OVINE β-ENDOGENOUS RETROVIRUSES : FROM INTEGRATION TO COPY-SPECIFIC TRANSCRIPTION

Marie Verneret, Caroline Leroux, Emmanuelle Lerat, Vincent Navratil, Jocelyn Turpin

Ovine β-Endogenous retroviruses : from integration to copy-specific transcription

Verneret Marie[1,2], Leroux Caroline[1], Lerat Emmanuelle[2], Navratil Vincent[3], Turpin Jocelyn[1]

1- UMR754 INRAE-University Lyon 1-Ecole Pratique des Hautes Etudes, "Viral Infections and Comparative Pathology", University of Lyon, France
2- UMR5558 University Lyon1-CNRS-VetAgroSup, "Biometrics and Evolutionary Biology", University of Lyon, France
3- Bioinformatics Rhône Alpes Pole – Biological System Modeling and Analysis Platform, University of Lyon, France

Introduction

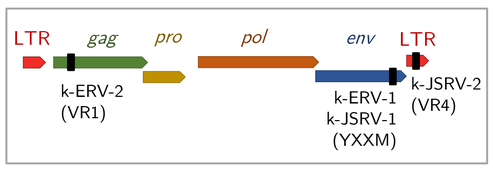
☑ JSRV (Jaagsiekte Sheep RetroVirus) is a transmissible exogenous retrovirus that induces lung cancers in sheep.
☑ Its endogenous counterparts (ERV), resulting from germinal infections, are widely present in sheep genomes and **coexist** with JSRV in infected animals.
☑ The study of JSRV and ERV transcription is a way to better understand their role in animals and to identify potential interactions between the two viral forms. Targeted tools need to be developped to **discriminate JSRV and ERV signals** due to their highly similar genomic sequences (>95% identity).
☑ We made a proof of concept analyzing the **specific transcription of JSRV-related ERV** using publicly available genome assembly and RNA-seq data from the Rambouillet Benz2616 ewe produced by the Ovine FAANG project.

What is the JSRV-related ERVs copies transcription landscape in sheep tissues ?

Methods and results

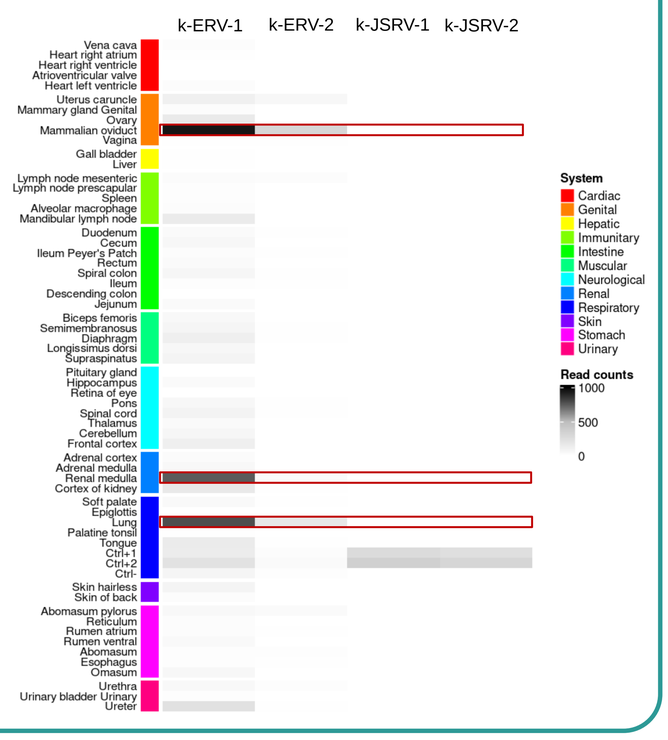
❶ ERVs are transcribed in specific organs

1. **Design of custom 21-mers in variable regions to discriminate endogenous and exogenous transcription**



2. ERV/JSRV signal in RNA-seq reads from 58 Benz2616 tissues

Reads including the different k-mers were counted with Bbduk tool (1 mismatch allowed) and their position on viral genome were checked using BWA-mem mapping approach. Three supplemental lung samples from JSRV-infected (Ctrl+) and non-infected (Ctrl-) animals were also analyzed (Karagianni J. Virol, 2019).

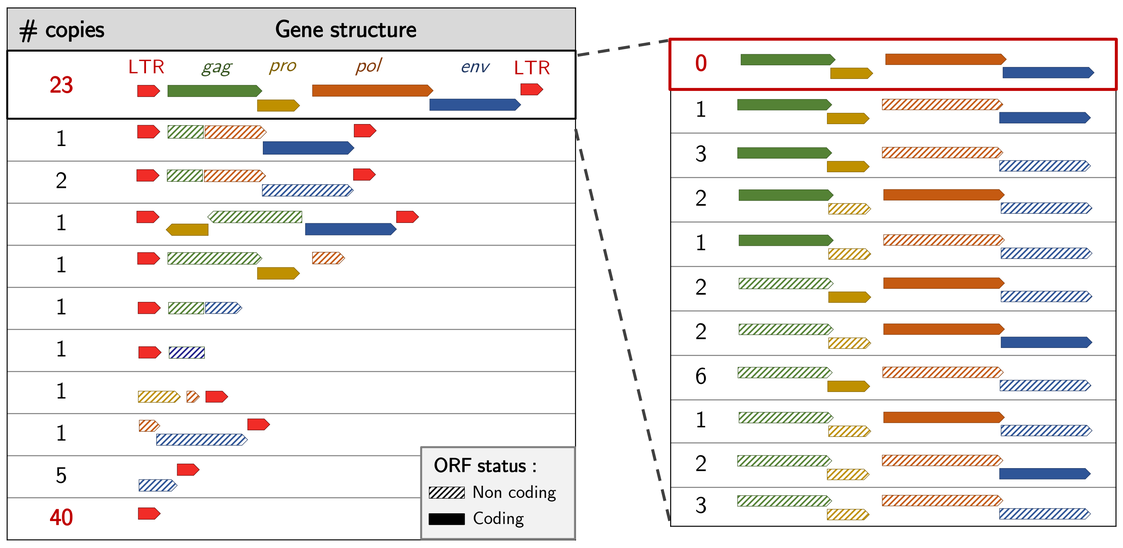


☑ JSRV-related ERVs are **higly transcribed in lung, kidney and oviduct**
☑ No JSRV signal was detected in Benz2616 samples

❷ Benz2616 genome assembly contains multi-structured ERVs

ERVs annotation in Benz2616 genome using a similarity based approach

ERV copies have been annotated in the sheep reference assembly (ARS-UI-Ramb2.0) with BLASTN and RepeatMasker using a panel of 28 complete JSRV-related ERVs sequences as query (Palmarini J. Virol, 2000 & Arnaud PLOS Pathogens, 2007). Gene and ORF (Open Reading Frame) annotations have been performed with MUSCLE and ORFfinder algorithms.

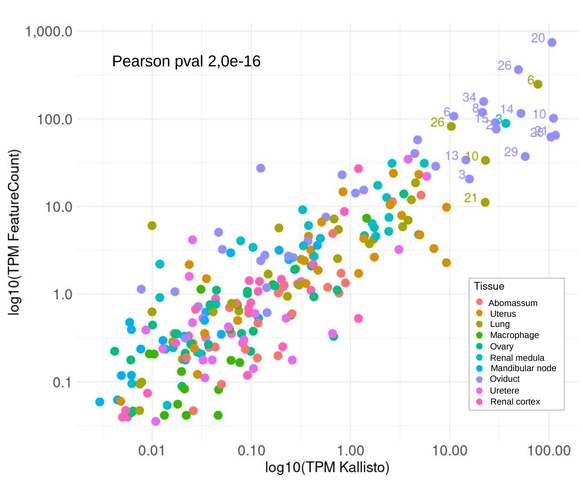


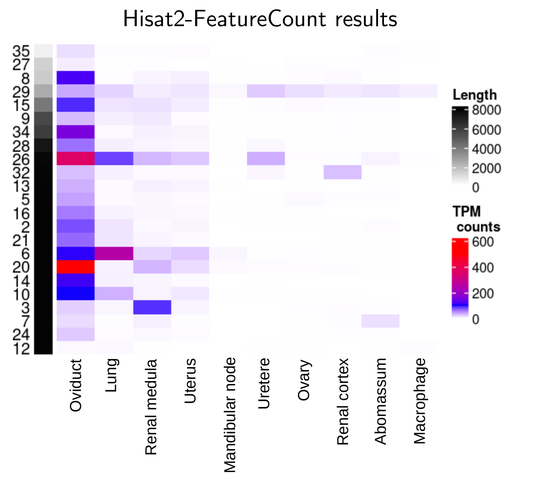
☑ 77 ERV copies were identified including 37 insertions with retroviral genes and 40 solo LTRs
☑ Among the 23 full insertions, **all have at least one non coding ORF**

❸ Transcription of specific ERV copies in tissues

Comparaison of alignment-based and pseudo-alignment methods to quantify copy-specific transcription

Two tools were tested on 10 selected tissues from Benz2616 : a mapping based approach using Hisat2 and FeatureCount and a pseudo-aligmnent one with Kallisto. The ARS-UI-Ramb2.0 genome has been taken as the reference for mapping. Only uniquely mapped reads have been conserved. The read counts are shown in TPM (Transcripts Per Kilobase Million).



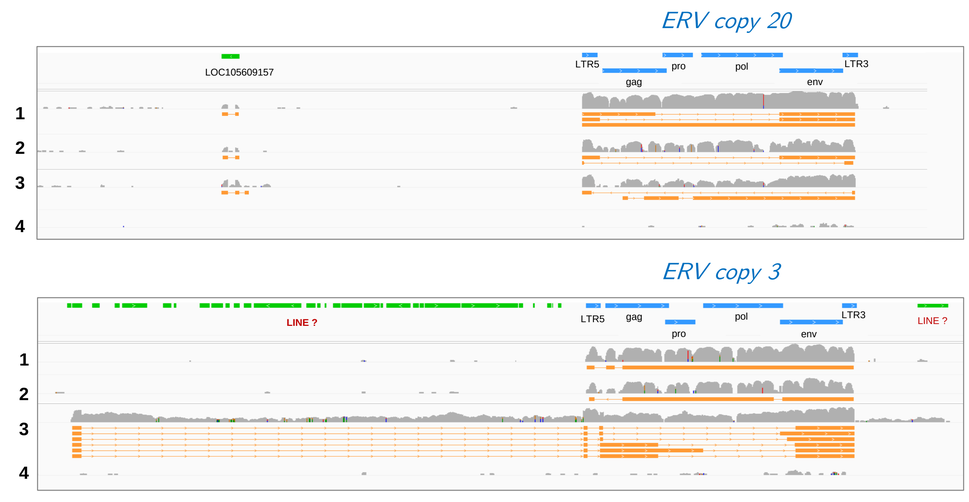


☑ Alignment based and pseudo-alignment based results are **positively correlated**
☑ The global transcription signal in oviduct, lung and renal medula is driven by the **transcription of 19 ERV copies** : 15 oviduct-specific, 1 kidney-specific and 4 transcribed in oviduct and lung

❹ ERV and flanking genome transcription interactions

Host genes and ERV RNA-seq reads distribution and transcripts

Hisat2 uniquely mapped reads associated to ERV (blue) and the flanking host genes (green) have been visualized on IGV. The read depth is shown in gray for oviduct (1), lung (2), renal medula (3) and macrophage (4) in log scale. Stringtie was used for transcript reconstruction (orange).



☑ **Both full length and *env* mRNA** are observed for highly transcribed ERV copies (copy 20)
☑ Chimeric transcription between specific ERV copies and the flanking genome can lead to alternative splicing (copy 3)

Conclusions and Perspectives

☑ JSRV-related ERVs are **globally transcribed in specific tissues** including the reproductive tract and the renal system. Transcription also has been identified in lungs in non-infected as well as in JSRV-infected sheep.
☑ The combination of alignment-based and pseudo-alignment based methods allowed the identification of **ERV copy-specific transcription** and confirmed that the presence of non-coding ORF does not prevent ERVs from being transcribed.
☑ **ERV isoforms** depend on the transcription level of the copies in tissues and their insertion sites in the host genome.
☑ To better understand the role of the ERVs, their impact on the genome and the correlation of their transcription activity with that of the host genes will be assessed.

Aknowledgments
The project is financed by ANR, INRAE and the BioEnvis federation of research. VM is supported by an ANR PhD fellowship.