, and when silenced, the growth of myeloid and lymphoblastic leukemia cells is inhibited. Both Np9 and Rec proteins interact with Promyelocytic Leukemia Zinc Finger (PLZF) tumour suppressor, a chromatin remodeler and transcriptional repressor that is implicated in spermatogonial stem cell renewal and cancer. The Np9 protein has also been shown to bind and regulate the

Correspondence to: Swieton Liane, Department of Genetics, Medical University of Bialystok, Bialystok, Poland, E-mail: lianeswi@gmail.com

Received: 04-Jul-2022, Manuscript No. JDMGP-22-17793; **Editor assigned:** 06-Jul-2022, PreQC No. JDMGP -22-17793 (PQ); **Reviewed:** 20-Jul-2022, QC No JDMGP-22-17793; **Revised:** 27-Jul-2022, Manuscript No. JDMGP-22-17793 (R); **Published:** 03-Aug-2022. DOI: 10.4172/2153-0602.22.13. 260.

Citation: Liane S (2022) Endogenous Retrovirus Gene Transcription Inhibition in Human Cancers. J Data Mining Genomics Proteomics. 13: 260.

ubiquitin ligase Murine Double Minute 2 (MDM2), which inhibits the activity of the p53 tumour suppressor. Transcripts of Np9 and Rec are also found in human embryonic stem cells and human induced pluripotent stem cell lines, where they are linked to pluripotency maintenance. With recent advances in high-throughput mRNA sequencing (RNA-seq), it is now possible to find novel genes and transcripts as well as measure transcription in a single assay. Small RNA-seq experiments can generate massive amounts of data: current sequencers can generate over 500 gigabases of raw sequencing reads per run.

Because of the sensitivity of RNA-seq, it is possible to detect alternative splice isoforms of transcripts, as well as rare and cell and context-specific transcripts. Furthermore, because the number of reads produced by an RNA transcript is proportional to its abundance, the depth of read coverage, or read density, can be used as a proxy for gene expression. However, there could be inherent biases in the RNA-seq methodology that must be considered. In general, most sequencing platforms generate short sequence reads that are computationally assembled to produce full-length transcripts. Even for well-characterized transcripts, such as those from protein-coding genes, ambiguity arises due to the expression of transcriptional variants that frequently share exons.