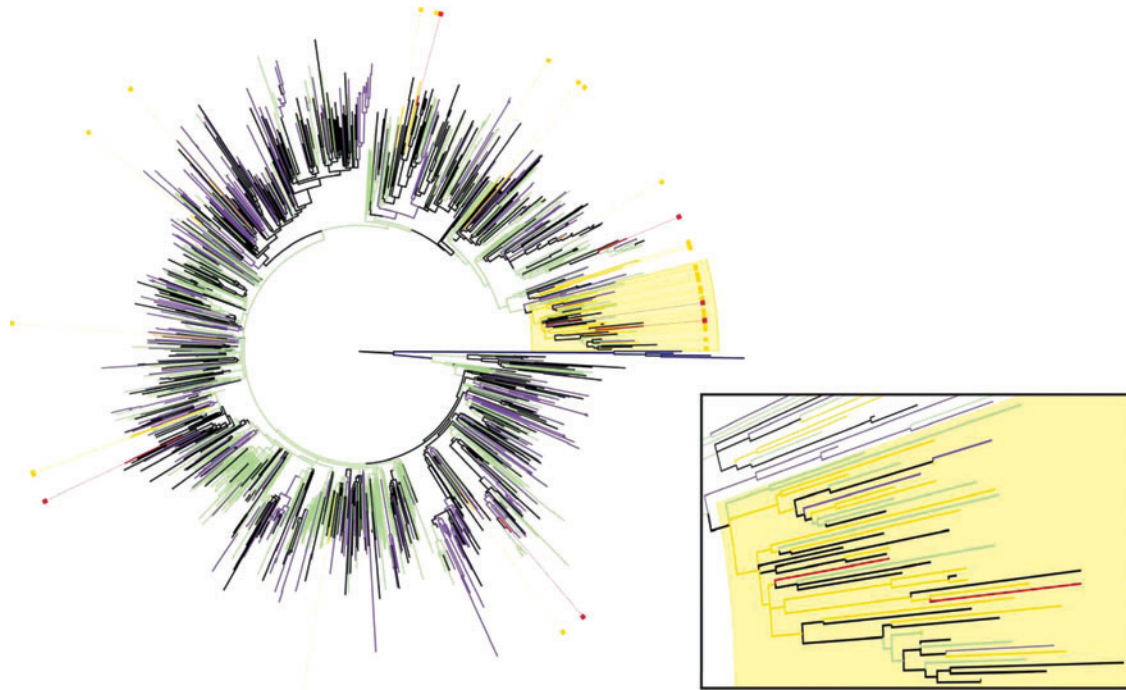# Supplementary Data

SUPPLEMENTARY FIG S1. Maximum likelihood tree of 1,121 subtype C sequences, constructed with the sequences from Portugal and Leuven cohorts complemented with controls retrieved from the BLAST search. Sequences enrolled in this study are colored *violet* (E138A absent), *yellow* (E138A present and SDRMs absent), or *red* (E138A present and SDRMs present). Sequences retrieved from BLAST are colored *green* (treatment-naive) or *black* (treatment experienced or treatment status unknown). The monophyletic cluster of 46 sequences is highlighted in *light yellow*, this cluster contained 51.4% (18/35) of the Portuguese treatment-naive sequences with the E138A. Most sequences were sampled in Portugal ($n=37$), but also in South Africa (3), Mozambique (2), Botswana (1), Poland (1), United States (1), and Zambia (1). SDRMs, surveillance drug resistance mutations.