A genomic analysis of 50 random *Salmonella* Concord isolates identifies it as polyphyletic and highly resistant in Ethiopia

1. Introduction

Salmonella enterica serovar Concord is a non-typhoidal Salmonella serovar that has garnered significant attention due to its association with severe infections and antimicrobial resistance (Concord et al., 2011) . Its emergence from Ethiopia, linked to international adoptions (Fabre et al., 2009), raised public health concerns due to the frequent occurrence of multidrug resistance, including extensively drug-resistant (XDR) and even pandrug-resistant (PDR) strains (Cuypers et al., 2023). This study focuses on building a phylogenetic tree for S. Concord lineage A, a concerning lineage with high levels of multidrug resistance, to understand its population structure and evolutionary history.

1. Significance

The emergence and spread of antimicrobial-resistant Salmonella Concord lineage A pose a significant threat to public health, particularly in regions like Ethiopia. Understanding the evolutionary history and AMR patterns of this lineage is crucial for developing effective surveillance and control strategies to mitigate the impact of this public health concern.

1. Aims

This study aims to reconstruct the evolutionary relationships within S. Concord lineage A and characterize the distribution and genetic basis of antimicrobial resistance among its isolates.

1. Methods and Pipeline

The analysis focused on S. Concord lineage A isolates that had passed quality control, serotype confirmation, and lineage identification. Genomes were acquired from the NCBI assembly database (https://www.ncbi.nlm.nih.gov/assembly), and Snippy was used to generate a full SNP alignment file (https://github.com/tseemann/snippy). Recombination was removed with Gubbins (Croucher et al., 2015), and a SNP-only alignment was generated using SNP-sites (https://github.com/sanger-pathogens/snp-sites). RAxML-NG(Kozlov et al., 2019) was then employed to construct a maximum likelihood phylogenetic tree, revealing the evolutionary relationships among the isolates. Finally, pairwise SNP distances were calculated using snp-dists (https://github.com/tseemann/snp-dists) to assess genetic diversity within the lineage. Data visualization was performed using R with the Ggtree (Yu et al., 2017) and ComplexHeatmap (Gu et al., 2016) packages to create the final figure.

1. Results and Discussion

A maximum likelihood phylogenetic tree of the high-quality random 50 S. Concord lineage A isolates based on the non-recombinant SNP sites was generated (Figure 1). The tree showed four distinct monophyletic groups with each having a common ancestor and one large paraphyletic group whose full lineages could not be traced. The diverse lineages of S. Concord make it a polyphyletic serovar and edge out its unclear origination from only Ethiopia as hinted by previous works. The majority of the isolates, though sourced from the USA and the UK traced their origin to Ethiopia and Kenya respectively. Isolates sourced from Austria also had an origin in Ethiopia. Most of the isolates with lineages in Ethiopia harboured a combination of MDR and PDR as dominant genes, aligning seamlessly with previous findings as Ethiopia being a hotspot destination of highly resistant genes. Conclusively, we identify S. Concord as a diverse and widespread isolate and call for intervention, especially in low and middle-income countries like Ethiopia where resistant genes highly prevail.

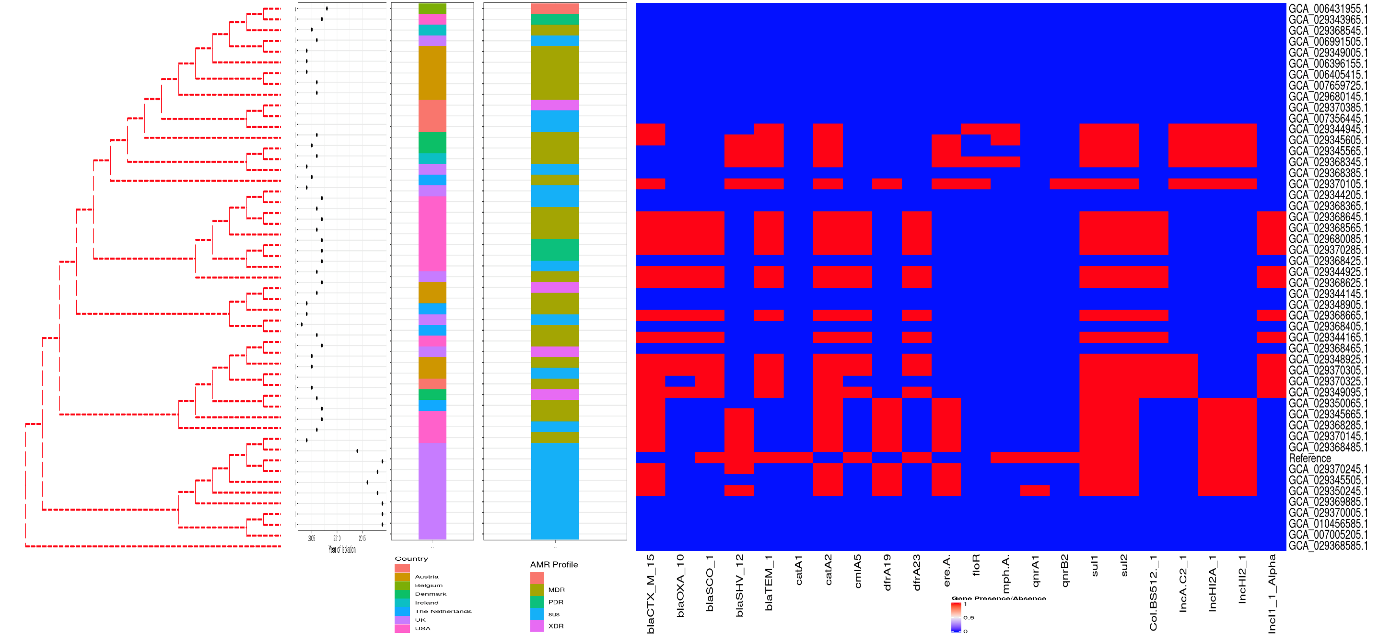


Figure 1

References:

Beyene, G. et al. (2011). *Multidrug resistant Salmonella Concord is a major cause of salmonellosis in children in Ethiopia* (Issue 85).

Croucher, N. J., Page, A. J., Connor, T. R., Delaney, A. J., Keane, J. A., Bentley, S. D., Parkhill, J., & Harris, S. R. (2015). Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Research*, *43*(3), e15. https://doi.org/10.1093/nar/gku1196

Cuypers, W. L., Meysman, P., Weill, F. X., Hendriksen, R. S., Beyene, G., Wain, J., Nair, S., Chattaway, M. A., Perez-Sepulveda, B. M., Ceyssens, P. J., de Block, T., Lee, W. W. Y., Pardos de la Gandara, M., Kornschober, C., Moran-Gilad, J., Veldman, K. T., Cormican, M., Torpdahl, M., Fields, P. I., … Van Puyvelde, S. (2023). A global genomic analysis of Salmonella Concord reveals lineages with high antimicrobial resistance in Ethiopia. *Nature Communications*, *14*(1). https://doi.org/10.1038/s41467-023-38902-x

Fabre, L., Delauné, A., Espié, E., Nygard, K., Pardos, M., Polomack, L., Guesnier, F., Galimand, M., Lassen, J., & Weill, F. X. (2009). Chromosomal integration of the extended-spectrum β-lactamase gene blaCTX-M-15 in Salmonella enterica serotype concord isolates from internationally adopted children. *Antimicrobial Agents and Chemotherapy*, *53*(5), 1808–1816. https://doi.org/10.1128/AAC.00451-08

Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, *32*(18), 2847–2849. https://doi.org/10.1093/bioinformatics/btw313

Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAxML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, *35*(21), 4453–4455. https://doi.org/10.1093/bioinformatics/btz305

Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T. Y. (2017). ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, *8*(1), 28–36. https://doi.org/10.1111/2041-210X.12628