

**Supplemental Material** (Supplemental Figures S1 – S2 and legends of Supplemental Tables S1 - S4)

**A whole genome-based gene-by-gene typing system for standardised high resolution strain typing of *Bacillus anthracis***

Mostafa Y. Abdel-Glil<sup>1,2#</sup>, Alexandra Chiaverini<sup>3</sup>, Giuliano Garofolo<sup>3</sup>, Antonio Fasanella<sup>4</sup>, Antonio Parisi<sup>4</sup>, Dag Harmsen<sup>5</sup>, Keith A Jolley<sup>6</sup>, Mandy C. Elschner<sup>1</sup>, Herbert Tomaso<sup>1</sup>, Jörg Linde<sup>1</sup>, Domenico Galante<sup>4</sup>

---

<sup>1</sup>Institute for Bacterial Infections and Zoonoses, Friedrich-Loeffler-Institut, Jena, Germany

<sup>2</sup>Department of Pathology, Faculty of Veterinary Medicine, Zagazig University, Sharkia Province, Egypt, Egypt

<sup>3</sup>Istituto Zooprofilattico Sperimentale dell'Abruzzo e del Molise "G. Caporale", Teramo, Italy

<sup>4</sup>Anthrax Reference Institute of Italy, Istituto Zooprofilattico Sperimentale della Puglia e Basilicata, Foggia, Italy

<sup>5</sup>Department of Periodontology and Operative Dentistry, University Hospital Muenster, Muenster, Germany

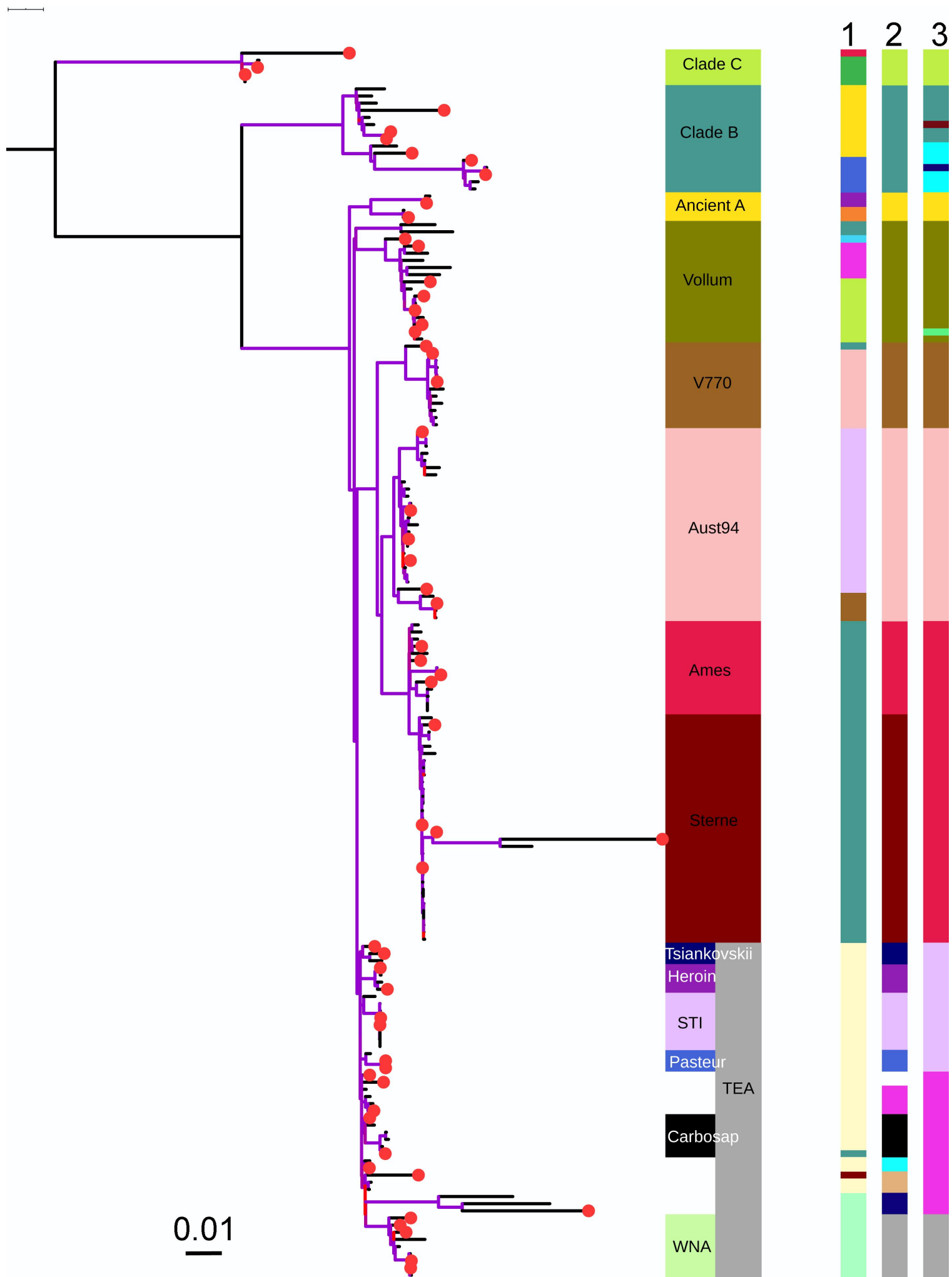
<sup>6</sup>Department of Zoology, University of Oxford, Oxford, United Kingdom.

---

## Supplemental Figures S1 – S2

---

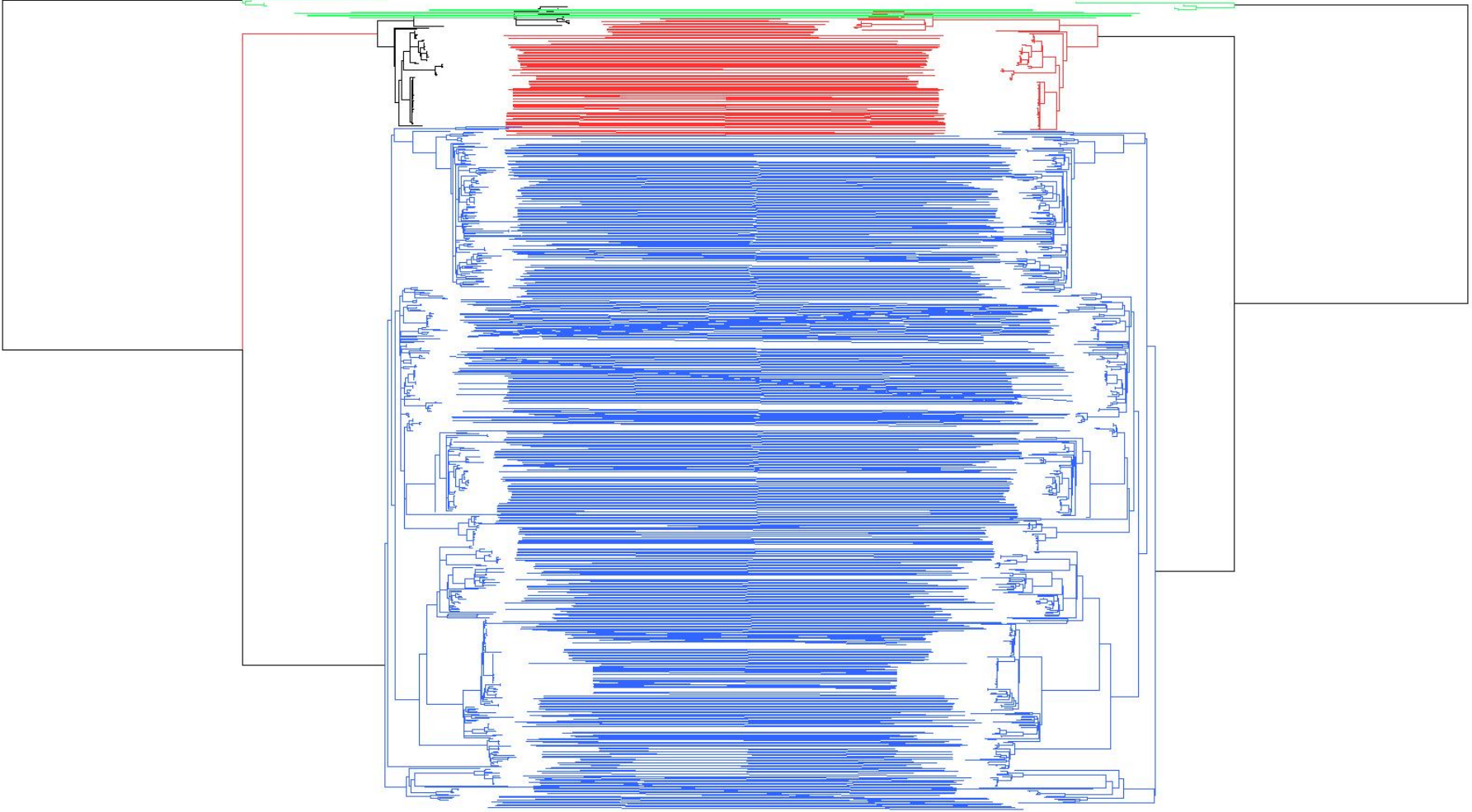
**Supplemental Figure S1:** Representativeness of *B. anthracis* strains selected and used for the calculation of the species core genome. Maximum likelihood phylogenetic tree constructed using 9,506 whole-genome SNPs based on a data set of 172 *B. anthracis* genomes that were downloaded from the RefSeq database. The corresponding genetic groups of the strains are shown in the inner column. Further annotation of the genomes in regard to genetic diversity was indicated in the outer columns numbered 1, 2 and 3 based on hierBAPS groups according to Bruce et al., 2019, updated canonical SNPs typing from Sahl et al., 2016 and traditional canonical SNPs grouping from Van Ert et al., 2007, respectively. Red dots at the tip of the branches indicate selected genomes for the calculation of the species core genome.



**Supplemental Figure S2:** Topological concordance between the cgMLST neighbor joining tree (left) and the SNP-based maximum likelihood tree (right) is represented as tanglegram. Tree branches were coloured based on the major clades of *B. anthracis*, clade A: blue, clade B: red and clade C: green.

1200.0

1200.0



## **Legends of Supplemental Tables S1 - S4**

**Supplemental Table S1:** Metadata information of strains and genomes of *B. anthracis* used in the setup and application of the core genome MLST system.

**Supplemental Table S2:** List and characteristics of the 3803 core genome MLST target genes used for *B. anthracis*.

**Supplemental Table S3:** Typeability percentages and variability of *B. anthracis* clades, average allele distance within and between each of the identified (sub)clades as well as the average allele distance of each (sub)clade to the closest sample in the other groups.

**Supplemental Table S4:** group-specific SNPs that are specifically present in each of the *B. anthracis* (sub)clade but simultaneously absent in all other genomes.