Whole genome sequence analysis of Staphylococcus saprophyticus involved in urinary tract infections

Kodjovi Mlaga¹, Cédric Abat², Jean-Marc Rolain*³

¹Aix Marseille Université, Umr7278-Umrite-Unité de Recherche Sur Les Maladies Infectieuses, Tropicales Et Emmergentes, Marseille, France
²Ihu Méditerranée Infection, Marseille, France
³Ihu Méditerranée Infection, Marseille Cedex 05, France

Background: *Staphylococcus saprophyticus* is a leading cause of Urinary Tract Infection (UTIs) diseases. Recently, our surveillance system identified an abnormal increase of *S. saprophyticus* UTIs in four university hospitals in Marseille. Suspecting a possible community *S. saprophyticus* UTIs outbreak, we conducted an epidemiological investigation to confirm the increase of *S. saprophyticus* UTIs in Marseille and decipher specific and evolutionary genomic features of the bacteria.

Material/methods: We analyzed retrospective 13 years historical clinical data from four university hospitals of Marseille. Using the R software, we built linear models on *S. saprophyticus*/*E. coli* UTIs ratios to analyse historical trend of the ratio with p-values <0.05 considered as statistically significant. We sequenced the genome of a *S. saprophyticus* strain isolated in our settings using the MiSeq Technology with the mate pair strategy. The reads were assembled using a5-Miseq program and annotated using Prokka. A pan genome analysis with others *S. saprophyticus* reference was performed using Genoscope platform. Hemaglutination activity test and real-time PCR targeting the core genome and accessory genome of urinary strains to discriminate between urinary and non-urinary strains were performed on 35 *S. saprophyticus* UTIs from Marseille.

Results: From January 2002 to December 2014, 776 patients experienced *S. saprophyticus* UTIs. Majority were females (93.9%) and aged between 11 and 30 years (76.7%). Over the same period, 39,095 patients experienced *E. coli* UTIs. *S. saprophyticus*/*E. coli* UTIs ratios increased 3.9 fold from 0.84 to 3.28, with an annual estimated ration trend of $6.10^{-4}$ (p-value < $10^{-3}$). The sequence length of the sequenced genome is 2,523,588 bp with 21 scaffolds (GC content: 33.27%), with 2,584 genomic ORFs. Genomic analysis revealed several putative virulence genes classified as biofilm formations, intercellular aggregation, toxin/antitoxin modules. Pangenome analysis contains 10,747 genes, with 3,541 families of genes, 8,677 core genes and 2,070 accessory genes. The genome of interest, *S. saprophyticus* G764, exhibited 51 specific genes with six regions of plasticity including 3 specifics regions and a genome reduction of 100kb (70% of them belonging to transcriptional regulatory COG group). We observe an expected translation termination within the uro-adherence gene (uafA) among non-urinary strains previously described as one of the determinant factor of uro-pathogenicity of *S. saprophyticus*. 85.71% (30/35) of the strains tested by hemagglutination showed positive hemagglutination, and 65.21% (15/23) positive to URP and BAP genes.

Conclusions: Our study showed the emergence of a community UTIs outbreak due to uro-pathogen *S. saprophyticus*. Uro-pathogenicity of *S. saprophyticus* correlates with full hemagglutination. Genome analysis shows high plasticity indicating possible evolution of uro-pathogen strains. Further analysis are ongoing to decipher the relationship between the clinical isolates from this outbreak.