Comparative whole genome analysis of *Staphylococcus saprophyticus* reveals mutation of uro-adherence factor A (*UafA)* gene in non-urinary strains and a clonal expansion in urinary tract infections in Marseille.

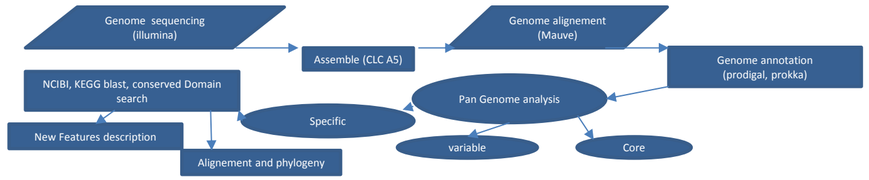
Kodjovi D. Mlaga; Cedric abat;Jean-Marc Rolain
Unité de recherche sur les maladies infectieuses et tropicales émergentes (URMITE) CNRS-IRD UMR 6236, Méditerranée Infection,
Faculté de Médecine et de Pharmacie, Aix-Marseille-Université, Marseille, France.

Introduction

- *Staphylococcus saprophyticus* is one of the leading cause of Urinary Tract Infections (UTIs) diseases in young ages, especially in female, accounting for up to 40% in this population.
- Recently, in Marseille, we have observed an abnormal increase of reported cases from January 2002 to December 20014, suspecting a possible outbreak or spread of clustered strains. Virulence factors of *S. saprophyticus* involved in UTIs includes adherence to urothelial cells by means of a surface-associated protein, lipase, lipotheicoic acid, hemagglutinin that bind to fibronectin, urease, and production of extracellular slime. A whole genome sequencing of *S. saprophyticus subsp. Saprophyticus 15305* isolated from a young women urine has revealed a single orf predictable as a cell-wall-anchored protein and it showed a positive hemagglutination and adherence to human bladder cells[11]
- We aim to conduct an epidemiological investigation within the suspected outbreak period to confirm the increasing of UTIs due to *Staphylococcus saprophyticus* in Marseille as well as performing a real time genomic to decipher any specific genomic features that may explained this phenomenon.

Materials and Methods

- We performed a retrospective statistical analysis of *S. saprophyticus* involved in UTIs diseases using the R software.
- A *S. saprophyticus_G764* whole genome was sequenced , assembled, annotated and analyzed using various bioinformatics tools.



Bioinformatics analysis workflow

Results

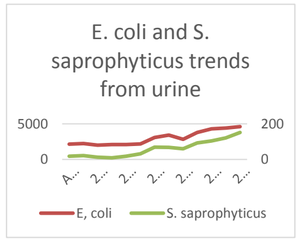


Figure 1: S. saprophyticus and Escherichia coli cases reported in Marseille during the study period

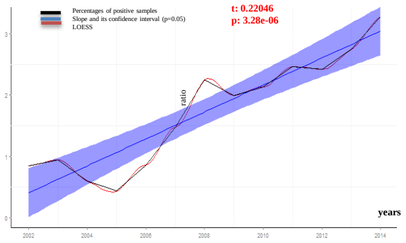


Figure 3: Ratios between the number of patients experiencing UTI infections due to *S. saprophyticus* versus *E. coli*

Over the same period, 39 ,095 patients have experienced *E. coli* UTI. Throughout the study period, *S. saprophyticus/E. coli* UTIs ratios increased 3.9 fold from 0.84 in 2002 to 3.28 in 2014, with an annual estimated trend of the ratio of 6.10^{-4} (p-value < 10^{-3}) (Figure 3). We observed a significant increasing of the number of *S. saprophyticus* UTIs (from 2 138 *vs.* 18 in 2002, to 4 599 *vs.* 151 in 2014, p-value < 10^{-3}).

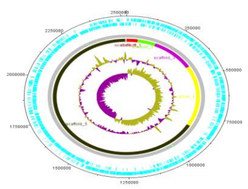


Fig. 3. Schematic circular diagram of the *S. saprophyticus* chromosome. From the outside, the five circles display: (i) scale in Mbp; (ii)predicted coding regions transcribed in the clockwise or counterclockwise (iii) tRNA and rRNA (green);(iv)GC deviation (G-C/G+C); (v) G+C content The genome size 2,523,588 bp ; 21 scaffolds. GC content is 33.27%. genomic features is 2660, accounting for 2523 CDSs, 63 tRNA, 23 rRNA; 767 CDSs with unknown function, 909 classified CDSs and 847 unclassified.

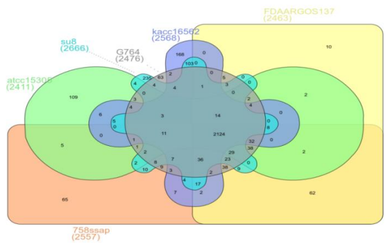
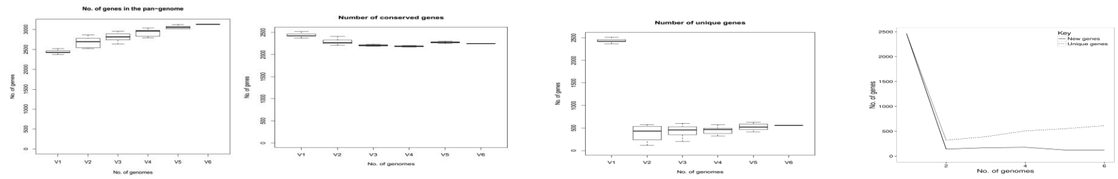


Figure 4:Venn diagram

*Show statistical difference compare to KACC 16562 and SU8 both non-clinical isolates p-value <0.0000001

- Pangenome analysis shows that clinical strains are phylogenically more closed.
- The highest average nucleotide similarity is among clinical isolates.
- Core genome is stable while Pangenome opened.
- New genes seem to follow the same trend as unique genes.
- These show that the population of S. saprophyticus can still evolve enabling the bacteria to adapt to changing environment.
- Majority of virulence genes belong to the core genome. Few of them are accessories.



Pangenome open | Core genome stable | Increasing unique genes | New/unique

Figure 5 : Graphs show distribution and evolution of various components of pan-genome: 1- total genes evolution of the Pan-genome 2-core or conserved genes evolution 3- unique genes evolution 4- number of blastp hit with different percentage identity 5- association between new genes and unique genes in the pan-genome.

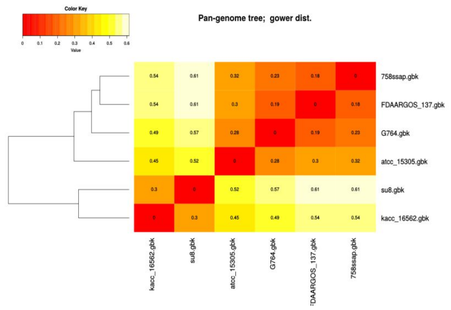


Figure 6: Heatmap derived from an average nucleotide identity matrix calculated with Get-homologues of S. saprophyticus isolated from various samples showing low average nucleotide similarity among clinical isolates

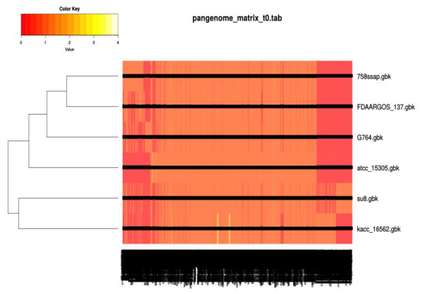


Figure 7: Heatmap of the previous pan genome matrix, with dendrogram sorting genomes and sequences

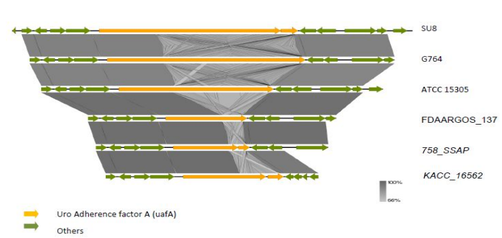


Figure 8: Genome sub-region visualization showing codon stop mutation that fragmented uro adherence protein into two fragment within non clinical strains and genes size variation.

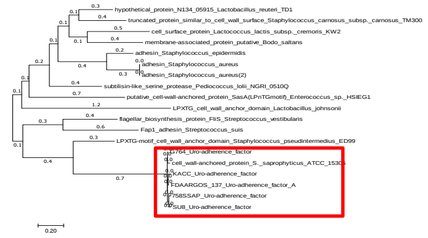


Figure 9: Molecular Phylogenetic analysis by Maximum Likelihood method of Uro-adherence protein

Table 1: Sex and Age distribution of patient from whom isolates were collected for hemagglutination test

			Marseille		Nice		
			Hemagglutinin				
			Neg	Pos	Neg	Pos	total
	male		0	1	1	1	3
Sex	female		4	29	54	23	110
total			4	30	55	24	113
	0-10		0	0	0	0	0
		44136	2	9	11	3	25
	21-30		1	14	25	12	52
	31-40		0	2	11	4	16
	41-50		1	1	6	4	12
	51-60		0	4	1	0	5
	61-70		0	0	0	1	1
	71+		0	0	1	0	1
total			4	30	55	24	112

	ATCC 15305	G764	SU8
hem	+	+	-

p-value<0.0000001, strains from Marseille showed positive hemagglutination test compare to that of Nice with a significant statistical difference.
Female were more affected than men p-value < 0.00000001

Conclusions

Our study confirmed clonal expansion of S. saprophyticus in a community UTIs outbreak. This clonal expansion was associated with positive hemagglutination enabling the bacteria to be more pathogenic. Negative hemagglutination correlates with mutation of UafA gene in non-clinical strains. Pan -genome analysis shows high plasticity, stable core genome and open pan-genome indicating possible evolution and emerging of uro-pathogen strains. No new virulence genes have been found to be associated with uro-pathogenicity of S. saprophyticus_G764. Hemagglutination test can be used as surrogate test and a surveillance marker for detecting S. saprophyticus uro-pathogen strains in the lab as well as biological surveillance tools.