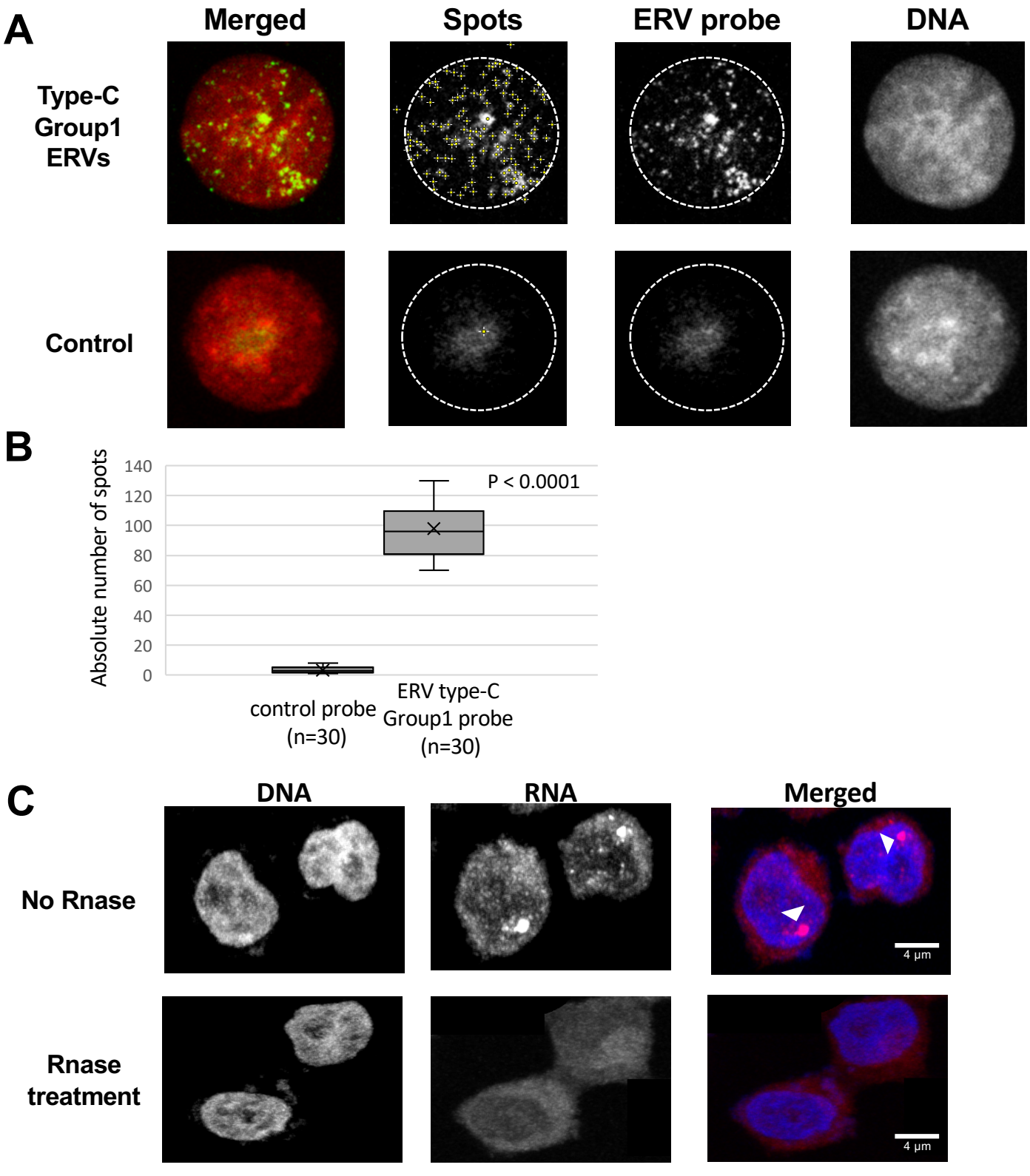
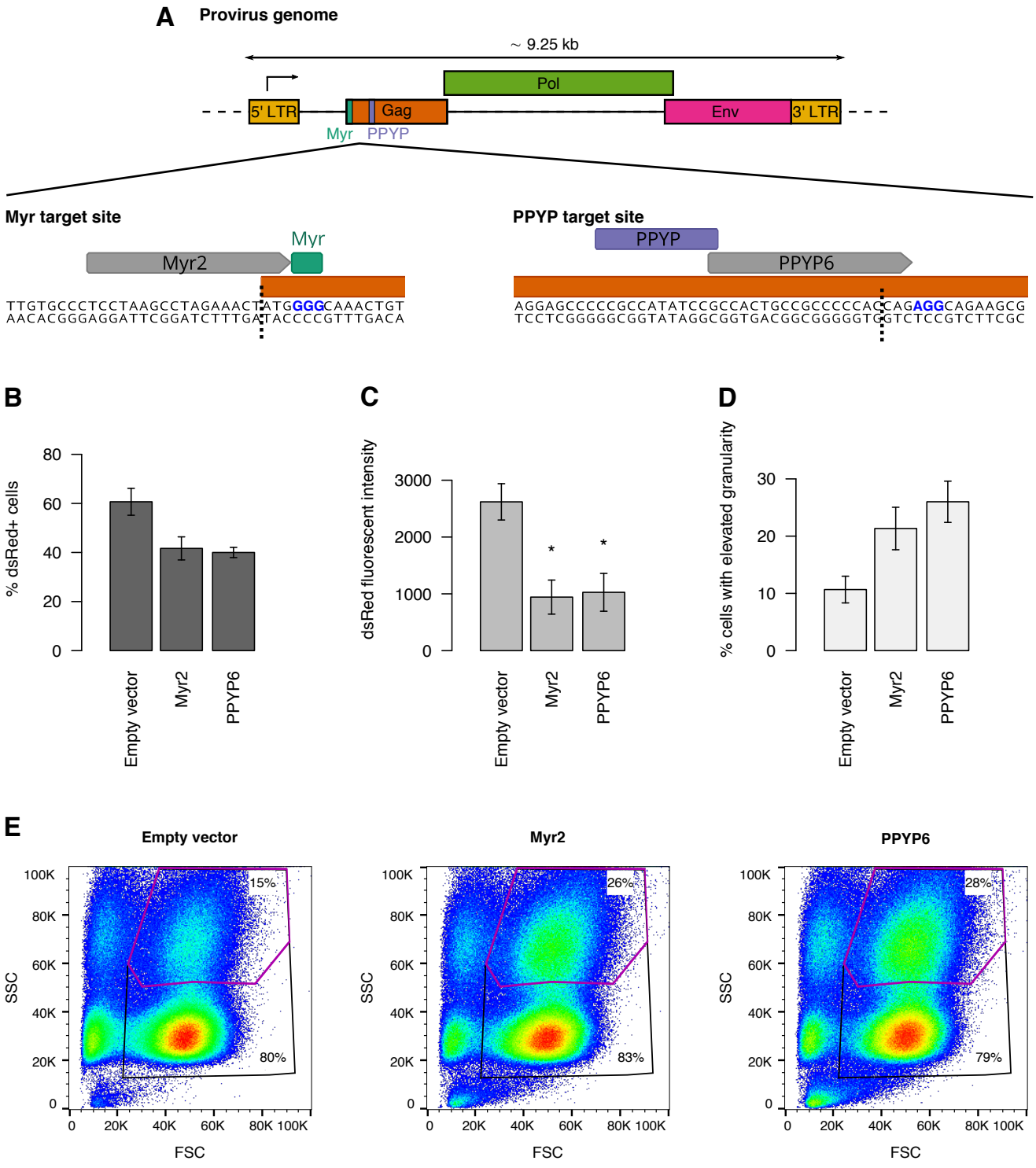
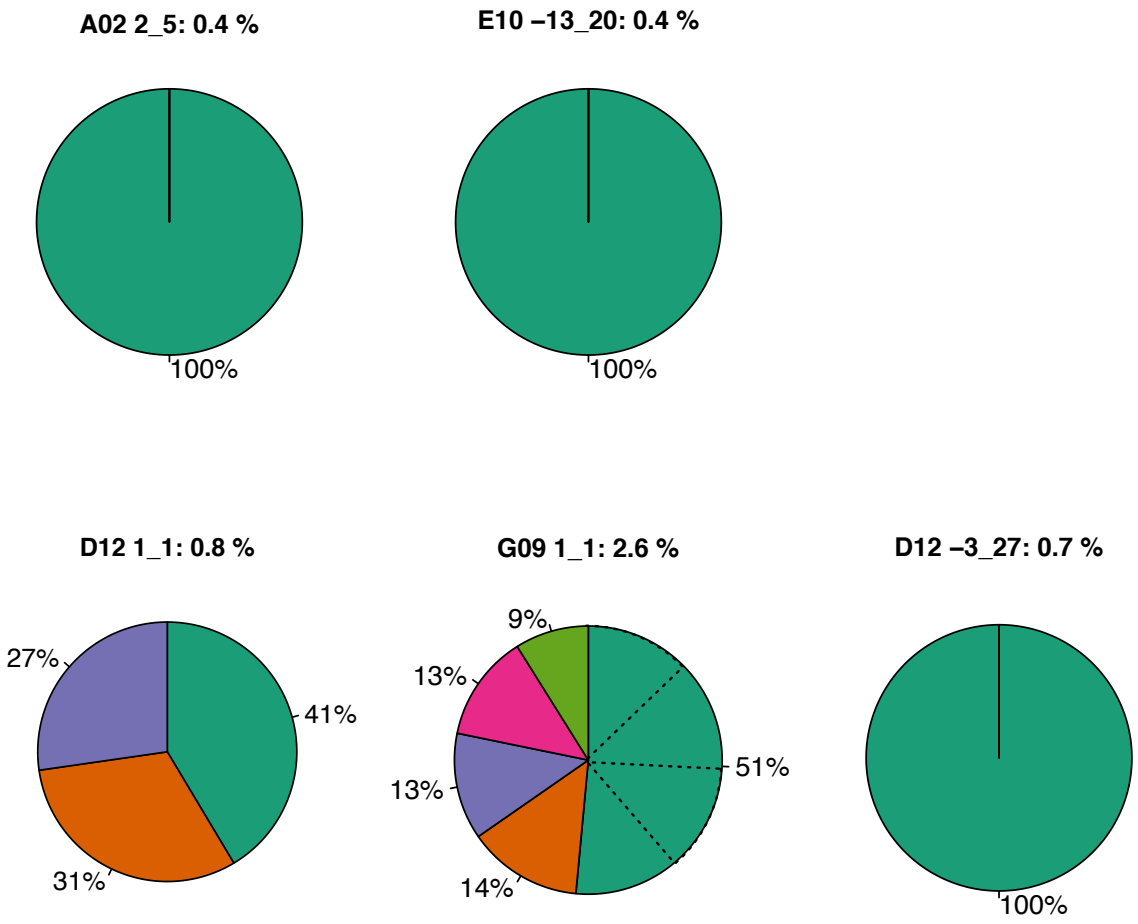



**Supplementary figure S1. Type-C group1 ERVs DNA-FISH or RNA-FISH** (A) CHO cell Type-C group1 ERVs imaging. Confocal pictures of interphase cells subjected to DNA-FISH using a probe targeting the group1 ERVs sequence (upper pictures) or a negative control probe (lower pictures) were taken with the indicated filter settings and pseudocolored for visualization purposes (total DNA in red, Type-C group1 ERV DNA-FISH signal in green, nuclear mask as dashed circles). Fluorescent images were treated using ImageJ function «Find Maxima». (B) Interphase cell type-C group 1 ERV DNA-FISH signal distributions versus a control probe are illustrated as box-plots diagrams. The box-plot edges represent 25–50% and 50–75% quartiles. The median value are represented by the middle bar. Box-plot whiskers (0–25% and 75–100% quartiles) indicate the maximum and minimum values in the distributions. The n values indicate the number of cells analyzed. Only cells with similar nucleus area were qualified for the analysis. Mean values are indicated by crosses. Statistical analysis (p values) were calculated using Mann-Whitney analysis. (C) Detection of type-C group1 ERV RNA by in situ RNA FISH using a probe directed against the Type-C group1 ERV subjected (bottom pictures) or not (top pictures) to a RNAse treatment (Blue: DNA staining, Red: ERV type C RNA).

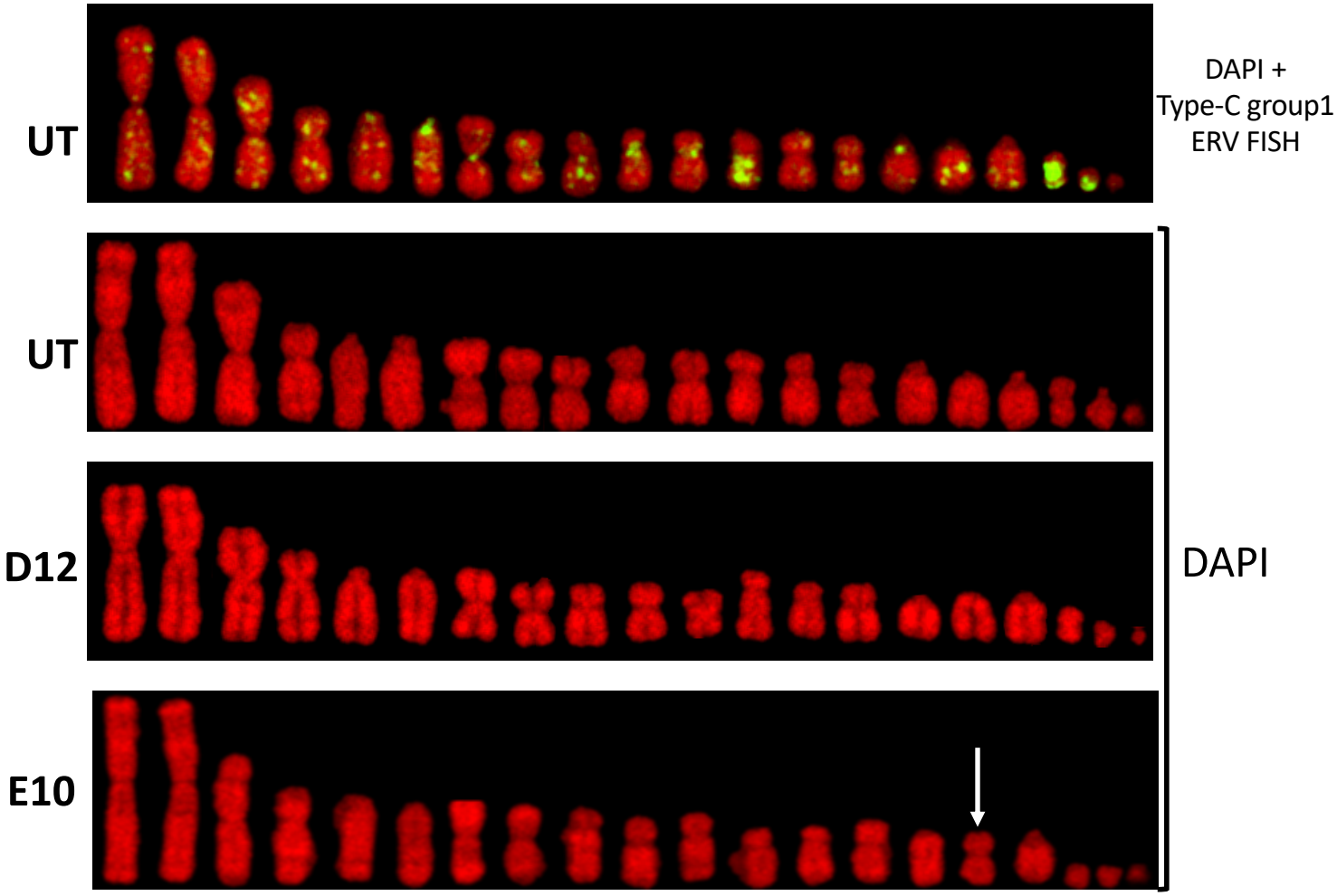



**Supplementary Figure S2. CRISPR-Cas9 target sites and assessment of cleavage efficiency by flow cytometry.** (A) The orientation and position of the Myr2 and PPYP6 sgRNA sequences designed to target the Myristoylation (Myr) and PPYP motifs of the *gag* group 1 type-C ERV consensus sequence, respectively, are illustrated by grey arrows. The CRISPR-Cas9 cleavage sites are shown by dotted lines, while the gag coding sequence, Myr and PPYP motifs are depicted by orange, green and blue boxes, respectively. The protospacer adjacent motif (PAM) sites are marked by blue bold letters. Analysis of the dsRed positive (dsRed+) cell frequency (B), the dsRed fluorescence intensity (C) and the frequency of high granularity cells (D and E) of CHO cells transfected with the CRISPR-Cas9 and Myr- or PPYP-specific sgRNAs or non-targeting empty control vector, and with the dsRed transfection control expression plasmids. Panel E shows size (FSC) vs granularity (SSC) flow cytometry density plots of the empty vector-, Myr2 or PPYP6 sgRNA-treated cells. The larger black gate selects for intact non-debris cells while the smaller purple gate marks the CHO cell subpopulation with an elevated granularity level, as quantified in panel D. Statistical significance relative to the empty vector control was calculated using the two-tailed unpaired Student's t-test with Benjamini and Hochberg false discovery rate correction (n = 3, error bars represent s.e.m, * $P < 0.05$).

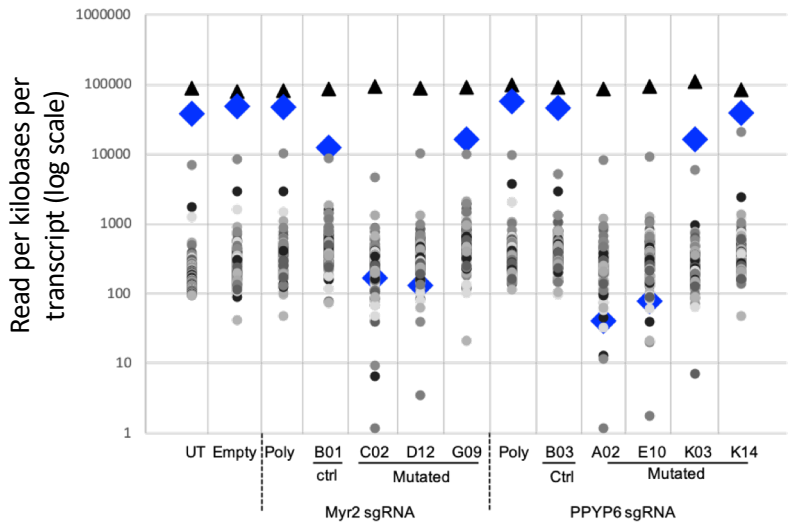

**Supplementary Figure S3. Clusters of Myr and PPYP motifs flanking sequence diversity relative a CHO cell Type-C ERV consensus sequence.** Colors correspond to the phylogenetic groups depicted in Figure 3B and 3C. Myr (A) and PPYP (B) clusters containing a sgRNA recognition site (black outlined arrow) with an adjacent PAM sequence are listed by bold letters. **The** Myr and PPYP motifs are indicated with turquoise and purple outlined boxes, respectively. The higher sequence complexity of the PPYP flanking region relative to the Myr flanking region is illustrated by missing sequences and colored lines depicting deletions, insertions and single nucleotide variants, respectively.



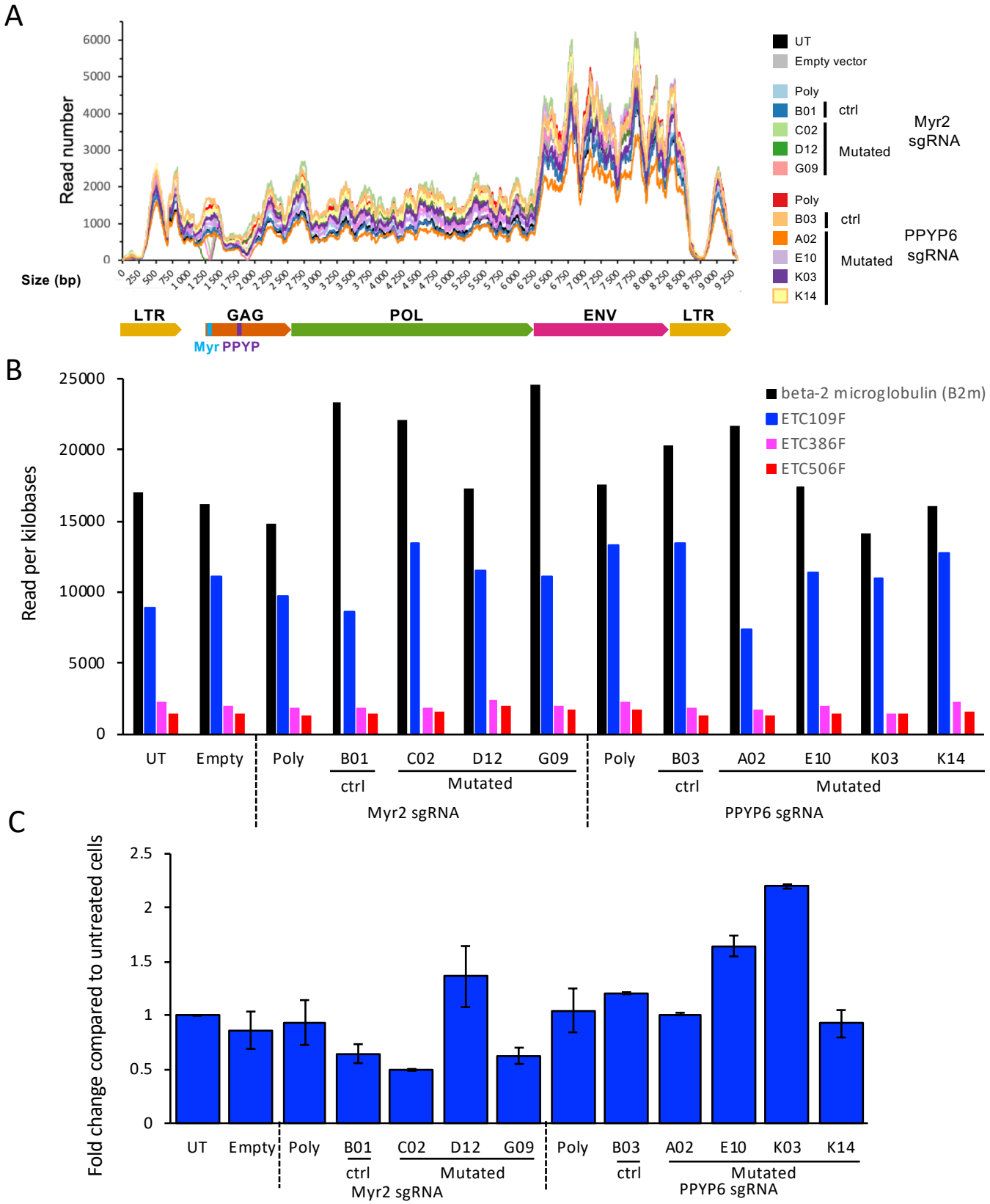

**Supplementary Figure S4. Characterization of ERV locus-specific mutations and their frequencies within clonal populations.** Analysis of Illumina raw reads of mutations detected at normal (0.2-0.4%) or high (> 0.4%) read frequencies in different clones. Pie charts represent the number and frequency of identified groups with identical CRISPR-derived mutation but distinct mutation-flanking sequences (e.g. in D12_1_1 and G09_1_1). Dotted lines indicate the number of predicted ERV loci that could not be distinguished based on their flanking sequences.



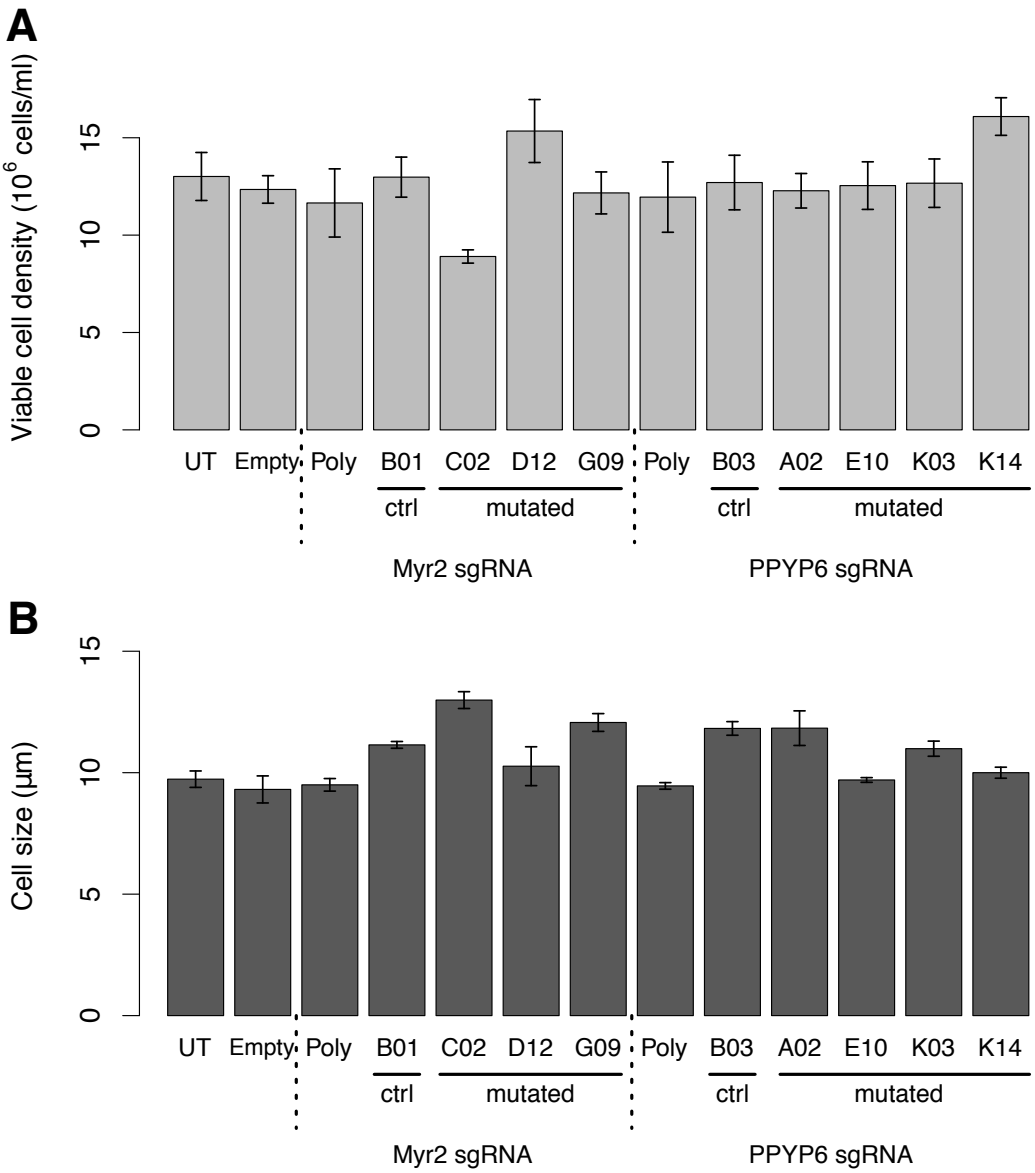

**Supplementary figure S5. CHO-K1 control or CRISPR treated cells karyotype**
Representative metaphase spread of untreated (UT) CHO-K1 chromosome FISH analysis using fluorescent probes specifically targeting group 1 type-C ERV (top panel). The chromosomal DAPI-stained DNA is represented in red and the FISH signals of integrated retroviral sequences, as obtained by type-C group1 ERV probes staining, are shown as green dots. The DAPI-stained chromosomal analysis of the untreated CHO cell and edited clones D12 and E10 are as indicated (lower three panels). The arrow indicates a chromosome subjected to a rearrangement.




**Supplementary Figure S6. Analysis of RNAseq sequencing data from the ERV-edited clones.** This graph represent the RNAseq analysis of Illumina sequencing reads obtained from VP RNA isolated from five-day cultures of untreated CHO cells (UT), empty sgRNA vector-treated cells (Empty), bulk-sorted polyclonal CRISPR-treated cells (Poly), as well as clones containing mutations in the expressed type-C group 1 ERV locus (C02, D12, G09, A02, E10, K03 and K14) or without a detected ERV mutation (B01, B03). The y-axis represents the number of reads per kilobase of transcript for each gene. The two most highly represented RNAs obtained from viral particle preparations, the 45S ribosomal RNA and the ETC109F type-C ERV, are represented by black triangles and blue diamonds, respectively, as expected from the co-purification of the viral particles with the ribosomes released into the culture supernatants during preparation. The CHO cell 100 genes having the highest number of mRNA reads mapping to their DNA sequences are depicted in this analysis as black and grey dots, corresponding to trace contaminating amounts of these cellular mRNAs, as the VP RNA preparation method did not include a RNAse step.




**Supplementary Figure S7. Assessment of the Type-C ERVs mRNA levels in ERV-mutated CHO cell clones.**
(A) Illumina sequencing reads of total cellular RNA obtained from ERV-edited clones and controls were mapped on the group 1 type-C ERV ETC109F sequence. The ERV sequence size are represented by the x-axis in bases pairs and the numbers of reads that begin at each ERV sequence nucleotide by the y-axis. (B) Cellular mRNA Illumina sequencing. Total RNA was isolated from five-day cultures of untreated CHO cells (UT), empty sgRNA vector-treated cells (Empty), bulk-sorted polyclonal CRISPR-treated cells (Poly), as well as clones containing mutations in the expressed type-C group 1 ERV locus (C02, D12, G09, A02, E10, K03 and K14) or without a detected ERV mutation (B01, B03). Samples were subjected a Ribo-Zero rRNA treatment, and poly-adenylated mRNAs were reverse transcribed and processed for Illumina sequencing. The obtained reads were mapped on the three expressed Type-C ERV sequences from group 1 (blue), 2 (purple) and 3 (red) and on the beta-2 microglobulin reference gene. The y axis represents the read numbers per kilobase. (C) Cellular mRNA was isolated from five-day cultures of the indicated CHO cells, as for panel A. Reverse transcription and q-PCR were performed on samples corresponding to 3 independent CHO-K1 cultures for each cell population. The type-C ETC109 ERV mRNA levels were quantified by RT-qPCR and normalized to those of the GAPDH gene. The fold changes were calculated relative to the untreated (UT) cells.




**Supplementary Figure S8. Assessment of cell growth and size of ERV-mutated CHO cell clones.** Viable cell density (A) and cell size (B) were measured for untreated CHO cells (UT), empty sgRNA vector-treated cells (Empty), bulk-sorted polyclonal CRISPR-treated cells (Poly), as well as for cells of clones containing mutations in the expressed group 1 ERV locus (C02, D12, G09, A02, E10, K03 and K14) or without a mutation (B01, B03) after five days of culture. No statistically significant difference relative to the empty vector control was determined using the two-tailed unpaired Student's t-test with Benjamini and Hochberg false discovery rate correction (n = 3, error bars represent s.e.m, $P > 0.05$).