Whole-Genome Sequencing of Seven Strains of *Staphylococcus lugdunensis* Allows Identification of Mobile Genetic Elements

Xavier Argemi[1,2,*], Véronique Martin[3], Valentin Loux[3], Sandrine Dahyot[4], Jérémie Lebeurre[4], Aurélien Guffroy[5,6], Mickael Martin[5,6], Aurélie Velay[7,8], Daniel Keller[2], Philippe Riegel[2], Yves Hansmann[1,2], Nicodème Paul[9], and Gilles Prévost[2]

[1]Hôpitaux Universitaires, Maladies Infectieuses et Tropicales, Strasbourg, France

[2]Université de Strasbourg, CHRU de Strasbourg, VBP EA7290, Fédération de Médecine Translationnelle de Strasbourg (FMTS), Institut de bactériologie, Hôpitaux Universitaires de Strasbourg, France

[3]INRA – Unité Mathématiques et Informatique Appliquées, du Genome à l'Environnement (MaIAGE), Jouy-en Josas, France

[4]Laboratoire GRAM EA2656, Université de Rouen – IRIB UFR Médecine-Pharmacie Batiment Recherche, Rouen, France

[5]Service d'Immunologie Clinique et de Médecine Interne, Centre National de Référence des Maladies Auto-immunes Rares, Hôpitaux Universitaires de Strasbourg, Strasbourg, France

[6]CNRS UPR 3572, Immunopathologie et Chimie Thérapeutique/Equipe, Tolérance Cellulaire B et Auto-immunité, Laboratoire d'excellence Medalis, Institute of Molecular and Cellular Biology (IBMC), Strasbourg, France

[7]Virology Laboratory, University Hospital of Strasbourg, Strasbourg, France

[8]2-INSERM, UMR_S1109, LabEx Transplantex, Fédération de Médecine Translationnelle de Strasbourg (FMTS), Université de Strasbourg, Strasbourg, France

[9]Laboratoire d'ImmunoRhumatologie Moléculaire, INSERM UMR_S 1109, Plateforme GENOMAX, Faculté de Médecine, Fédération Hospitalo-Universitaire OMICARE, Fédération de Médecine Translationnelle de Strasbourg (FMTS), Université de Strasbourg, Strasbourg, France

*Corresponding author: E-mail: xavier.argemi@chru-strasbourg.fr.

Accepted: April 20, 2017

Data Deposition: The complete genome sequences have been deposited at GenBank under the accessions CP020769-CP020770, CP020761, CP020764-CP020765, CP020766-CP020767-CP020768, CP020735-CP020736, CP020763, CP020762. Available from: https://www.ncbi.nlm.nih.gov/genome/genomes/2548?.

Abstract

Coagulase negative staphylococci are normal inhabitant of the human skin flora that account for an increasing number of infections, particularly hospital-acquired infections. *Staphylococcus lugdunensis* has emerged as a most virulent species causing various infections with clinical characteristics close to what clinicians usually observe with *Staphylococcus aureus* and both bacteria share more than 70% of their genome. Virulence of *S. aureus* relies on a large repertoire of virulence factors, many of which are encoded on mobile genetic elements. *S. lugdunensis* also bears various putative virulence genes but only one complete genome with extensive analysis has been published with one prophage sequence (ϕSL2) and a unique plasmid was previously described. In this study, we performed de novo sequencing, whole genome assembly and annotation of seven strains of *S. lugdunensis* from VISLISI clinical trial. We searched for the presence of virulence genes and mobile genetics elements using bioinformatics tools. We identified four new prophages, named ϕSL2 to ϕSL4, belonging to the *Siphoviridae* class and five plasmids, named *pVISLISI_1* to *pVISLISI_5*. Three plasmids are homologous to known plasmids that include, amongst others, one *S. aureus* plasmid. The two other plasmids were not described previously. This study provides a new context for the study of *S. lugdunensis* virulence suggesting the occurrence of several genetic recombination' with other staphylococci.

Key words: *Staphylococcus lugdunensis*, plasmids, prophages, virulence, next generation sequencing.

Background

Staphylococcus lugdunensis is a coagulase negative staphylococci (CoNS) that displays an unusual rate of virulence close to *Staphylococcus aureus* (Babu and Oropello 2011; Argemi et al. 2015; Douiri et al. 2016). This commensal bacterium has been mainly involved in nosocomial infections affecting debilitated patients with epithelial barrier breaches. But this bacteria also causes severe community-acquired infection such as endocarditis, skin, and soft tissues infections with necrosis or septic shock that might be due to virulence factors production (Delaunay et al. 2014; Hung et al. 2012; Sabe et al. 2014; Woznowski et al. 2010; Pareja et al. 1998). Until now, four complete genome sequences have been published and seven partially finished sequences using next generation sequencing (NCBI). The genome sequence of *Staphylococcus lugdunensis* N920143 (NCBI reference sequence NC_017353.1) by Heilbronner et al. identified various putative virulence factors, a single prophage named ϕSL1, and 14 insertion sequences (Heilbronner et al. 2013). One cadmium resistance plasmid named pLUG10 (NCBI reference sequence NC_002093.1) was also described in several strains of *S. lugdunensis* homologous to pOX6 *S. aureus* plasmid (Poitevin-Later et al. 1992). Mobile genetics elements (MGE) such as phages, plasmids, and pathogenicity islands have been widely studied in *S. aureus* as they encode numerous virulence factors but their occurrence in CoNS remains scarce (Malachowa and DeLeo 2010). This accessory genetic material might represent up to 25% in the *S. aureus* genome and contributes to the phenotypic plasticity of this pathogen but also its virulence as a majority of the virulence factors described in this bacteria are located on MGE (Otto 2014). Phage-encoded virulence factors have not been observed in CoNS and only few prophages have been described in clinical isolates, mainly from *Staphylococcus epidermidis* and *S. hominis* showing close relationships with *S. aureus* phages and prophages (Deghorain and Van Melderen 2012; Deghorain et al. 2012). Plasmids are implicated in the dissemination of multidrug resistance genes especially in *S. aureus* in hospital settings but can also bear toxin genes (Shintani et al. 2015; McCarthy and Lindsay 2012). Their descriptions in CoNS are rare and probably underestimated, most of the plasmids described were found in *S. epidermidis* strains (European Nucleotide Archive 2016). Finally, pathogenicity islands are widespread in *S. aureus* and usually carry one or more virulence factors such as superantigens (Sato'o et al. 2013). Although reports of toxigenic genes presence in CoNS pathogenicity islands can be found in the literature many author still question their existence and only three strains of *S. epidermidis* with an enterotoxin C-like-bearing pathogenicity island have been described (Madhusoodanan et al. 2011; Nanoukon et al. 2016). Thus, we conducted a genomic study on *S. lugdunensis* aiming to find elements that might explain its virulence. Bioinformatics tools have been developed to identify *in silico* prophages, plasmids or pathogenicity islands in newly assembled genomes (Che et al. 2014; Arndt et al. 2016; Carattoli et al. 2014). We describe then the whole-genome sequencing and annotation of seven clinical strains of *S. lugdunensis* and identification of MGE using computational approaches. We identified four new prophages with similarities with CoNS and *S. aureus* phages, but also five plasmids previously identified in other CoNS and in *S. aureus*. We did not identify pathogenicity islands bearing virulence factors, but those data give a totally new insight in CoNS genetic plasticity and argue for the possibility of horizontal genetics transfers with other CoNS and *S. aureus*.

Materials and Methods

Bacterial Strains

Six strains of *Staphylococcus lugdunensis* came from clinical samples issued from VISLISI trial (Virulence of *Staphylococcus lugdunensis* in Severe Infections) (Argemi et al. 2016). This prospective study was conducted from November 2013 to March 2016 at the University Hospital of Strasbourg, France that promoted the study (PRI 2013–HUS n° 5616). The study was carried out in accordance with the French Ethical Committee recommendations and written informed consent from all subjects in accordance with the Declaration of Helsinki (study registration number: IDCRB-2013-A01057-38). This study was also registered at clinicaltrial.gov under number NCT02026895. The six selected strains for whole-genome sequencing were named in accordance to the trial name and the rank of inclusion. Clinical origins of the strains are reported in table 1.

Genome Sequencing and Annotation

Whole-genome sequencing was performed using Illumina technology: Illumina HiSeq 2500 (GATC Biotech AG, Konstanz, Germany). It produced paired end sequences of 125 bp. Adapters were removed and low quality sequences excluded by GATC and final quality of the fastq files controlled with FastQC (v 0.11.4). Then, sequence assemblies were performed using SPAdes (v 2.9.0) with the following kmer: 21-33-55 (Bankevich et al. 2012). SPAdes output contigs $<$ 500 pb or with coverage $<10\times$ were removed. Sequences were finished using PAGIT toolkit from Sanger Institute (Swain et al. 2012). ABACAS (Algorithm-Based Automatic Contiguation of Assembled Sequences) software (v 1.3.1) was used to orientate and order contigs using a reference genome N920143 (NCBI reference sequence NC_017353.1) (Assefa et al. 2009). IMAGE (Iterative Mapping and Assembly for Gap Elimination) software (v 2.4.1) allowed gap closing using raw fastq files and scaffolding in a second step (Tsai et al. 2010). Final sequence annotation was performed using AGMIAL pipeline (Bryson et al. 2006). This workflow produced *in fine* a unique chromosome for each sequence and short

Table 1

Whole-Genome Sequencing of Seven Strains of *Staphylococcus lugdunensis* (Paired-Base Sequencing, Illumina HiSeq 2500) and Comparison with *S. lugdunensis* Reference Strain N920143 (NCBI Reference Sequence: NC_017353.1)

Strain	Ref N920143	VISLISI_21	VISLISI_22	VISLISI_25	VISLISI_27	VISLISI_33	VISLISI_37	C33
Clinical origin	Breast abscess	Bacteremia	Endocarditis	Knee prosthesis	Knee prosthesis	Liver abscess	Endocarditis	Cutaneous swab
Contigs number	69 (V)	34 (S)	27 (S)	19 (S)	26 (S)	26 (S)	24 (S)	40 (S)
N50 (kb)	72	167	352	599	402	245	444	170
Full length of the scaffold (bp)	2.595.888	2.546.158	2.567.776	2.491.642	2.594.484	2.662.421	2.579.522	2.529.316
GC content (%)	33.8	33.7	33.7	33.8	33.6	33.7	33.7	33.8
Genome fraction VS reference (%)	—	94.7	97.4	93.7	97.1	97.3	98.9	94.3
Coding sequences	2359	2451	2483	2422	2427	2524	2415	2429
tRNA	56	47	59	48	60	56	52	58
rRNA	14	6	7	4	7	5	6	7
tmRNA	0	1	1	1	1	1	1	1
Functional annotations								
All proteins	—	2438	2476	2416	2498	2599	2483	2420
GO terms	—	1610 (66%)	1608 (64.5%)	1596 (66.1%)	1632 (65.3%)	1649 (63.4%)	1622 (65.3%)	1602 (66.2%)

NOTE.—Ref, reference; V, velvet and S, Spades; kb, kilobase; bp, base pairs; tRNA, transfer RNA; rRNA, ribosomal RNA; tmRNA, transfer-messenger RNA; GO terms, number of gene ontology terms found by InterProScan software.

nonaligned contigs that were further analysed to search for plasmids, or any other genetic elements of interest. Each nonaligned contig was loaded into ARTEMIS software (v 16.0.0) to identify open reading frames (ORF) that were successively analyzed using protein BLAST (Carver et al. 2008). Genomes functional annotations were performed using InterProScan (v 4.8) (Jones et al. 2014).

Genome Comparison and MGE Search

CSI Phylogeny (v 1.4) from the Center for Genomic Epidemiology (Lyngby, Danemark) was used to produce a phylogenetic tree of both sequenced strains and available annotated genomes. This web based tool calls and filters single nucleotide polymorphisms (SNP), does site validation and infers a phylogeny based on the concatenated alignment of the high quality SNPs (Kaas et al. 2014). The reference strain chosen to compare with was *S. lugdunensis* N920143 (NCBI Reference Sequence: NC_017353.1) and we also included in the tree the three other fully sequenced and annotated genomes of *S. lugdunensis*: HKU09-01 (NCBI Reference Sequence: NC_013893.1), FDAARGOS_141 (NCBI Reference Sequence: NZ_CP014022.1), and FDAARGOS_143 (NCBI Reference Sequence: NZ_CP014023.1). Detailed parameters and command line of the tools used are available in supplementary material part S1, Supplementary Material online.

Prophage search and annotation was performed using PHASTER (Phage Search Tool Enhanced release) (Zhou et al. 2011; Arndt et al. 2016). This web-based tool allows rapid identification of putative prophages sequences and provides annotations. A quality score $>$ 90 defined an intact prophage sequence. PHASTER also provided sequence analysis parameters: region length and position, GC content and the most common related prophages, and phages with available sequences.

Plasmid search was performed on all nonaligned contigs remaining after genome assembly particularly in contigs displaying an unusual high level of coverage possibly linked to sequence duplication in bacterial genome as it might be seen with plasmids. Those additional contigs were annotated using PROKKA (v 1.11) and sequence similarities were searched through BLAST® database. Plasmids categorization was done in accordance to the terminology used by Smillie et al. that distinguish mobilizable from conjugative plasmids depending on the presence of a type IV secretion system (T4SS) (Smillie et al. 2010). Homogeneity of the coverage depth of each contig was controlled by mapping the reads from the raw fastq files against the assembly itself using Bowtie2 (v 2.2.6) and IGV (v 2.3) for alignments visualization (Robinson et al. 2011; Langmead and Salzberg 2012).

Pathogenicity island identification was performed using IslandViewer3 (Dhillon et al. 2015). This web based tool associates IslandPick, SIGI-HMM, and IslandPath-DIMOB for pathogenicity islands identification and displays results in circular graphical images (Langille et al. 2008). Input data are *de novo* annotated genomes issued from AGMIAL pipeline.

Results

Genome Sequences

Genome sequences are detailed in table 1. Illumina paired end sequencing produced 98.8–99.3% high quality reads covering more than 94% of the reference genome with an average depth of 590 $\times$ Total reads range from 10.9 to 14.5 M per sample. Discarded reads remain scarce, representing 1.1–3.2%

of all reads. SPAdes assembly produced seven genomes with 19–34 contigs and N50 ranging from 167 to 444 kb after removing contigs < 500 pb or with coverage <10×. ABACAS contigs ordering and IMAGE gap closing followed by scaffolding produced seven final genomes with sequence length ranging from 2.4 to 2.6 Mbp with GC content comprised between 33.6% and 33.8%. All genome assemblies shared from 93.7 to 98.9% of their sequence with *S. lugdunensis* N920143.

Genome Annotations

Annotations reports are displayed in table 1. All strains contain 2373 to 2524 coding sequences (CDS), with 47–60 tRNA, 5–7 rRNA, and 1 tmRNA for each strain. Those results are closed to annotations reports available for the previously sequenced strains. InterProScan identified 63.4–66.2% ontologies among all identified putative proteins Regarding biological process, 49% of GO terms did not belong to any GO slims terms, 17% of proteins were dedicated to metabolic processes, 5% to biosynthetic processes, 5% to carbohydrate metabolic processes and 24% to various other processes. For cellular component 31% were integral membrane proteins, 31% membrane proteins, 23% cytoplasmic, and 15% were from various locations. Finally, regarding molecular functions, 41% of the proteins did not belong to GO slim terms, 12% displayed catalytic activity, 9% were ATP binding proteins, 6% nucleotide binding, 6% DNA binding, and 26% had various molecular functions. Detailed results are displayed in supplementary material part S2, Supplementary Material online.

Genome Comparisons

All seven genomes were loaded into CSI Phylogeny with the three available genomes and the reference strain as described in the method section. Results are displayed in supplementary material part S3, Supplementary Material online. It shows that the seven strains from VISLISI trial are not isolated in a cluster compared to the strains coming from the literature. VISLISI_21 and FDAARGOS_141 belong to the same cluster, close to VISLISI_25 and a cluster comprising VISLISI_22 and FDAARGOS_143. C33 and VISLISI_37 belong to two separate and distant clusters, as VISLISI_277, VISLISI_33, and HKU09-01 that are closely related.

Prophage Identifications

PHASTER allowed the identification of four putative prophages in assembled genomes: one prophage in VISLISI_22, two in VISLISI_33, and one in VISLISI_37. Prophage annotations and similarities search with other phages/prophages are displayed in table 2. Those prophages were named ϕSL2–ϕSL5 due to the existence of a unique prophage in the literature, ϕSL1 (Heilbronner et al. 2011). All four prophages displayed quite similar lengths from 44.4 to 53.5 Kb and a GC content from 33.8% to 34.5%. Those characteristics and their modular organization is characteristics of phages from the *Siphoviridae* class according to the classification proposed by Kwan et al. (2005). The five functional modules described in this class of phages are observed in those four annotated sequences: lysogeny, DNA metabolism, DNA packaging, and head, tail, and finally a lysis module that is absent from ϕSL4 (Deghorain and Van Melderen 2012). All annotated sequences are flanked by two attachment sites (left and right). The DNA packaging and head module are organized similarly in the four prophages with a small and large subunit of the terminase, then a portal protein followed by a minor head protein, two hypothetical proteins, one scaffold protein, a major head protein and finally a DNA packaging then another minor head protein. The lysis module was absent in ϕSL4 but included in the three other strains amidase and holing proteins. In the tail module, ϕSL2, 3, and 5 displayed similar ORF product with major tail protein, an endopeptidase and a Zn^{2+} carboxypeptidase. ϕSL3 and 5 contain in this module a putative peptidoglycan hydrolase. In the lysogeny module four integrase sequences were identified with the presence of a putative cro-like repressor in ϕSL3. Cro repressors work in temperate bacteriophages in opposition to the phage' repressor that controls the genetic switch and determines whether a lytic or lysogenic cycle will happen after infection (Schubert et al. 2007). Virulence factors were not identified in all four sequences.

Plasmid Identifications

We identified one full plasmid sequence in VISLISI_22 and VISLISI_27, one full plasmid sequence and one partial sequence in VISLISI_33 and two full plasmids sequences in C33. As expected, GC content of the plasmids sequences are lower than *S. lugdunensis*, ranging from 28.7% to 31.8% except pVISLISI3 from VISLISI_33 that display a 33.6% GC content as seen in *S. lugdunensis* (Shintani et al. 2015). The plasmid sizes range from 3,310 to 12,579 bp, what is usually observed in firmicutes plasmids (Shintani et al. 2015). Results are displayed in table 2.

Pathogenicity Island Identifications

All seven annotated genomes were loaded into Island Viewer3 but fail to identify any pathogenicity island in all sequences. Several putative genomic islands were identified with IslandPath-DIMOB that search for genes that are functionally related to mobile elements but IslandPick and SIGI-HMM failed to identify such motifs. IslandPick is a tool that identifies putative genomic islands using comparative methods considering related species and already known genomic islands. SIGI-HMM is a prediction method that uses a Hidden Markov Model and measures codon usage. After careful examination of each genomic region signaled by IslandPath-DIMOB, we confirm the absence of any pathogenicity island in our seven assembled genomes of *S. lugdunensis*.

Table 2

Prophages and Plasmids Identification after Whole-Genome Sequencing of *Stpahylococcus lugdunensis*

	Plasmids					Prophages			
Name	***pVISLISI_1***	***pVISLISI_2***	***pVISLISI_3***	***pVISLISI_4***	***pVISLISI_5***	***ϕSL2***	***ϕSL3***	***ϕSL4***	***ϕSL5***
Host	VISLISI_22	VISLISI_27	VISLISI_33	C33	C33	VISLISI_22	VISLISI_33	VISLISI_33	VISLISI_37
Genome size (kb)	3.1	6.5	4.3	7	12.6	49.4	44.4	53.5	47
GC content (%)	29.9	31.8	33.6	31.11	28.7	34.3	33.8	34.4	34.9
CDS	3	7	5	6	13	66	59	52	65
Virulence factor	None	None	None	None	None	None	None	None	None
Resistance gene	*cadD*	None	None	None	None	None	None	None	None
Related phage	—	—	—	—	—	StB12 (*S. hominis*)	StB12 (*S. hominis*)	PH15 (*S. epidermidis*)	StB12 (*S. hominis*)
Shared proteins	—	—	—	—	—	23	23	13	29
Replication gene	*repL*	*repE*	*repL*	*repE*	*repA*	—	—	—	—
Mobilization module	None	MOBc	MOBc	None	MobA/MobL	—	—	—	—
T4CP/T4SS genes	None	None	None	None	None	—	—	—	—
Plasmid category	Nonmobilizable	Mobilizable	Mobilizable	Nonmobilizable	Mobilizable	—	—	—	—
Related plasmid (strain, nucleotide length, CDS)	pLUG 10 (*S. lugdunensis*, 3117 bp, 3CDS)	SAP108B (*S. epidermidis*, 6206 bp, 7 CDS)	pRIVM6519_1 (*S. aureus*, 4264 bp, 6 CDS)	PM221 p-3 (*S. epidermidis*, 11252 bp, 17)	VRSAp (*S. aureus*, 25107 bp, 30)	—	—	—	—
Nucleotide similarities:									
Sequence cov (%)	98	97	100	41	44	—	—	—	—
Identities (%)	98	97	100	94	91				
E-value	0.0	0.0	0.0	0.0	0.0				

NOTE.—kb, kilo bases; CDS, number of coding sequences; cov, coverage.

Discussion

The whole-genome sequencing of seven strains of *S. lugdunensis* reveal several unexpected characteristics for this virulent CoNS which clinical significance is probably close to *S. aureus*. The core genome of staphylococci is well conserved and nearly 80% of coding sequences from *S. lugdunensis* have reciprocal FASTA matches with *S. aureus* or other CoNS (Heilbronner et al. 2011). Conservation between *S. lugdunensis* strains was expected to be high and this study shows that 94–99% of this genome is conserved in comparison to the actual reference genome available from strain N920143. We identified several mobile genetic elements. Regarding phages, their description in CoNS remain scarce compared to *S. aureus* and if their relationship with *S. aureus* phages was suggested with the characterization of StB12, StB20, and StB27 the possibility of direct genetic transfer *in vivo* from *S. aureus* to CoNS was not proven (Deghorain et al. 2012). Our study identified four new prophages in the seven sequenced strains using computational approaches. Those prophages belong as expected to the *Siphoviridae* family with a GC content close to their host and a usual modules organization (Deghorain and Van Melderen 2012). We did not identify any virulence factor in the CDS but it is of interest to note that the closest prophages sequence from ϕSL4 of VISLISI_33 come from two phages: PH15 from *S. epidermidis* and 187 from *S. aureus*, sharing, respectively, 13 and 12 proteins, respectively, supporting the existence of reciprocal exchange between phages originating from *S. aureus* and *S. lugdunensis* (Kwan et al. 2005; Daniel et al. 2007). This observation is of high importance because even if we did not identify any pathogenicity islands in our genome assemblies this sort of MGE in *S. aureus* are supported by phage helper sequences which provide the genetic machinery for their horizontal transfer (Novick et al. 2010). The *in silico* approach for prophage sequences identification in bacteria is a promising tool and will probably show that phages are not only widespread in *S. aureus*, but also in CoNS, providing them a fundamental tool to evolve and adapt themselves to the environment for example in hospital settings (Xia and Wolz 2014). Plasmids play a major role particularly in bacteria for horizontal genetic transfers such as for antibiotic resistance genes particularly in *Firmicutes* where extrachromosomal replicons are frequently involved in the transmission of antibiotics resistance genes (Shintani et al. 2015). Most of the plasmids described in Staphylococci belong to *S. aureus* (European Nucleotide Archive 2016). *S. epidermidis* has been the most studied CoNS regarding its pathogenicity and up to now, 21 plasmids have been identified. The second CoNS coming after in terms of known plasmids are *S. haemolyticus* and *S. simulans* with five known plasmids. Regarding *S. epidermidis*, the transfer of MGE from *S. epidermidis* to *S. aureus* have been frequently described but in return *S. aureus* did not seem able to transfer genetic material to CoNS (Otto 2009). Nevertheless, the description of a pathogenicity island like regions bearing enterotoxin gene in *S. epidermidis* might suggest the possibility of such events (Madhusoodanan et al. 2011; Méric et al. 2015). Surprisingly, we identified in our study the plasmid pVISLISI_3 (strain VISLISI_33) that has 100% homologies with pRIVM_1, a *S. aureus* plasmid of 4,264 bp, and a second plasmid pVISLISI_5 (strain C33) covering nearly half of the sequence of VRSAp, another *S. aureus* plasmid of 25,107 bp. Those data are limited by the computational approach used to get them, but this emphasizes the need to extend the search of the ability of *S. aureus* to transfer genetic materials to CoNS, an eventuality of the highest importance owing to the capacity of this strain to bear virulence factors on MGE.

Finally, this study allows the identification of several putative mobile genetic elements as prophages and plasmids in a virulent CoNS, *S. lugdunensis*. It did not allow the identification of virulence factors other than already described, but the prophages and plasmids we describe here are important to consider horizontal gene transfer, potentially from *S. aureus* to *S. lugdunensis* as a way for these commensal bacteria to gain its pathogenicity. It reveals the central role than could play NGS for a rapid detection of those MGE. We believe that underrepresentation of those MGE in CoNS compare to *S. aureus* might change rapidly with those fast and reliable methods with a decreasing cost. It cannot replace direct identification of plasmids or prophages, because finding a gene does not mean that this gene will be expressed and how, but this provides a significant gain of time for future research.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Author contributions

X.A., Y.H., P.R., and G.P. designed the experiment, X.A. assembled the genomes, V.M. and V.L. annotated the genomes, N.P. performed ontologies analyses, S.D., J.L., M.M., and A.G. analyzed whole genomic data, D.K. performed DNA extractions, A.V. assisted for data format manipulations.

Acknowledgments

This work was funded by grants from the Collège des Universitaires des Maladies Infectieuses et Tropicales (CMIT), from EA7290 Virulence Bactérienne Précoce, Faculté de Médecine, Strasbourg and supported both by the University Hospital of Strasbourg and Public funding of EA7290.

Literature Cited

Argemi X, et al. 2015. Implementation of matrix-assisted laser desorption ionization-time of flight mass spectrometry in routine clinical laboratories improves identification of coagulase-negative staphylococci and reveals the pathogenic role of *Staphylococcus lugdunensis*. J Clin Microbiol. 53:2030–2036.

Argemi X, et al. 2016. VISLISI trial, a prospective clinical study allowing identification of a new metalloprotease and putative virulence factor from *Staphylococcus lugdunensis*. Clin Microbiol Infect. doi: 10.1016/j.cmi.2016.12.018.
Arndt D, et al. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 44:W16–W21.
Assefa S, Keane TM, Otto TD, Newbold C, Berriman M. 2009. ABACAS: algorithm-based automatic contiguation of assembled sequences. Bioinformatics 25:1968–1969.
Babu E, Oropello J. 2011. *Staphylococcus lugdunensis*: the coagulase-negative staphylococcus you don't want to ignore. Expert Rev Anti Infect Ther. 9:901–907.
Bankevich A, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19:455–477.
Bryson K, et al. 2006. AGMIAL: implementing an annotation strategy for prokaryote genomes as a distributed system. Nucleic Acids Res. 34:3533–3545.
Carattoli A, et al. 2014. *In silico* detection and typing of plasmids using plasmid finder and plasmid multilocus sequence typing. Antimicrob Agents Chemother. 58:3895–3903.
Carver T, et al. 2008. Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. Bioinformatics 24:2672–2676.
Che D, Hasan MS, Chen B. 2014. Identifying pathogenicity islands in bacterial pathogenomics using computational approaches. Pathogens 3:36–56.
Daniel A, Bonnen PE, Fischetti VA. 2007. First complete genome sequence of two *Staphylococcus epidermidis* bacteriophages. J Bacteriol. 189:2086–2100.
Deghorain M, et al. 2012. Characterization of novel phages isolated in coagulase-negative staphylococci reveals evolutionary relationships with *Staphylococcus aureus* phages. J Bacteriol. 194:5829–5839.
Deghorain M, Van Melderen L. 2012. The Staphylococci phages family: an overview. Viruses 4:3316–3335.
Delaunay F, Pegot A, Coquerel-Beghin D, Aktouf A, Auquit-Auckbur I. 2014. *Staphylococcus lugdunensis* necrotizing fasciitis after abdominal dermolipectomy: report of two cases and review of the literature. Ann Chir Plast Esthet. 59:136–139.
Dhillon BK, et al. 2015. IslandViewer 3: more flexible, interactive genomic island discovery, visualization and analysis. Nucleic Acids Res. 43:W104–W108.
Douiri N, et al. 2016. *Staphylococcus lugdunensis*: a virulent pathogen causing bone and joint infections. Clin Microbiol Infect. 22:747–748.
European Nucleotide Archive. 2016. Genomes pages: plasmid [cited 2016 Sep 29] Available from: http://www.ebi.ac.uk/genomes/plasmid.html
Heilbronner S, et al. 2011. Genome sequence of *Staphylococcus lugdunensis* N920143 allows identification of putative colonization and virulence factors: *Staphylococcus lugdunensis* genome sequence. FEMS Microbiol Lett. 322:60–67.
Heilbronner S, Monk IR, Foster TJ. 2013. The phage integrase vector pIPI03 allows RecA-independent, site-specific labelling of *Staphylococcus lugdunensis* strains. Plasmid 70:377–384.
Hung T, Zaghi S, Yousefzadeh J, Leibowitz M. 2012. Necrotizing fasciitis associated with *Staphylococcus lugdunensis*. Case Rep Infect Dis. 2012:453685.
Jones P, et al. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30:1236–1240.
Kaas RS, Leekitcharoenphon P, Aarestrup FM, Lund O. 2014. Solving the problem of comparing whole bacterial genomes across different sequencing platforms. PLoS One 9:e104984.
Kwan T, Liu J, DuBow M, Gros P, Pelletier J. 2005. The complete genomes and proteomes of 27 *Staphylococcus aureus* bacteriophages. Proc Natl Acad Sci U S A. 102:5174–5179.
Langille MG, Hsiao WW, Brinkman FS. 2008. Evaluation of genomic island predictors using a comparative genomics approach. BMC Bioinformatics 9:329.
Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359.
Madhusoodanan J, et al. 2011. An enterotoxin-bearing pathogenicity island in *Staphylococcus epidermidis*. J Bacteriol. 193:1854–1862.
Malachowa N, DeLeo FR. 2010. Mobile genetic elements of *Staphylococcus aureus*. Cell Mol Life Sci. 67:3057–3071.
McCarthy AJ, Lindsay JA. 2012. The distribution of plasmids that carry virulence and resistance genes in *Staphylococcus aureus* is lineage associated. BMC Microbiol. 12:104.
Méric G, et al. 2015. Ecological overlap and horizontal gene transfer in *Staphylococcus aureus* and *Staphylococcus epidermidis*. Genome Biol Evol. 7:1313–1328.
Nanoukon C, et al. 2016. Pathogenic features of clinically significant coagulase-negative staphylococci in hospital and community infections in Benin. Int J Med Microbiol. 307:75–82.
NCBI. Genomes. NCBI [cited 2016 Sep 29]. Available from: https://www.ncbi.nlm.nih.gov/genome/genomes/2548?
Novick RP, Christie GE, Penadés JR. 2010. The phage-related chromosomal islands of Gram-positive bacteria. Nat Rev Microbiol. 8:541–551.
Otto M. 2014. *Staphylococcus aureus* toxins. Curr Opin Microbiol. 17:32–37.
Otto M. 2009. *Staphylococcus epidermidis*: the 'accidental' pathogen. Nat Rev Microbiol. 7:555–567.
Pareja J, Gupta K, Koziel H. 1998. The toxic shock syndrome and *Staphylococcus lugdunensis* bacteremia. Ann Intern Med. 128:603–604.
Poitevin-Later F, Vandenesch F, Dyke K, Fleurette J, Etienne J. 1992. Cadmium-resistance plasmid in *Staphylococcus lugdunensis*. FEMS Microbiol Lett. 78:59–63.
Robinson JT, et al. 2011. Integrative genomics viewer. Nat Biotechnol. 29:24–26.
Sabe MA, Shrestha NK, Gordon S, Menon V. 2014. *Staphylococcus lugdunensis*: a rare but destructive cause of coagulase-negative staphylococcus infective endocarditis. Eur Heart J Acute Cardiovasc Care 3:275–280.
Sato'o Y, Omoe K, Ono HK, Nakane A, Hu D-L. 2013. A novel comprehensive analysis method for *Staphylococcus aureus* pathogenicity islands. Microbiol Immunol. 57:91–99.
Schubert RA, Dodd IB, Egan JB, Shearwin KE. 2007. Cro's role in the CI-Cro bistable switch is critical for λ's transition from lysogeny to lytic development. Genes Dev. 21:2461–2472.
Shintani M, Sanchez ZK, Kimbara K. 2015. Genomics of microbial plasmids: classification and identification based on replication and transfer systems and host taxonomy. Front Microbiol. 6:242.
Smillie C, et al. 2010. Mobility of plasmids. Microbiol Mol Biol Rev. 74:434–452.
Swain MT, et al. 2012. A post-assembly genome-improvement toolkit (PAGIT) to obtain annotated genomes from contigs. Nat Protoc. 7:1260–1284.
Tsai IJ, Otto TD, Berriman M. 2010. Improving draft assemblies by iterative mapping and assembly of short reads to eliminate gaps. Genome Biol. 11:R41.
Woznowski M, et al. 2010. Fulminant *Staphylococcus lugdunensis* septicaemia following a pelvic varicella-zoster virus infection in an immune-deficient patient: a case report. Eur J Med Res. 15:410–414.
Xia G, Wolz C. 2014. Phages of *Staphylococcus aureus* and their impact on host evolution. Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis. 21:593–601.
Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. 2011. PHAST: a fast phage search tool. Nucleic Acids Res. 39:W347–W352.

Associate editor: Howard Ochman