Supplemental Digital Content 2. Methods

All the *Streptococcus pyogenes* isolates were cultured on Columbia agar plate with 5% sheep blood (bioMérieux, Marcy-l'Étoile, France) and incubated at 37 °C overnight in 5% CO_2 atmosphere. Colonies grown on agar plates were identified by using Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry (MALDI-TOF MS; Bruker Daltonics, Bremen, Germany). *S. pyogenes* strains were tested for their antimicrobial susceptibility by VITEK®2 (bioMérieux, Marcy-l'Étoile, France) automated system and results were interpreted according to clinical breakpoints based on the European Committee on Antimicrobial Susceptibility Testing (EUCAST) tables (version 13.0)[1].

Bacterial DNA was extracted using the automatic extractor EZ1 (Qiagen BioRobot EZ1), with the extraction kit (EZ1&2 DNA tissue kit, Qiagen, Germany), following the manufacturer's instructions and setting the elution volume at 50 µl. Extracted DNA was quantified using Bioanalyzer Instrument and Next Generation Sequencing library preparation was performed according to manufacturer's protocol with DNAprep kit (Illumina, San Diego, California, USA). Prepared libraries were sequenced with an Illumina NextSeq 550 sequencing platform using a NextSeq 500/550 v2.5 Kits in paired end (150x2).

Whole genome sequencing analysis was conducted using the Bactopia pipeline (v3.0.0)[2], including read pre-processing with Fastp[3] and FastQC[4], *de novo* assembly with Shovill[5] and Pilon[6], and assembly with Prokka[7]. Quality of the assemblies was evaluated using Quast (v5.1)[8]. Emm typing and Multilocus Sequence Typing (MLST) prediction were performed with emm-typer (v.0.2.0)[9] and mlst (v2.23.0)[10], respectively. Antibiotic resistance (AMR) genes were investigated with ABRicate (v1.0.1)[11], using the Comprehensive Antibiotic Resistance Database (CARD)[12] with 90% coverage and identity, while virulence factors with a combination of ABRicate, using the Virulence Factor Database (VFDB)[13] with 70% coverage and identity, and BLAST, searching for the main *S. pyogenes* virulence factors described in literature. Prophages were investigated using PhiSpy (v4.2.21)[14] and manually curated by searching the identified sequences with the Microbial Nucleotide BLAST

(available at https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&BLAST_SPEC=MicrobialGenomes), using the bacteriophages database. Mutations in *csrR/S* genes were evaluated by mapping reads against the MGAS5005 reference strain (GenBank accession number NC_007297) using bwa mem algorithm[15], followed by variant calling with freebayes (v1.3.6)[16] and conversion of DNA to amino acid sequences by the EMBOSS transeq tool (EMBL-EBI).

Single nucleotide polymorphism (SNP) calling was performed with Snippy (v4.6.0)[17], using MGAS5005 as reference. The coreSNP obtained after removing putative recombinogenic regions predicted by Gubbins (3.3.1)[18] was used to infer phylogenetic relationships among whole genome sequences by Maximum-Likelihood (ML) using IQTREE (v2.2.5)[19] with 1,000 bootstrap replicates under the best nucleotide substitution model (GTR+F+G4) determined by ModelFinder[20]. The ML tree was visualized and annotated using iTOL (v6.5.2)[21]. Pairwise SNP distances were calculated using snp-dists tool (https://github.com/tseemann/snp-dists).

REFERENCES

1. EUCAST. https://www.eucast.org/clinical_breakpoints/.

2. Petit RA 3rd, Read TD. Bactopia: a Flexible Pipeline for Complete Analysis of Bacterial Genomes. mSystems. 2020 Aug 4;5(4):e00190-20. doi: 10.1128/mSystems.00190-20.

3. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018 Sep 1;34(17):i884-i890. doi: 10.1093/bioinformatics/bty560.

4. Andrews, S. (2010). FastQC: a Quality Control Tool for High Throughput Sequence Data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

5. Seemann, T. Shovill: Faster SPAdes (or better SKESA/Megahit/Velvet) Assembly of Illumina Reads. https://github.com/tseemann/shovill (2018)

6. Walker, B. J. et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9, e112963 (2014).

7. Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069. doi: 10.1093/bioinformatics/btu153

8. Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. Bioinformatics 29, 1072–1075. doi: 10.1093/bioinformatics/btt086

9. Microbiological Diagnostic Unit Public Health Laboratory (2021). Emmtyper—Emm Automatic Isolate Labeller (v0.2.0). Available online at: https://github.com/MDU-PHL/emmtyper

10. Seemann, T. (2022). mlst Tool. Available online at: https://github.com/tseemann/mlst (accessed May 24, 2023).

11. Seemann, T. (2020). Abricate. Available online at: https://github.com/tseemann/abricate.

12. Jia, B., Raphenya, A. R., Alcock, B., Waglechner, N., Guo, P., Tsang, K. K., et al. (2017). CARD 2017, expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucl. Acids Res. 45, D566–D573. doi: 10.1093/nar/gkw1004

13. Chen, L., Zheng, D., Liu, B., Yang, J., and Jin, Q. (2016). VFDB 2016, hierarchical and refined dataset for big data analysis−10 years on. Nucl. Acids Res. 44, D694–D697. doi: 10.1093/nar/gkv1239

14. Akhter S, Aziz RK, Edwards RA. PhiSpy: a novel algorithm for finding prophages in bacterial genomes that combines similarity- and composition-based strategies. Nucleic Acids Res. 2012 Sep;40(16):e126. doi: 10.1093/nar/gks406.

15. Heng Li, Richard Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform, Bioinformatics, Volume 25, Issue 14, July 2009, Pages 1754–1760, https://doi.org/10.1093/bioinformatics/btp324

16. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. arXiv 2012;1207.3907. https://doi.org/10.48550/arXiv.1207.3907

17. Seemann, T. (2015). Snippy: Fast Bacterial Variant Calling From NGS Reads. Available online at: https://github.com/tseemann/snippy

18. Croucher, N. J., Page, A. J., Connor, T. R., Delaney, A. J., Keane, J. A., Bentley, S. D., et al. (2015). Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucl. Acids Res. 43, e15. doi: 10.1093/nar/gku1196

19. Nguyen, L. T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2017). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274. doi: 10.1093/molbev/msu300

20. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589. doi: 10.1038/nmeth.4285

21. Letunic, Ivica, and Peer Bork. “Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation.” *Nucleic acids research* vol. 49,W1 (2021): W293-W296. doi:10.1093/nar/gkab301