Supplemental Figure S2

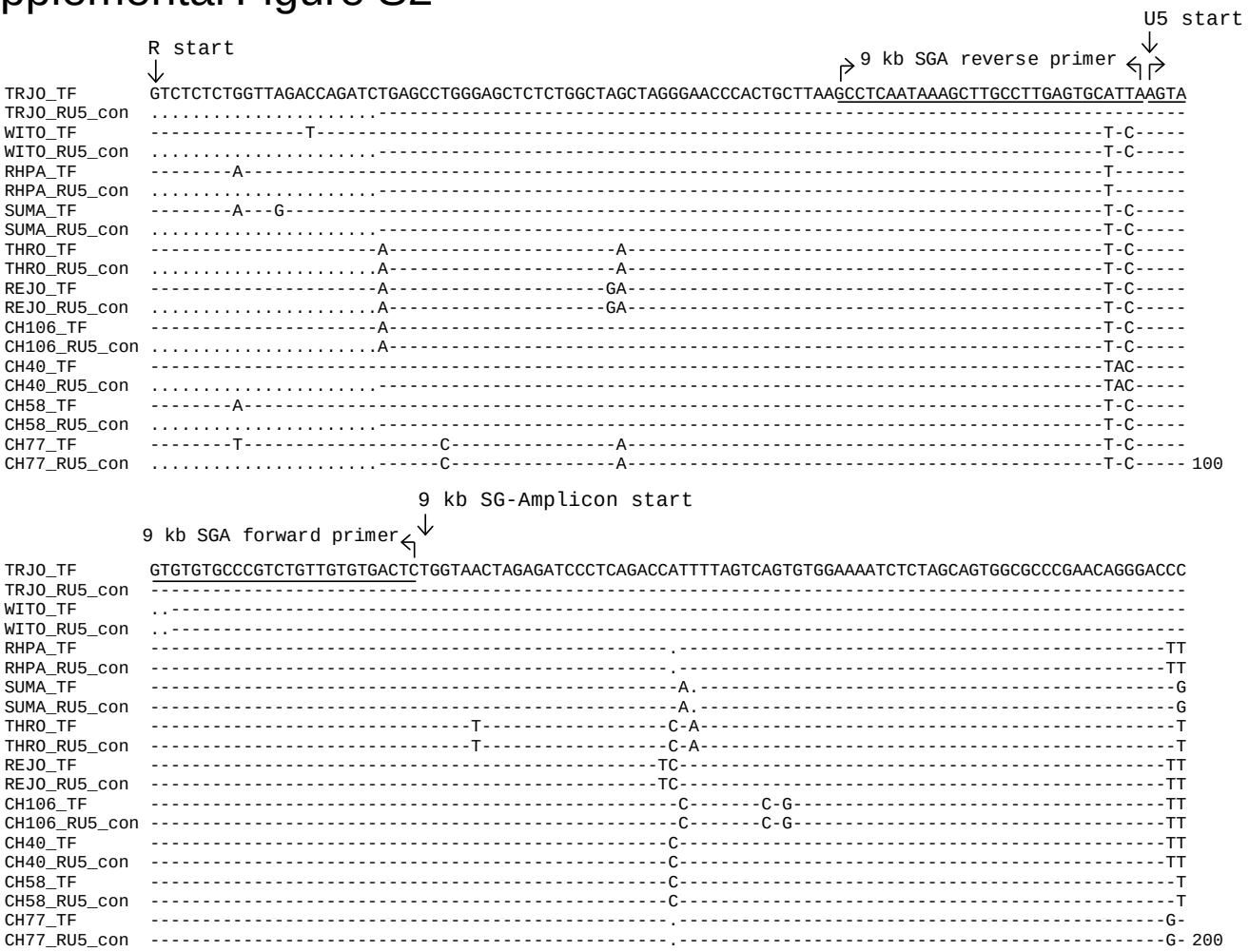

Figure S2. Identification of T/F sequence obscured by SGA primers in the LTR. Illustrated for each of the 10 subjects is an alignment of the R-U5-containing T/F virus nucleotide sequence deduced from 9 kb SG-amplicons, and a consensus sequence deduced from ~800 nt SG-amplicons (Con) that were independently generated and comprise most of R, all of U5, the primer binding site (PBS), and a 5' portion of *gag* (only the first 200 nt encompassing the R-U5 junction are shown). The underlined 60 nucleotides at the R/U5 junction represent the segment of each T/F sequence that could not be identified from the initial analysis of SG-amplicons due to the use of HXB2-specific primers. Furthermore note that in proviral DNA, repeat elements of R and U5 exist in the 5' and 3' LTRs, and in the initial generation and analysis of SG-amplicons a cDNA copy of the virus genome (5'-R/U5—U3/R-3') was amplified by annealing of the (reverse) R primer to the 3' R. Therefore, the consensus sequence deduced by analyzing R-U5-containing SG-amplicons generated from the 5' end of the genome, served to identify the missing T/F sequence in R.