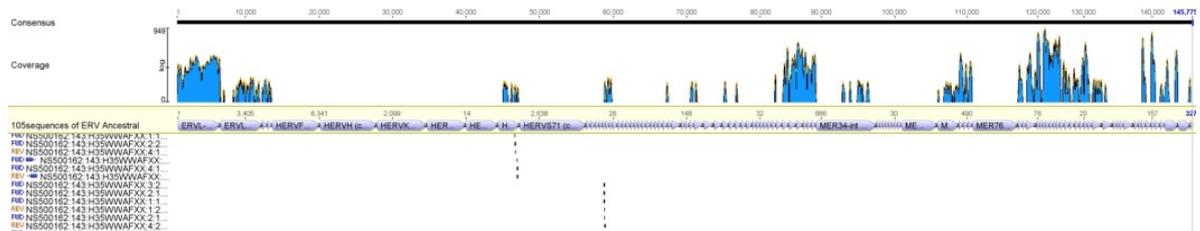


## Supplementary Materials

**Table S1.** DNA samples and sample locations used for high-throughput sequencing (HTS) data. F; female and M; male.

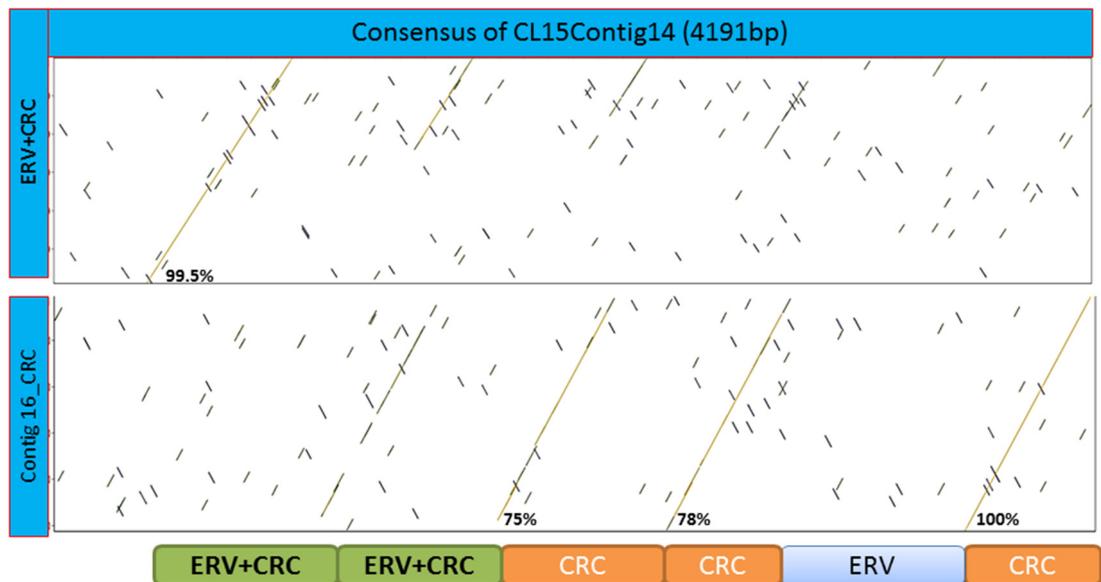
Sheep breed (Sample code)	Sample location	Sex
Hamdani (HamM)	Erbil-Iraq	M
Hamdani (HamJ2)	Duhok-Iraq	M
Hamdani (HamJ1)	Erbil-Iraq	M
Karadi (KarM)	Erbil-Iraq	M
Karadi (KarJ)	Duhok-Iraq	F



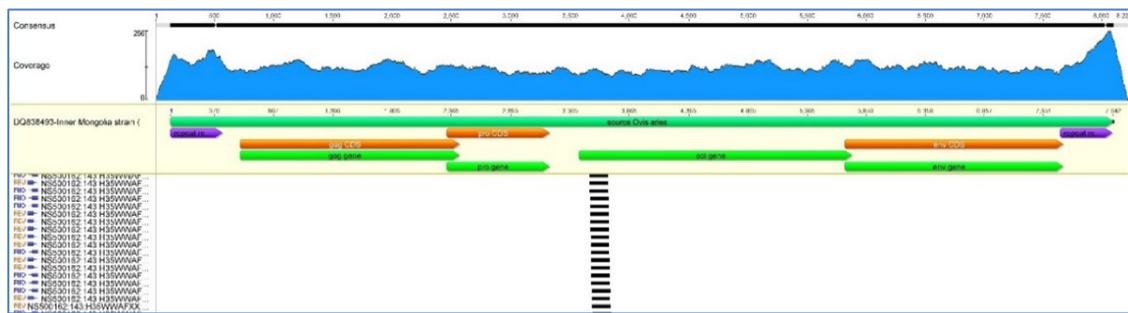
**Figure S1.** Mapping of high-throughput sequencing (HTS) data of sheep to ancestral sequences of ERVs



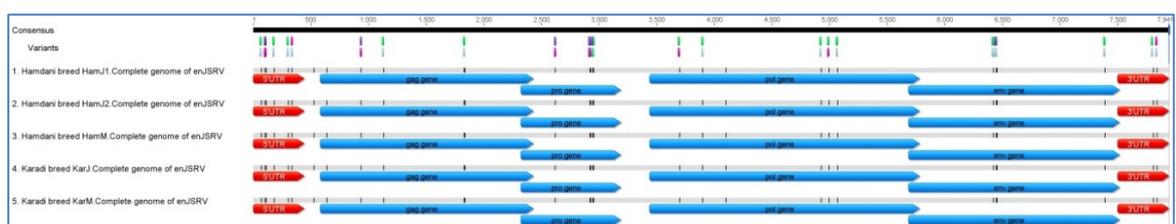
**Figure S2.** Mapping of high-throughput sequencing (HTS) data of sheep to *Bos taurus* sequences of ERVs



**Figure S3.** Consensus of CL15C14 (4191bp) of RepeatExplorer including combined ERV1+CRC sequences, three copies of 32merC16\_Sat\_CRC satellite like sequences and ERVs. Accordingly, the 32merC16\_Sat\_CRC was named 32mer\_ERV1+CRC Figure 5. Furthermore, combined ERV1+CRC sequences and three copies of 32merC16\_Sat\_CRC satellite like sequences were found in the same consensus of CL15C14 (4191bp).



**Figure S4.** Assembly of high-throughput sequencing (HTS) data (black lines) to reference complete genome of enJSRV DQ838493 showing the 5' and 3' LTRs (purple) and the four gene open reading frames, ORFs (brown for coding sequence and genes for green). Reads cover the whole sequence with an average depth of c. 120x (blue, see Table 6.2) and increased depth in the LTRs. Illumina sequencing gave good shotgun coverage (equal forward FWD and reverse REV) reads, and matched left and right paired-end reads to the sequence (shown by symbol after REV/FWD and before read code NS500162...).



**Figure S5.** Distribution of 50 SNPs along the complete genome of enJSRV of five samples of the Iraqi sheep breeds Hamdani and Karadi.