

Practicability of Molecular Typing Schemes for *Staphylococcus aureus* in the Genomic Era



Nir Gilad¹, Michal Ziv-Ukelson¹, Vered Chalifa-Caspi¹ and Jacob Moran-Gilad^{1,2,3}

¹Ben-Gurion University of the Negev, Beer Sheva, Israel; ²Public Health Services, Ministry of Health, Jerusalem, Israel;

³ESCMID Study Group for Molecular Diagnostics

Contact details:
Prof. Jacob Moran-Gilad
giladko@post.bgu.ac.il

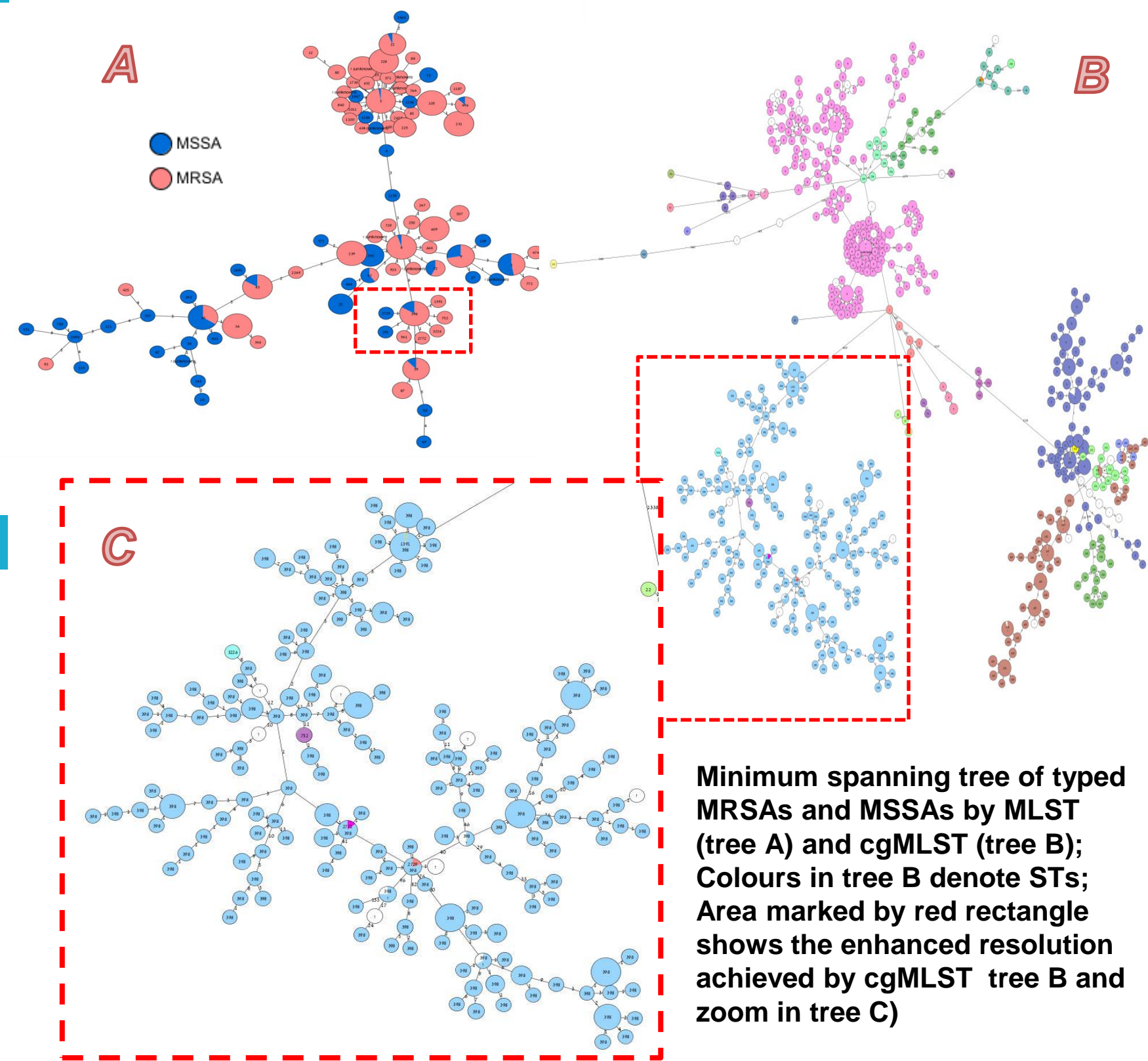
Introduction

- There are several accepted typing schemes for *S. aureus* which may be used for outbreak investigation and microbial population analysis such as *spa* typing and MLST.
- The introduction of whole genome sequencing (WGS) mandates the development of new typing schemes whilst maintaining reverse compatibility.
- Indeed, a core genome (cg) based MLST has recently been proposed as a highly discriminatory yet standardised approach.
- We sought to evaluate the practicability of applying these typing scheme on a large convenience sample of publicly available *S. aureus* genomes.

Methods

- *S. aureus* genome assemblies were downloaded from NCBI and a local database was created
- Genomes were checked for quality using QUAST, and subsequently analysed using an in house *S. aureus* pipeline (Fig. 1).
- Typing was performed by extracting the *spa* locus as well as the 7 MLST gene loci and querying against the respective databases.
- cgMLST analysis and construction of minimum spanning trees was performed using SeqSphere v.3.0.1 (Ridom GmbH, Munster, Germany).
- The Simpson's diversity index (DI) was calculated for each method and scheme performance was compared (Table 1).

Fig. 2 Phylogenetic analysis by MLST and cgMLST



Results

- Overall, 4,207 genomes were analysed, of which 4,032 were typeable by MLST (95.8%).
- There were 84 known STs identified, of which ST5 accounted for 40%, ST8 18%, ST398 15% and ST105 13% of samples.
- Additional 38 STs included 2 or more samples (range 2-84) and 42 STs were singletons (Fig. 2A).
- Nine novel STs were found in 28 samples.
- *spa* typing was unsuccessful for 557 samples and in 274 presumably new types were found.
- For cgMLST, a 'cluster type' was assigned for 2,412 samples (57%) and included 121 different types according to the current scheme (Fig. 2B/C).
- However, the vast majority of samples were typeable when 'unknown targets' were not excluded.

Fig. 1 *S. aureus* bioinformatics pipeline

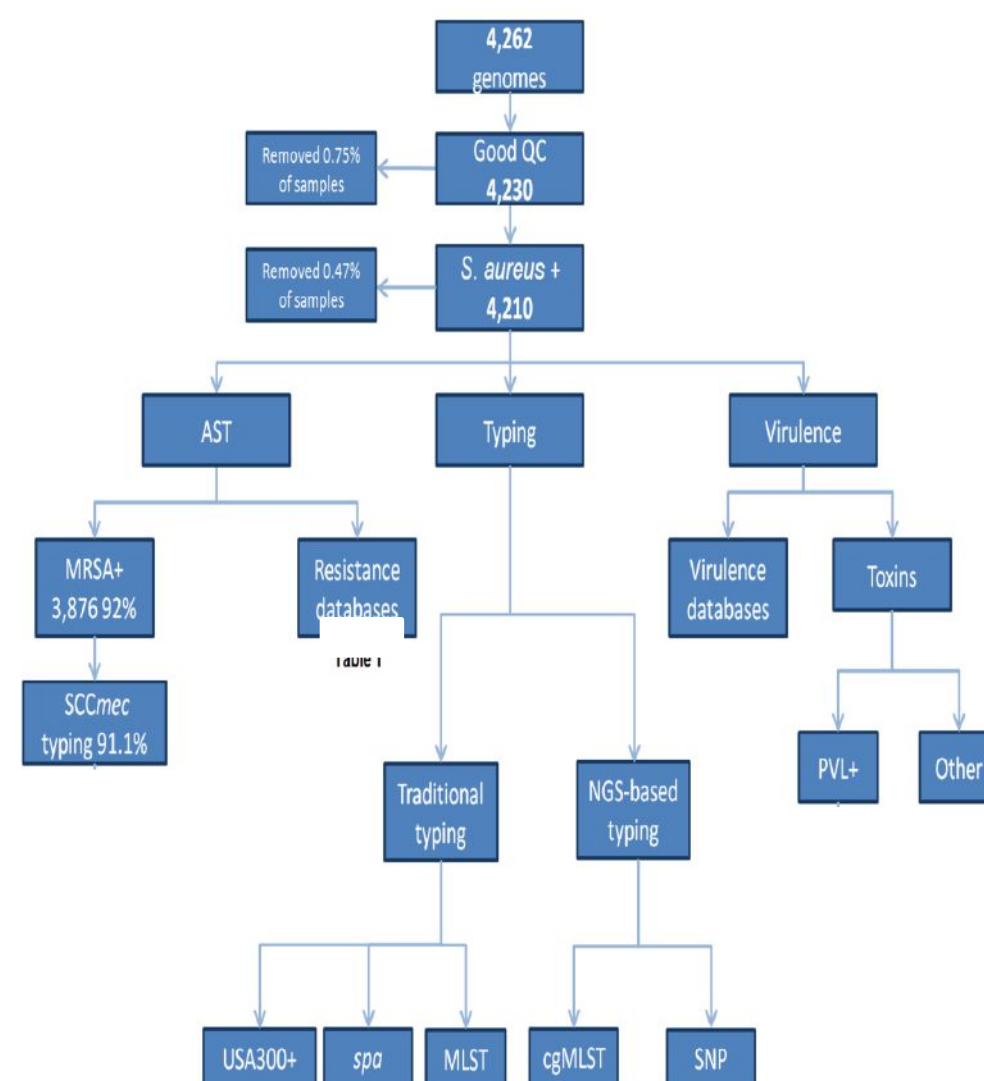


Table 1. Simpson's DI for typing schemes

Method	loci in scheme	Typeability	No. of types	DI
MLST	7	95.8%	93	0.767
<i>spa</i>	1	86.8%	487	0.880
cgMLST Cluster Types	1,861	56.9%	121	0.888
cgMLST with unknown targets	1,861	100%	4,029	1.0

Conclusions

- Analysis of a large sample of genomes demonstrated satisfactory reverse compatibility of WGS with MLST, whereas *spa* typing showed a higher DI compared to MLST but also a higher failure rate that should be further investigated.
- cgMLST exhibited a maximal DI that emphasises its potential to become a standard typing tool for WGS typing of *S. aureus* should the nomenclature is implemented.
- Application of the schemes on a large collection of publicly available genomes adds to the understanding of the global diversity of *S. aureus*.