Background: Tuberculosis (TB) remains a major health problem within the Community of Portuguese Language Speaking Countries (CPLP). Despite the marked variation in TB incidence across its member-states and continued human migratory flux between countries, a considerable gap in the knowledge on the Mycobacterium tuberculosis population structure and strain circulation between the countries still exists.

Materials/methods: We have assembled and analyzed the largest CPLP M. tuberculosis molecular and drug susceptibility dataset, comprised by a total of 1447 clinical isolates, including 423 multidrug-resistant isolates, from five CPLP countries. Genotyping analysis was carried out by 15/24 Mycobacterial Interspersed Repetitive Unit – Variable Number of Tandem Repeat (MIRU-VNTR) and Spoligotyping. Drug Susceptibility testing was
performed using standardized BACTEC 960 MGIT methodology or through the resazurin microtiter assay (REMA).

**Results:** The data herein presented reinforces Latin American and Mediterranean (LAM) strains as the hallmark of *M. tuberculosis* populational structure in the CPLP coupled with country-specific differential prevalence of minor clades. Moreover, using high-resolution typing by 24-loci MIRU-VNTR, six cross-border genetic clusters were detected, thus supporting recent clonal expansion across the Lusophone space.

To make this data available to the scientific community and public health authorities we developed CPLP-TB (available at http://cplp-tb.ff.ulisboa.pt), an online database coupled with web-based tools for exploratory data analysis.

**Conclusions:** As a public health tool, CPLP-TB is expected to contribute to improved knowledge on the *M. tuberculosis* population structure and strain circulation within the CPLP, thus supporting risk assessment of strain-specific trends.