



<u>Flockerzi et al. Additional Figure 1:</u> Pustell matrix comparisons of HERV-K(HML-2) proviruses identified as transcriptionally active in this study. Proviral sequences were compared to the previously published HERV-K(HML-2.HOM) sequence. Location of LTR and proviral *gag*, *pro*, *pol* and *env* genes within the latter sequence is indicated. The following parameters were applied for comparisons: Window size = 30 nt; Min. % Score = 55; Hash Value = 6; Jump = 1; Strands = both. Note that provirus $c7_A$ —the actual HERV-K(HML-2.HOM) provirus—is represented in tandem structure in the human genome reference sequence. HERV-K(HML-2) type 1 loci display a gap at nt 6500 with respect to the HERV-K(HML-2.HOM) sequence, as opposed to type 2 loci. Provirus sequences were reverse complemented prior to comparison if located in antisense on the respective human chromosome. The following parameters were applied for comparisons: Window size = 30 nt; Min. % Score = 55; Hash Value = 65; Jump = 1; Strands = both. Thus, each position in one sequence that shares on the DNA level at least 55% similarity in a 30 nt window to the other sequence is marked by a dot, and producing a line for longer stretches of DNA.