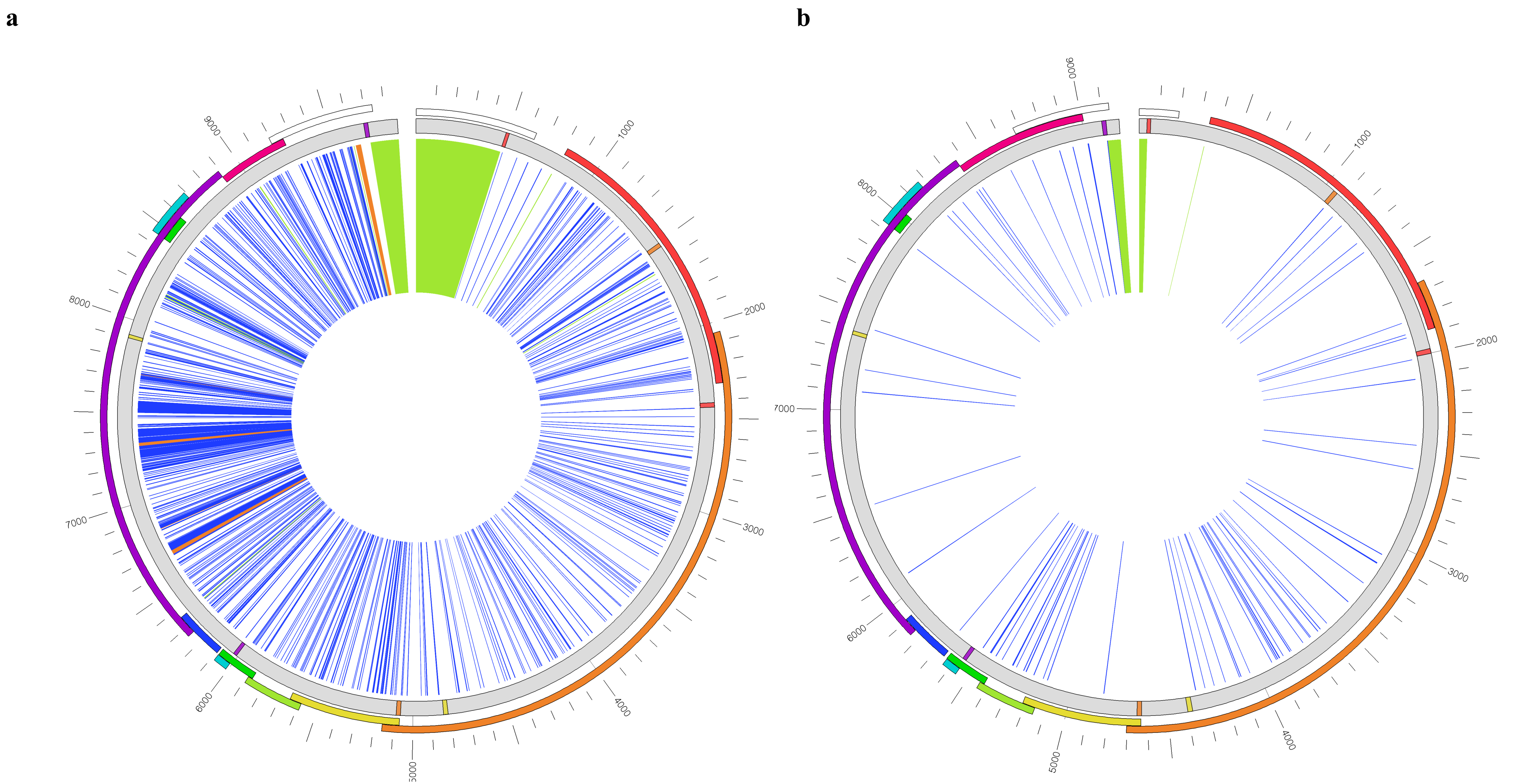

FIG S6 Genetic diversity of HIV-1 genes, groups and subtypes. Representative HIV-1 genomes of group O and N sequenced in this study are shown in a circular format and relative to the relevant reference sequence. They are displayed in the inner tracks. For each sequence, every nucleotide differing from the reference strain is shown as a blue line, an insertion is shown in orange and a deletion in green. The positions of primer sets 1 to 4 for the 'pan'-HIV-1 RT-PCR are shown as coloured lines in the reference strain. The genomes contain the complete coding sequence of HIV-1, a complete U5 and partial R region of the 5'-LTR, and a partial U3 region of the 3'-LTR. The outer tracks show the open reading frames in rainbow colours and a scale bar.

- **a** HIV-1 genome sequence of group O (strain #7) obtained by NGS is compared to the group O reference strain MVP5180 (acc. no. L20571) shown in grey.
- **b** HIV-1 genome sequence of group N reference strain YBF30 obtained by NGS (strain #5) is compared to the YBF30 GenBank sequence (acc. no. AJ006022) shown in grey. The sequences are 99.1% identical.