**Data associated with the manuscript:**

**Pan genome-based delineation of species and subspecies in the genus *Mycobacterium***

**Authors:** Arup Panda1,2, Pierre Pontarotti2, Michel Drancourt1\*

\*Corresponding author

**Affiliations:**

1. Institut Hospitalo-Universitaire (IHU) Méditerranée Infection, Pôle des Maladies Infectieuses et Tropicales Clinique et Biologique, Fédération de Bactériologie-Hygiène-Virologie, Centre Hospitalo-Universitaire Timone, Assistance Publique-Hôpitaux de Marseille, Marseille, France.
2. Aix-Marseille Université, I2M, UMR-CNRS 7373, Evolution Biologique et Modélisation, Marseille, France

**Description:**

**Text files:**

**1. Genome\_cluster\_all\_23\_genomes.txt:** OrthoMcl clustering results for the genus level study with 23 completely sequenced mycobacterium genomes (one from each species).

**2. Abscessus\_cluster.txt:** OrthoMcl clustering results for *Mycobacetrium abscessus* genomes*.*

**3. Avium\_cluster.txt:** OrthoMcl clustering results for *Mycobacetrium avium* genomes.

**4. Bovis\_cluster.txt:** OrthoMcl clustering results for *Mycobacetrium bovis* genomes*.*

**5. Intracellulare\_cluster.txt:** OrthoMcl clustering results for *Mycobacetrium intracellulare.*

**6. Kansasii\_cluster.txt:** OrthoMcl clustering results for *Mycobacetrium kansasii* genomes*.*

**7. Smegmatis\_cluster.txt:** OrthoMcl clustering results for *Mycobacetrium smegmatis* genomes.

**8. Tuberculosis\_cluster.txt:** OrthoMcl clustering results for *Mycobacetrium tuberculosis* genomes.

**Excel files:**

**1. core\_pan\_values\_23\_mycobacterium\_genomes.xlsx:** Number of lineage-restricted, accessory, and core HGCs at the genus level study with 23 mycobacterium species.

**2. core\_pan\_values\_M\_abscessus.xlsx:** Number of lineage-restricted, accessory, and core HGCs at the species level study with *M. abscessus* genomes.

**3. core\_pan\_values\_M\_avium.xlsx:** Number of lineage-restricted, accessory, and core HGCs at the species level study with *M. avium* genomes.

**4. core\_pan\_values\_M\_bovis.xlsx:** Number of lineage-restricted, accessory, and core HGCs at the species level study with *M. bovis* genomes.

**5.** **core\_pan\_values\_M\_** **intracellulare.xlsx:** Number of lineage-restricted, accessory, and core HGCs at the species level study with *M. intracellulare* genomes.

**6. core\_pan\_values\_M\_kansasii.xlsx:** Number of lineage-restricted, accessory, and core HGCs at the species level study with *M. kansasii* genomes.

**7. core\_pan\_values\_M\_smegmatis.xlsx:** Number of lineage-restricted, accessory, and core HGCs at the species level study with *M. smegmatis* genomes.

**8. core\_pan\_values\_M\_tuberculosis.xlsx:** Number of lineage-restricted, accessory, and core HGCs at the species level study with *M. tuberculosis* genomes.

**9. functional\_enrichment\_result\_23\_mycobacterium\_genomes.xlsx:** Results of functional enrichment analysis between core and accessory HGCs of 23 mycobacterium species (genus level study).

**10. functional\_enrichment\_result\_M\_abscessus.xlsx:** Results of functional enrichment analysis between core and accessory HGCs of *M. abscessus* genomes (species level study).

**11. functional\_enrichment\_result\_M\_avium.xlsx:** Results of functional enrichment analysis between core and accessory HGCs of *M. avium* genomes (species level study).

**12. functional\_enrichment\_result\_M\_bovis.xlsx:** Results of functional enrichment analysis between core and accessory HGCs of *M. bovis* genomes (species level study).

**13. functional\_enrichment\_result\_M\_intracellulare.xlsx:** Results of functional enrichment analysis between core and accessory HGCs of *M. intracellulare* genomes (species level study).

**14.** **functional\_enrichment\_result\_M\_kansasii.xlsx:** Results of functional enrichment analysis between core and accessory HGCs of *M. kansasii* genomes (species level study).

**15.** **functional\_enrichment\_result\_M\_smegmatis.xlsx:** Results of functional enrichment analysis between core and accessory HGCs of *M. smegmatis* genomes (species level study).

**16. functional\_enrichment\_result\_M\_tuberculosis.xlsx:** Results of functional enrichment analysis between core and accessory HGCs of *M. tuberculosis* genomes (species level study).