

A peer-reviewed version of this preprint was published in PeerJ on 20 November 2019.

[View the peer-reviewed version](https://doi.org/10.7717/peerj.8068) (peerj.com/articles/8068), which is the preferred citable publication unless you specifically need to cite this preprint.

Cabrera-Contreras R, Santamaría RI, Bustos P, Martínez-Flores I, Meléndez-Herrada E, Morelos-Ramírez R, Barbosa-Amezcu M, González-Covarrubias V, Silva-Herzog E, Soberón X, González V. 2019. Genomic diversity of prevalent *Staphylococcus epidermidis* multidrug-resistant strains isolated from a Children's Hospital in México City in an eight-years survey. PeerJ 7:e8068 <https://doi.org/10.7717/peerj.8068>

Genomic diversity of antibiotic multi-resistant *Staphylococcus epidermidis* isolated from a tertiary care hospital in México City

Roberto Cabrera-Contreras^{Corresp., 1}, Rosa I Santamaría², Patricia Bustos², Irma Martínez-Flores², Enrique Meléndez¹, Rubén Morelos¹, Martín Barbosa-Amezcu³, Vanessa González-Covarrubias³, Xavier Soberón³, Víctor González^{Corresp. 2}

¹ Laboratorio de Patogenicidad Bacteriana, Departamento de Salud Pública, Facultad de Medicina, Universidad Nacional Autónoma de México, Ciudad de México, México

² Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México

³ Instituto Nacional de Medicina Genómica, Mexico City, México

Corresponding Authors: Roberto Cabrera-Contreras, Víctor González
Email address: cabreracontrerasr@yahoo.com, vgonzal@ccg.unam.mx

Staphylococcus epidermidis is a human commensal and opportunistic pathogen worldwide distributed. To ascertain which pathogenic *S. epidermidis* clones are circulating in a local tertiary hospital setting, we sequenced the complete genomes of 17 *S. epidermidis* isolates obtained from neonatal infections at a Hospital Care Unit in México City. Genomic comparisons between *S. epidermidis* isolates revealed high pairwise whole genome nucleotide identities of about 97% to 99% and essentially a clonal structure. We inferred eight Multilocus Sequence Types (MLST's), six of them of worldwide distribution, and two showing allelic variants, not in MLST databases. The profile of virulence includes genes involved in biofilm and modulin formation; most of the strains are multi-resistant to methicillin and several other beta-lactams, fluoroquinolones, and macrolides. Uneven distribution of insertion sequences, phages, and CRISPR-Cas immunity phage systems suggest frequent horizontal gene transfer. Rates of recombination between *S. epidermidis* strains were more frequent than the mutation rate and affected the whole genome. Therefore, recombination properties shape the population structure of local nosocomial *S. epidermidis* strains, formed by pathogenic and probably, non-pathogenic clones.

Genomic diversity of antibiotic multi-resistant *Staphylococcus epidermidis* isolated from a tertiary care hospital in México City.

Roberto Cabrera-Contreras¹, Rosa I. Santamaría², Patricia Bustos², Irma Martínez-Flores², Enrique Meléndez-Herrada¹, Rubén Morelos-Ramírez¹, Martín Barbosa-Amezcu³, Vanessa González³, Xavier Soberón³ and Víctor González².

¹ Laboratorio de Patogenicidad Bacteriana del Departamento de Salud Pública de la Facultad de Medicina, Universidad Nacional Autónoma de México. CP 04510. Ciudad de México. México.

² Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Av. Universidad s/n Col. Chamilpa, Cuernavaca, Morelos CP 62210. México.

³ Instituto Nacional de Medicina Genómica, Periférico Sur No. 4809, Col. Arenal Tepepan, Delegación Tlalpan, Ciudad de México, C.P. 14610, México.

Corresponding authors:

Roberto Cabrera-Contreras

Av. Universidad 3000, Coyoacán, Ciudad de México. CP 04510. México.

E-mail: cabreracntrerasr@yahoo.com

Xavier Soberón

Periférico Sur 4809, Arenal Tepepan, Ciudad de México. 14610. México.

E-mail: xsoberon@inmegen.gob.mx

Víctor González

Av. Universidad N/C, Chamilpa, Morelos. 62219. México

E-mail: vgonzal@cgc.unam.mx

Abstract

Staphylococcus epidermidis is a human commensal and opportunistic pathogen worldwide distributed. To ascertain which pathogenic *S. epidermidis* clones are circulating in a local tertiary hospital setting, we sequenced the complete genomes of 17 *S. epidermidis* isolates obtained from neonatal infections at a Hospital Care Unit in México City. Genomic comparisons between *S. epidermidis* isolates revealed high pairwise whole genome nucleotide identities of about 97% to 99% and essentially a clonal structure. We inferred eight Multilocus Sequence Types (MLST's), six of them of worldwide distribution, and two showing allelic variants, not in MLST databases. The profile of virulence includes genes involved in biofilm and modulin formation; most of the strains are multi-resistant to methicillin and several other beta-lactams, fluoroquinolones, and macrolides. Uneven distribution of insertion sequences, phages, and CRISPR-Cas immunity phage systems suggest frequent horizontal gene transfer. Rates of recombination between *S. epidermidis* strains were more frequent than the mutation rate and affected the whole genome. Therefore, recombination properties shape the population structure of local nosocomial *S. epidermidis* strains, formed by pathogenic and probably, non-pathogenic clones.

Introduction

Staphylococcus epidermidis (*SE*) is a normal commensal bacteria forming part of the human skin microbiome (Byrd et al., 2018). However, some *SE* strains behave as pathogen colonizing surgery wounds, medical devices, and in some circumstances, they can reach the human bloodstream causing severe bacteremia and potential mortality (Otto, 2009; Chessa et al., 2015). Children are prone to acquire methicillin-resistant *SE* strains in contrast to *S. aureus*, but it is especially true in perinatal hospitals, where children are in high risk to contact with *SE* (Villari et al., 2000; Marchant et al., 2013). The genetic basis for pathogenicity of *SE* strains includes genes related to biofilm formation (adhesion) and multiple genes for antibiotic resistance, that are also present in *S. aureus* Elements (MGEs) like phages, insertion sequences (ISs), and pathogenicity islands, that may be associated with the transfer of antibiotic and virulence traits (Miragaia et al., 2009a, 2016; Rolo et al., 2017). (Kozitskaya et al., 2005; Conlan et al., 2012; Meric et al., 2015). There is no clear genetic distinction between pathogenic and commensal non-pathogenic *SE* strains, even though nosocomial strains are enriched in virulence and antibiotic resistance genes (Kozitskaya et al., 2005; Conlan et al., 2012; Meric et al., 2018). It has been proposed that these genes, are within the pool of accessory genome, mobilized within and between species (Meric et al., 2015; Rolo et al., 2017). Genome analysis of strains of *SE* revealed the presence of diverse Mobile Genetic Elements (MGEs) like phages, insertion sequences (ISs), and pathogenicity islands, that may be associated with the transfer of antibiotic and virulence traits (Miragaia et al., 2009; Bouchami et al., 2016; Rolo et al., 2017).

The nosocomial *SE* strains showed high genetic diversity among isolates from distant geographic locations and, even from a single clinical origin (Miragaia et al., 2007). There is evidence by Multilocus Sequence Typing (MLST) that the population structure of *SE* is essentially clonal (Miragaia et al., 2002; Miragaia et al., 2007; Thomas et al., 2007). The clonal complexes ST2 and ST23 are the most frequently found in clinical environments (Miragaia et al., 2007; Lee et al., 2018). Besides the clonal structure, recombination estimated in a few genetic loci occurs two times more than mutation (Miragaia et al., 2007). It has been estimated that 40% of the core genes of *SE* has undergone recombination (Meric et al., 2015). Then, recombination might be a general property of *SE* allowing rapid evolution in clinical settings but conserving high linkage disequilibrium between alleles. In this case, recombination might act as a cohesive force, maintaining the clonal population structure but allowing clones to diverge.

In this work, we addressed the genomic characterization of a collection 17 *SE* isolates from catheters and blood from hospitalized patients at the National Institute of Perinatology (Instituto Nacional de Perinatología “INPer” “Dr. Isidro Espinosa de los Reyes” SSA) in México City. We aimed to determine which *SE* clones are flowing in local hospital environments and to characterize their virulence and antibiotic resistance genetic profile. In this context, we also provide a genome-wide measure of recombination, to shed light on the mechanisms of *SE* genetic diversification of *SE* strains in local sites.

Materials & Methods

Strains and phenotypic analysis. This study was carried out following the recommendations of the ethics review committee of the “Facultad de Medicina de la UNAM”. The *Staphylococcus epidermidis* strains used in this work were obtained by donation from the microbiological collection of the Instituto Nacional de Perinatología “Isidro Espinosa de los Reyes”(INPer). Consent form was not required. Original identification keys and clinical data concerning the isolates are maintained under control of INPer. The INPer strain identifiers were substituted by conventional names in the present study to prevent further interpretations of the data. Authors do not have access in any form to specific clinical information of strains and patients.

In the public health laboratory at the Faculty of Medicine, conventional phenotypic analyses were performed of the 17 *SE* strains. The biotype and antibiotype for 17 antibiotics were determined using the automatic VITEK equipment. The antimicrobials tested were as follows: Amoxicillin/Clavulanic Acid, Ampicillin, Cefazolin, Ciprofloxacin, Clindamycin, Erythromycin, Gentamicin, Imipenem, Levofloxacin, Linezolid, Oxacillin, Penicillin, Rifampicin, Synercid, Tetracycline, Trimethoprim/Sulfamethoxazole, and Vancomycin.

Genomic sequencing and annotation. Genomic DNA was isolated by using Lysostaphin-Lysozyme and isopropanol precipitation (Coll, 2005). Nextera libraries were generated from genomic DNA and sequenced using a paired-end base dual index, performed on a MySeq Illumina unit producing 1-2 millions of reads pairs per library by ~80x genome coverage reads (Andrews, 2010) (Bolger et al., 2014). Assemblies were made with Spades 3.6.0 in genomes sequenced in high coverage (50x to 100x). Genome annotations were obtained from the PATRIC server (Wattam et al., 2018) (<https://www.patricbr.org/>), through the automated bioinformatic method RAST (Rapid Annotation Subsystem Technologies) (Antonopoulos et al., 2017). Gene

annotation of antibiotic resistance and virulence-related genes were obtained from the section of Special Genes of PATRIC. Annotations for virulome were supported with other sources of information such as PATRIC VF, VFDB, and Victors, the information obtained refers to genes identified or those associated with virulence proposed in the literature. For the prediction of antibiotic resistance genes, the PATRIC server was used to search specialty genes and is supported for this search with the external databases CARD (Comprehensive Antibiotic Resistance Database) and NDARO (National Database of Antibiotic Resistant Organism) (Davis et al., 2016).

Genome comparisons and pangenome modeling. The Average Nucleotide Identity by MUMmer (ANIm) and the Genomic Coverage (G_{cov}) values were calculated with the JSspecies program, with MUMmer used as a pairwise comparison tool for pairs of *SE* INPer genomes (Richter, 2016). The pangenome was modeled using the GET_HOMOLOGUES package and its defaults values (Contreras-Moreira and Vinuesa, 2013). Briefly, groups of orthologues proteins were computed using the Ortho-MCL integrated into GET_HOMOLOGUES. Accessory genes (unique genes and genes present in at least in two genomes but not in the all genomes), were obtained from the cloud, shell, and soft-core pangenome components according to GET_HOMOLOGUES (Contreras-Moreira and Vinuesa, 2013). The distribution of accessory genes in *SE* genomes was performed with the heatmap.2 function of the R's ggplot2 package.

Phylogeny. A consensus core genome calculated from the pangenome model of 29 *SE* genomes was obtained with GET_HOMOLOGUES (Contreras-Moreira and Vinuesa, 2013). The resulting 1575 core protein clusters were subject to multiple alignments with MUSCLE (Edgar, 2004a, b), gaps removed with TrimAl v2.1, and were concatenated using homemade Perl scripts (Capella-Gutiérrez and Gabaldón, 2013). Phylogenetic trees were constructed by the maximum likelihood (ML) method based on the substitution matrix of JTT (Jones, Taylor, Thornton), with 1000 bootstrap replicates, using RaxML program (Stamatakis, 2015). To draw and edit the phylogenetic trees we used the iTol program (Letunic and Bork, 2019).

Recombination. Inference of recombination was performed with ClonalFrameML (Didelot and Wilson, 2015). First, we ran GET-HOMOLOGUES to obtain the common protein clusters encoded by the genomes of each set of *SE* strains (Contreras-Moreira and Vinuesa, 2013). Second, they were converted to nucleotide sequences and concatenated using homemade Perl scripts. Multiple alignments were made with MUSCLE and gaps were eliminated with TrimAl (Capella-Gutiérrez and Gabaldón, 2013; Edgar et al., 2004). Third, a RAXML (Stamatakis, 2015) tree was done to obtain the Newick format and the transition/transversion parameter κ for running ClonalFrameML under default parameters (Didelot and Wilson, 2015). Z-score statistics was obtained for all the sets of *SE* genomes and p-values using the web application Z Score Calculator (<https://www.socscistatistics.com/tests/ztest/zscorecalculator.aspx>). BoxPlots were performed with the R ggplot2 system.

Mobile elements identification. Prophages were identified with PHAST (Arndt et al., 2017). Only predictions ranked as “intact prophages” were considered for analysis. IS, CRISPR-Cas elements and spacer sequences were obtained from PATRIC annotations (Antonopoulos et al., 2017). Then, IS were classified into families by BLASTx comparison with the ISfinder (Siguiet et al., 2006). To determine the identity of the DNA sequences of the spacers located in the

CRISPR-Cas elements, they were compared by BLASTn with the NCBI virus database. Only identical matches with a phage sequence in the database were recorded.

GenBank accession numbers. *S. epidermidis* of the INPer collection used in this work were uploaded in GenBank with the following Biosample identifiers: SAMN11086744, SAMN11086745, SAMN11086746, SAMN11086747, SAMN11086748, SAMN11086749, SAMN11086750, SAMN11086751, SAMN11086752, SAMN11086753, SAMN11086754, SAMN11086755, SAMN11086756, SAMN11086757, SAMN11086758, SAMN11086759, SAMN11086760. The accession numbers for the genomes of reference *S. epidermidis* strains are listed in Table S3.

Results

Isolation of *SE* nosocomial strains. A total of 17 clinical *SE* INPer coagulase negative (CoNs) strains, were selected according to its outstanding antibiotic resistant profile from the laboratory bacterial collection of the Public Health Laboratory (Facultad de Medicina, UNAM). These particular strains were isolated in a period since 2006 to 2013 from 14 newborns and 3 adults, with nosocomial infection at the National Institute of Perinatology “INPer” (Instituto Nacional de Perinatología “Dr. Isidro Espinosa de los Reyes”, SSA). The isolates came from three different infection sites: blood (8 strains), catheters (7 strains), cerebrospinal fluid (1 strain), and soft tissue (1 strain) (Table 1).

Broad gene catalog from draft genomes. To understand the genetic basis of pathogenicity and antibiotic resistance of nosocomial *SE* INPer strains, we obtained the genome sequence of the 17 strains. After testing different parameters with the assembler programs Velvet and Spades (Zerbino, 2010; Bankevich et al., 2012), we obtained draft genomes assemblies consisting of about 92 up to 432 contigs with a 60-70x average sequence coverage per genome (Table S1). To assert that the 17 assemblies represent a substantial part of the *SE* genomes, we compared the total genome length, and the number and length of the predicted ORFs, with 11 complete and draft genomes of *SE* downloaded from GenBank (Table S2). There was no differences between the genome length of the *SE* INPer genomes and the complete genomes from the GenBank (unpaired t-test = 2.33; p-value = 0.022); in ORFs number (unpaired t-test = -1.68; p-value = 0.117); or ORFs length (unpaired t-test = 2.60; p-value = 0.014) (Table S2). Therefore, the *SE* INPer genome sequences obtained here provide a broad catalog of genes per genome, useful for comparative genomics.

Genomic, pangenomic and phylogenetic relationships among *SE* INPer isolates. To define the genomic similarity between the 17 clinical *SE* INPer strains, we did whole pairwise nucleotide identity estimates (Average Nucleotide Identity by Mummer, ANIm) (Richter et al., 2016). The *SE* INPer, showed high genomic ANIm values about 99%, covering more than 90% of the genome length (Fig. S1). One exception is the strain S10, which showed ANIm values about 97% respect to the rest of the strains.

Besides the genomic identity between *SE* INPer isolates, we wanted to investigate the extent of their genetic variability by performing a pangenome model. To this end, the *SE* INPer collection was complemented by the inclusion of 12 complete genomes of *SE* strains downloaded from GenBank (Table S3, NCBI complete genomes). The model obtained using the

GET_HOMOLOGUES software package (Contreras-Moreira and Vinuesa, 2013), indicated an open pangenome (Fig. 2). The core genome component for the 29 *SE* strains was predicted to consist of about 1575 gene clusters, whereas the sum of genes unevenly distributed in the 29 *SE* genomes (accessory component) contains 4360 gene clusters.

To know the phylogenetic relationship of the *SE* INPer strains in the context of reference *SE* strains, we did an un-rooted ML phylogenetic tree using the predicted 1575 concatenated core proteins (Fig. 1). There were three clades separated by the largest branches in the tree that comprise most of the *SE* INPer strains, and one or more *SE* strains isolate worldwide (Fig. 1, clades B, C, D). The clade marked as D in the tree consists of two groups, one of which includes only reference *SE* strains, whereas the other had most of the *SE* INPer strains. The strain S10, the most different strain by ANIm in the *SE* INPer collection, was grouped in the clade A with the commensal strains *SE* ATCC12228 (2) and *SE* 14.1.R1 isolated from USA and Denmark.

The distribution of accessory genes in individual *SE* strains coincided with the phylogenetic clades. The *SE* INPer strains grouped in tree phylogenetic clades are also related by their similitude in accessory genes presence/absence profile (Fig. 2). Despite the high identity of the *SE* strains, there is a great individual variation that may account for adaptations to local environments.

Clonal structure. To investigate to which clonal ST complex belongs the *SE* INPer strains, we look for the seven proteins of the *S. epidermidis* MLST scheme in the genomes, and compared them with their respective alleles in the *Staphylococcus epidermidis* MLST database (<https://pubmlst.org/sepidermidis/>; Table 1; see methods) (Feil et al., 2004; Thomas et al., 2007). The analysis showed a total of 8 different STs; seven of them already recorded in the database. The S10 strain had an unassigned ST in the database and only differed from the rest by a unique amino acid substitution in the YqiL protein (370L to C). The ST2, ST5, and ST23 are the most represented in our sample, and they are worldwide distributed (Miragaia et al., 2007; Lee et al., 2018). Others STs (ST35, ST59, ST81, ST89) are less frequent in our *SE* collection and other countries as well. The clonal relationships among STs determined by eBURST, indicate that founder clones are ST2 and ST5, whereas the other four STs (ST 59, 81, 89 and 23), are peripheral clones mostly related each other than to the main founder clones (Fig. S3).

Virulence genes. There are few virulence genes characterized in the *SE* INPer strains, and several known virulence genes are often common between commensal and pathogenic strains (Otto, 2009). The major virulence genes of *SE* strains already reported, are those contained in the cluster *icaADBCR* (Arciola et al., 2005). They encode adhesins that allow for biofilm formation, a property shared with some commensal *SE* (Conlan et al., 2012). The five *ica* genes were specifically found in nine of the *SE* INPer strains analyzed here (53%), except the *icaC* that was absent in the strain S05 (Table 1). Another class of virulence genes encoding soluble phenol proteins (modulins), with cytolytic activity against leukocytes, are present in the genomes of all *SE* INPer strains analyzed (Otto, 2014). On the contrary, the secretion system ESAT-6 consisting of the genes *esaAB*, *essABC*, and *esxAB*, related to killing polymorpho-nuclear leukocytes or PMN, are uniquely present in the S10 strain (Burts et al., 2005; Wang et al., 2016; Dai et al., 2017).

Antibiotic resistance genotype and phenotype. All the *SE* INPer strains examined here were first phenotypically characterized in the laboratory by their resistance to methicillin and other β -lactams antibiotics. The *mecA* encodes for a penicillin-binding protein (PBP) carried in a mobile element known as Staphylococcal Cassette Chromosome *mec* or SCC*mec*. We found the community acquired SCC*mec* type IV cassette in 15 out of 17 *SE* strains analyzed, and five of them contain an additional SSC*mec* type II cassette. Indeed, *SE* strains S12, and S15 harbor the SSC*mec* cassettes II, IV, and V, whereas the strains S10 and S16 both lack the SSC*mec* cassette, including the *mecA* gene. Concurrently, all the 17 *SE* strains have the β -lactamase gene (*blaZ*) and their regulators (*blaR* and *I*), which are probably responsible for the broad resistance spectrum to penicillin, carbapenems, and cephalosporins determined by VITEK system (Table 2).

The genomic analysis revealed that all the *SE* studied here, included genes for resistance to fluoroquinolones, macrolides, sulfonamides, aminoglycosides, tetracycline, and other antibiotics not used as the first choice in clinical therapy (Table 2). The analysis also shows several mechanisms of action for antibiotic resistance, including non-synonymous substitutions in housekeeping genes (e. g. amino acid S84F change in GyrA), the presence of membrane extrusion pumps (*norAB*), and modifying enzymes that inactivate the antibiotic (AAC, APH). Indeed, the genomic antibiotic resistance spectrum of these *SE* strains extends to include several genes present only in some strains. For instance, *fosB* the gene for resistance to fosfomycin was found in 9 out of 17 strains. In several of these instances, the resistance phenotype coincided with the presence of one or more genes, encoding modifying or degrading enzymes and mutant protein targets for some antibiotics (Table 2).

Mobile genetic elements. The high degree of genetic identity and low variation in gene content among the *SE* strains suggested closed clonal complexes with few gene exchange and recombination. To study this concern, we first looked for prophages and CRISPR-Cas related systems in the genomes. Phage footprints were found in 10 out of 17 genomes of the *SE* strains. The most significant prophage hits detected by PHAST program, were for genomic regions spanning about 28 to 95 kb that includes an attachment site, a signature of lysogenic phages (Arndt et al., 2017) (Table S4). In this analysis, prophages were found integrated into the genomes of some *SE* strains, such as CNPH82 found in the *SE* strains S14, S17, and S18 (Daniel et al., 2007). Some other prophages such as StB20 and SpBeta were present in S03 and S16 strains respectively and, the prophages IPLA5 in strain S07 and IPLA7 in S12 and S15 strains (Gutierrez et al., 2012). In the remaining *SE* strains, prophages sequences were not detected.

Along with phage infections, resistance mechanisms to phages have also been acquired by *SE* strains. The quest for CRISPR-Cas immune systems results in 9 out of 17 *SE* INPer strains carrying a CRISPR-Cas Type III system. It is constituted by Cas1 and Cas2 responsible for spacer processing and insertion, the ribonuclease Cas6 and the cascade proteins Csm1 to Csm6 involved in the processing of the target transcript (Table S5). The CRISPR-Cas Type III system has been already reported in *SE* to confer immunity to phages as well as to conjugative plasmids (Marraffini and Sontheimer, 2008; Marraffini, 2015). Although the enzymatic organization of the CRISPR-Cas systems is remarkably conserved in *SE*, there are variations in the array of repeats and spacers in the CRISPR loci. There are three distinct types of identical repeated units of 30 or 36 nucleotides, associated with specific sequence spacers. Both repeated and spacers units are

present in variable numbers (3 to 13, for repeated units; 2 to 12 for spacer units). Three spacers that correspond to the CRISPR loci found in the strains S02, S05, and S24, match precisely with a sequence in the *Staphylococcus* phage PH15 genome for the first two strains and the *Staphylococcus* phage 6ec genome for the last (Daniel et al., 2007; Aswani et al., 2014).

SE strains harbor many ISs, belonging to different families (Table S6). The presence of IS256 has been found within pathogenic *SE*, associated with biofilm formation and virulence (Kozitskaya et al., 2004; Murugesan et al., 2018). Among the strains of our collection, there is a clear relationship between strains having the IS256 and the presence of the *ica* operon, confirming previous observations. Exceptionally, only the strain S21 has the *ica* genes but lack of the IS256.

Recombination. The above results suggested that high frequency of HGT and recombination might promote diversification of local *SE* populations (Fig. 1). To evaluate this proposition, we measured the ratio of recombination to mutation (r/m), using the ClonalFrameML program (Didelot and Wilson, 2015) in 17 *SE* INPer genomes and several sets of *SE* genomes downloaded from GenBank (Fig. 2; Table S3). A median average r/m rate about 6.9 was calculated when the 17 *SE* INPer were tested, suggesting that nucleotide substitutions by recombination are more frequent than random point mutations (Vos and Didelot, 2009). Every COG class shows r/m values equal or higher than the estimate for the complete set of 17 *SE* genomes. Indeed, the r/m values on virulence or antibiotic resistance gene class results similar to the other COGs involved in housekeeping functions.

To determine whether or not the recombination estimates were affected by the sample composition of *SE* strains, we design several control tests, with distinct groups of genomes. First, we discarded the most divergent S10 strain of the *SE* INPer collection and ran the ClonalFrameML test only with the 16 *SE* most related *SE* strains of the collection. As it is shown, the r/m rate for this set reduced to a median of four, and this value is not significantly different respect to the r/m of *SE* of the complete genome of *SE* strains obtained from the GenBank (z -score = 1.4, P -value 0.135) (Fig. 3, boxplot 25). Second, we computed the r/m rate in eight different sets randomly selected among 260 complete and draft genomes of *SE* strains from the GenBank (Fig. 3, 26- 33; Table S3). Some sets (Ctr2 and 8) display the lowest r/m values, whereas the rest control sets have r/m upper than two up to four. The results indicate that the strains sample composition influence the recombination estimates.

The RaxML nucleotide phylogenetic tree used as a reference to estimate recombination looks similar to the core protein phylogeny presented in Fig. 1; *SE* strains within clades maintain cohesive relationships (Fig. 4). However, multiple recombination events were detected in the ancestral nodes leading to the *SE* INPer strains. The most prominent branch (red dot line in Fig. 4) divided the *SE* strains in two large clades, one including the clade D and the other constituted by clade B and C. Within the different divergent lineages, recombination introduces much more nucleotide variants than mutation as presented before. The results indicate that at local hospital settings, *SE* strains may contain enough genomic diversity despite their close relationship with the main clonal ST complexes of worldwide distribution.

Discussion

SE is among the most common bacterial isolates found in the human skin microbiome (Byrd et al., 2018). It is also frequently recovered from bacteremia and sepsis samples in neonatal care clinic units, being the most probable etiological-agent (Otto, 2009; Byrd et al., 2018). In clinical practice, it is difficult to assess if *SE* clones are the causal agents of the disease or they are accidental or opportunistic pathogens (Miragaia et al., 2008; Otto, 2017). Although limited, there is essential knowledge about the genetic basis of virulence, antibiotic resistance, and the prevalence of *SE* clones at a worldwide level and in local hospital environments. In this work, we analyzed the genomes of 17 *SE* strains isolated mostly from neonatal patients, over a period spanning of seven years from the Instituto Nacional de Perinatología “INPer” in México City. Our findings respect the virulence genetic determinants, are concordant with those found in *SE* isolates from hospitals and commensal strains isolated worldwide. The most prominent feature of these isolates is the ability to form biofilms, the presence of modulins and the multidrug resistance displayed (Otto, 2014; Xu et al., 2018). However, eight strains of the *SE* collection analyzed here, harbor neither the *ica* operon nor the IS256, considered pathogenicity markers in *SE* (Kozitskaya et al., 2004; Murugesan et al., 2018). Likely, they are commensal *SE* that invaded the patients in the course of their hospital stay.

In the phylogenetic trees reported, pathogenic *SE* strains are intermingled with commensal *SE* strains and do not form a cluster of pathogenic clones (Miragaia et al., 2005; Meric et al., 2018). Recently, Meric *et al.*, suggest, that pathogenic *SE* subpopulations occur within the commensal *SE* strains that may contain genes and alleles for colonization infection sites (Meric et al., 2018). Genome-wide-association (GWAS) studies showed the enrichment of several genes involved with methicillin resistance (*mecA*), biofilm formation (*ica*), cell toxicity, and inflammation response in pathogenic *SE* isolates (Meric et al., 2018). Therefore, it is likely that pathogenic and commensal *SE* strains concur in the same infection site, but actual clinical methods of isolation do not allow to distinguish one from the other. Despite some commensal *SE* strains lack the capacity to form biofilms, they were still recovered from ill patients (Table 1).

There are various footprints of MGEs like phages, ISs, and the phage immunity CRISPR-Cas systems in the *SE* genomes that likely contribute to the adaptability by the acquisition of virulence and antibiotic resistance factors. As expected, due to its mobile nature, these elements do not follow a uniform distribution in the phylogeny, indicating frequent genetic exchange in the *SE* population. Together HGT, homologous recombination may be a factor for genetic diversification of *SE* at hospital settings. In our work, extensive genome analysis of the rates of recombination versus mutation suggests that recombination affects the whole genome and not a particular class of genes. It has been estimated that recombination would involve 40% of the genome of *SE* epidermidis, whereas in *S. aureus* recombination comprise the 24% portion (Meric et al., 2015). Although recombination depends strongly on the sample of strains used for the analysis, the estimated *r/m* values agreed with other recombination test performed with distinct samples of *SE*, few or whole genome markers and distinct methods (Miragaia et al., 2005; Meric et al., 2015). Therefore, the *SE* population despite its whole low levels of nucleotide variation (ANIm > 97%) shows cohesive clonal behavior but frequent gene exchange and recombination.

Several areas concerning the study of the *SE* populations should be addressed to improve diagnostic methods for identifying the risk of carriage virulent *SE* (Meric et al., 2018). One is to

know, the population structure of commensal and pathogenic *SE* strains in healthy and sick individuals; and, second the identification of genes showing high rates of recombination and HGT and their influence in the emergence of pathogenic *SE* clones.

Conclusions

At local hospital settings concurs pathogenic and commensal *SE* strains but it is hard to discern if they are contaminants, commensal colonizers, or true virulent strains (Widerström, 2016). Indeed, single-colony testing of *SE* isolates hamper to know the extent of multi clonal or multi species infections (Van Eldere et al., 2000; Harris et al., 2016). In the present study, some analyzed *SE* INPer strains came from nosocomial patients, but lack *ica* genes, a classical virulence marker. Likely they are part of the intra hospital non-pathogenic microbiome. The comparative genomic analysis presented here, suggest that HGT and recombination might play a crucial role in the origin of the pathogenic clones, moving and recombining antibiotic resistance and virulence genes in distinct genomic clonal backgrounds, including non-pathogenic.

Acknowledgements

Authors recognized the Instituto Nacional de Perinatología (INPer)—for providing the *S. epidermidis* strains used in this work. Thanks to Gabriela Guerrero, Luis Lozano and José Espíritu for computational help. We express thanks to Olga M. Pérez-Carrascal for her advice on the use of ClonalFrameML, and to Santiago Castillo for discussion on recombination analysis.

References

- Andrews S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
- Antonopoulos, D.A., Assaf, R., Aziz, R.K., Brettin, T., Bun, C., Conrad, N. et al. (2017) PATRIC as a unique resource for studying antimicrobial resistance. Brief Bioinform 10.1093/bib/bbx083.
- Arciola, C.R., Gamberini, S., Campoccia, D., Visai, L., Speziale, P., Baldassarri, L., and Montanaro, L. (2005) A multiplex PCR method for the detection of all five individual genes of *ica* locus in *Staphylococcus epidermidis*. A survey on 400 clinical isolates from prosthesis-associated infections. J Biomed Mater Res A 75: 408-413.
- Arndt, D., Marcu, A., Liang, Y., and Wishart, D.S. (2017) PHAST, PHASTER and PHASTEST: Tools for finding prophage in bacterial genomes. Brief Bioinform 10.1093/bib/bbx121.
- Aswani, V.H., Tremblay, D.M., Moineau, S., and Shukla, S.K. (2014) Complete Genome Sequence of a *Staphylococcus epidermidis* Bacteriophage Isolated from the Anterior Nares of Humans. Genome Announc 2.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S. et al. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19: 455-477.

- 444 Bolger, A.M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina
445 sequence data. *Bioinformatics* 30: 2114-2120.
- 447 Bouchami, O., de Lencastre, H., and Miragaia, M. (2016) Impact of Insertion Sequences and
448 Recombination on the Population Structure of *Staphylococcus haemolyticus*. *PLoS One* 11:
449 e0156653.
- 451 Burts, M.L., Williams, W.A., DeBord, K., and Missiakas, D.M. (2005) EsxA and EsxB are
452 secreted by an ESAT-6-like system that is required for the pathogenesis of *Staphylococcus*
453 *aureus* infections. *Proc Natl Acad Sci U S A* 102: 1169-1174.
- 455 Byrd, A.L., Belkaid, Y., and Segre, J.A. (2018) The human skin microbiome. *Nat Rev Microbiol*
456 16: 143-155.
- 458 Chessa, D., Ganau, G., and Mazzarello, V. (2015) An overview of *Staphylococcus epidermidis*
459 and *Staphylococcus aureus* with a focus on developing countries. *J Infect Dev Ctries* 9: 547-550.
- 461 Coll, P., Coque, T.M., Domínguez, M.A., Vázquez, J., Vila, J. (2005) Métodos Moleculares de
462 tipificación epidemiológica en bacteriología. En *Procedimientos en Microbiología Clínica*
463 Recomendaciones de la Sociedad Española de Enfermedades Infecciosas y Microbiología
464 Clínica. Cercenado, E., y Cantón, R. (eds).
- 466 Conlan, S., Mijares, L.A., Program, N.C.S., Becker, J., Blakesley, R.W., Bouffard, G.G. et al.
467 (2012) *Staphylococcus epidermidis* pan-genome sequence analysis reveals diversity of skin
468 commensal and hospital infection-associated isolates. *Genome Biol* 13: R64.
- 469 Contreras-Moreira, B., and Vinuesa, P. (2013) GET_HOMOLOGUES, a versatile software
470 package for scalable and robust microbial pangenome analysis. *Appl Environ Microbiol* 79:
471 7696-7701. Dai, Y., Wang, Y., Liu, Q., Gao, Q., Lu, H., Meng, H. et al. (2017) A Novel ESAT-6
472 Secretion System-Secreted Protein EsxX of Community-Associated *Staphylococcus aureus*
473 Lineage ST398 Contributes to Immune Evasion and Virulence. *Front Microbiol* 8: 819.
- 474 Daniel, A., Bonnen, P.E., and Fischetti, V.A. (2007) First complete genome sequence of two
475 *Staphylococcus epidermidis* bacteriophages. *J Bacteriol* 189: 2086-2100.
- 476 Davis, J.J., Boisvert, S., Brettin, T., Kenyon, R.W., Mao, C., Olson, R. et al. (2016)
477 Antimicrobial Resistance Prediction in PATRIC and RAST. *Sci Rep* 6: 27930.
- 479 Didelot, X., and Wilson, D.J. (2015) ClonalFrameML: efficient inference of recombination in
480 whole bacterial genomes. *PLoS Comput Biol* 11: e1004041.
- 482 Feil, E.J., Li, B.C., Aanensen, D.M., Hanage, W.P., and Spratt, B.G. (2004) eBURST: inferring
483 patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus
484 sequence typing data. *J Bacteriol* 186: 1518-1530.
- 486

- Gutiérrez, D., Martínez, B., Rodríguez, A., and García, P. (2012) Genomic characterization of two *Staphylococcus epidermidis* bacteriophages with anti-biofilm potential. BMC Genomics 13: 228.
- Harris LG, Murray S, Pascoe B, Bray J, Meric G, Mageiros L, et al. (2016) Correction: Biofilm Morphotypes and Population Structure among *Staphylococcus epidermidis* from Commensal and Clinical Samples. PLoS ONE 11(4): e0154510. <https://doi.org/10.1371/journal.pone.0154510>
- Kozitskaya, S., Cho, S.H., Dietrich, K., Marre, R., Naber, K., and Ziebuhr, W. (2004) The bacterial insertion sequence element IS256 occurs preferentially in nosocomial *Staphylococcus epidermidis* isolates: association with biofilm formation and resistance to aminoglycosides. Infect Immun 72: 1210-1215.
- Kozitskaya, S., Olson, M.E., Fey, P.D., Witte, W., Ohlsen, K., and Ziebuhr, W. (2005) Clonal analysis of *Staphylococcus epidermidis* isolates carrying or lacking biofilm-mediating genes by multilocus sequence typing. J Clin Microbiol 43: 4751-4757.
- Lee, J.Y.H., Monk, I.R., Goncalves da Silva, A., Seemann, T., Chua, K.Y.L., Kearns, A. et al. (2018) Global spread of three multidrug-resistant lineages of *Staphylococcus epidermidis*. Nat Microbiol 3: 1175-1185.
- Letunic, I., and Bork, P. (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 10.1093/nar/gkz239.
- Marchant, E.A., Boyce, G.K., Sadarangani, M., and Lavoie, P.M. (2013) Neonatal sepsis due to coagulase-negative staphylococci. Clin Dev Immunol 2013: 586076.
- Marraffini, L.A. (2015) CRISPR-Cas immunity in prokaryotes. Nature 526: 55-61.
- Marraffini, L.A., and Sontheimer, E.J. (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science 322: 1843-1845.
- Meric, G., Miragaia, M., de Been, M., Yahara, K., Pascoe, B., Mageiros, L. et al. (2015) Ecological Overlap and Horizontal Gene Transfer in *Staphylococcus aureus* and *Staphylococcus epidermidis*. Genome Biol Evol 7: 1313-1328.
- Meric, G., Mageiros, L., Pensar, J., Laabei, M., Yahara, K., Pascoe, B. et al. (2018) Disease-associated genotypes of the commensal skin bacterium *Staphylococcus epidermidis*. Nat Commun 9: 5034.
- Miragaia, M., Couto, I., and de Lencastre, H. (2005) Genetic diversity among methicillin-resistant *Staphylococcus epidermidis* (MRSE). Microb Drug Resist 11: 83-93.
- Miragaia, M., Thomas, J.C., Couto, I., Enright, M.C., and de Lencastre, H. (2007) Inferring a population structure for *Staphylococcus epidermidis* from multilocus sequence typing data. J Bacteriol 189: 2540-2552.

- 533 Miragaia, M., Carrico, J.A., Thomas, J.C., Couto, I., Enright, M.C., and de Lencastre, H. (2008)
- 534 Comparison of molecular typing methods for characterization of *Staphylococcus epidermidis*:
- 535 proposal for clone definition. J Clin Microbiol 46: 118-129.
- 536
- 537 Miragaia, M., Couto, I., Pereira, S.F., Kristinsson, K.G., Westh, H., Jarlov, J.O. et al. (2002)
- 538 Molecular characterization of methicillin-resistant *Staphylococcus epidermidis* clones: evidence
- 539 of geographic dissemination. J Clin Microbiol 40: 430-438.
- 540
- 541 Miragaia, M., de Lencastre, H., Perdreau-Remington, F., Chambers, H.F., Higashi, J., Sullam,
- 542 P.M. et al. (2009) Genetic diversity of arginine catabolic mobile element in *Staphylococcus*
- 543 *epidermidis*. PLoS One 4: e7722.
- 544
- 545 Murugesan, S., Mani, S., Kuppusamy, I., and Krishnan, P. (2018) Role of insertion sequence
- 546 element is256 as a virulence marker and its association with biofilm formation among
- 547 methicillin-resistant *Staphylococcus epidermidis* from hospital and community settings in
- 548 Chennai, South India. Indian J Med Microbiol 36: 124-126.
- 549
- 550 Otto, M. (2009) *Staphylococcus epidermidis*--the 'accidental' pathogen. Nat Rev Microbiol 7:
- 551 555-567.
- 552
- 553 Otto, M. (2014) Phenol-soluble modulins. Int J Med Microbiol 304: 164-169.
- 554
- 555 Otto, M. (2017) *Staphylococcus epidermidis*: a major player in bacterial sepsis? Future Microbiol
- 556 12: 1031-1033.
- 557
- 558 Richter, M., Rossello-Mora, R., Oliver Glockner, F., and Peplies, J. (2016) JSpeciesWS: a web
- 559 server for prokaryotic species circumscription based on pairwise genome comparison.
- 560 Bioinformatics 32: 929-931.
- 561
- 562 Rolo, J., Worning, P., Nielsen, J.B., Bowden, R., Bouchami, O., Damborg, P. et al. (2017)
- 563 Evolutionary Origin of the Staphylococcal Cassette Chromosome mec (SCCmec). Antimicrob
- 564 Agents Chemother 61.
- 565
- 566 Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., and Chandler, M. (2006) ISfinder: the
- 567 reference centre for bacterial insertion sequences. Nucleic Acids Res 34: D32-36.
- 568 Stamatakis, A. (2015) Using RAxML to Infer Phylogenies. Curr Protoc Bioinformatics 51: 6 14
- 569 11-14.
- 570
- 571 Thomas, J.C., Vargas, M.R., Miragaia, M., Peacock, S.J., Archer, G.L., and Enright, M.C. (2007)
- 572 Improved multilocus sequence typing scheme for *Staphylococcus epidermidis*. J Clin Microbiol
- 573 45: 616-619.
- 574
- 575 J. Van Eldere, W. E. Peetermans, M. Struelens, A. Deplano, and H. Bobbaers (2000) Polyclonal
- 576 Staphylococcal Endocarditis Caused by Genetic Variability, CID 31: 24-30.

- Villari, P., Sarnataro, C., and Iacuzio, L. (2000) Molecular epidemiology of *Staphylococcus epidermidis* in a neonatal intensive care unit over a three-year period. J Clin Microbiol 38: 1740-1746.
- Vos, M., and Didelot, X. (2009) A comparison of homologous recombination rates in bacteria and archaea. ISME J 3: 199-208.
- Wang, Y., Hu, M., Liu, Q., Qin, J., Dai, Y., He, L. et al. (2016) Role of the ESAT-6 secretion system in virulence of the emerging community-associated *Staphylococcus aureus* lineage ST398. Sci Rep 6: 25163.
- Wattam, A.R., Brettin, T., Davis, J.J., Gerdes, S., Kenyon, R., Machi, D. et al. (2018) Assembly, Annotation, and Comparative Genomics in PATRIC, the All Bacterial Bioinformatics Resource Center. Methods Mol Biol 1704: 79-101.
- Widerström M. Significance of *Staphylococcus epidermidis* in Health Care-Associated Infections, from Contaminant to Clinically Relevant Pathogen: This Is a Wake-Up Call! J Clin Microbiol. 2016;54(7):1679–1681.
- Xu, Z., Misra, R., Jamrozy, D., Paterson, G.K., Cutler, R.R., Holmes, M.A. et al. (2018) Whole Genome Sequence and Comparative Genomics Analysis of Multi-drug Resistant Environmental *Staphylococcus epidermidis* ST59. G3 (Bethesda) 8: 2225-2230.
- Zerbino, D.R. (2010) Using the Velvet *de novo* assembler for short-read sequencing technologies. Curr Protoc Bioinformatics Chapter 11: Unit 11 15.

Figure 1(on next page)

Phylogenetic relationships of *SE* INPer strains and *SE* selected from GenBank

The ML tree consists of four main clades defined by the longest branches (A to D). They are indicated with color dots: red, clade A; blue, clade B; green, clade C; yellow, clade D. The tree was constructed with 1575 common core proteins using RaxML program as described in methods. The results of bootstrap performed with 1000 replicates are indicated in the branches. Acronyms specify the isolation site of the *SE* strains: MEX, México; USA, United States of America; FIN, Finland; GER, Germany; DEN, Denmark; IRE, Ireland; CHI, China; AUS, Australia. *SE*INPer strains are denoted in red.

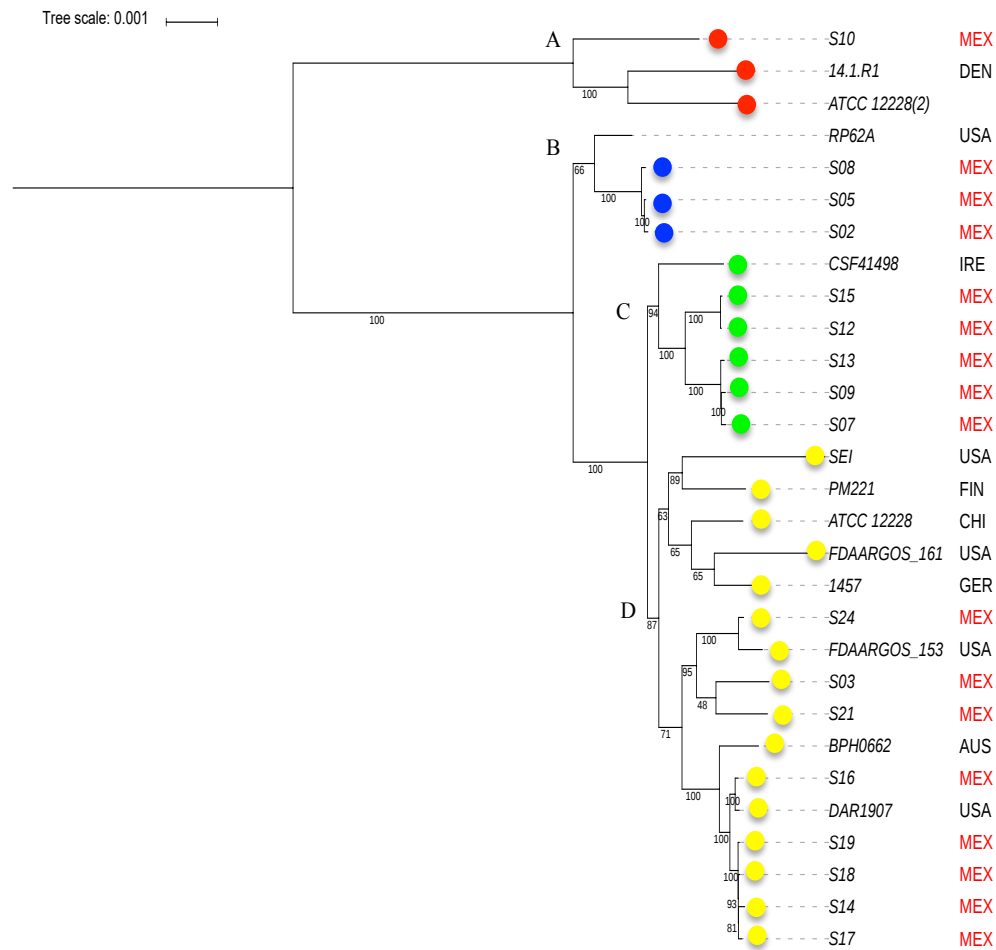


Figure 2 (on next page)

Distribution of accessory genes in the *SE* genomes.

The heat-map profile was performed with the ggplot2 function in R using the Bray-Curtiss dissimilarity matrix. The Bray-Curtiss dendrogram is indicated at the left. The heat-map at the middle indicates gene presence (blue color); empty cells represent the absence of genes. *SE* strains are shown at the left with the same colors of the clades in the core proteins phylogeny (Fig. 1): red, clade A; blue, clade B; green, clade C; yellow, clade D. In the last column, the presence/absence of genomic regions similar to known prophages is indicated (Table S4).

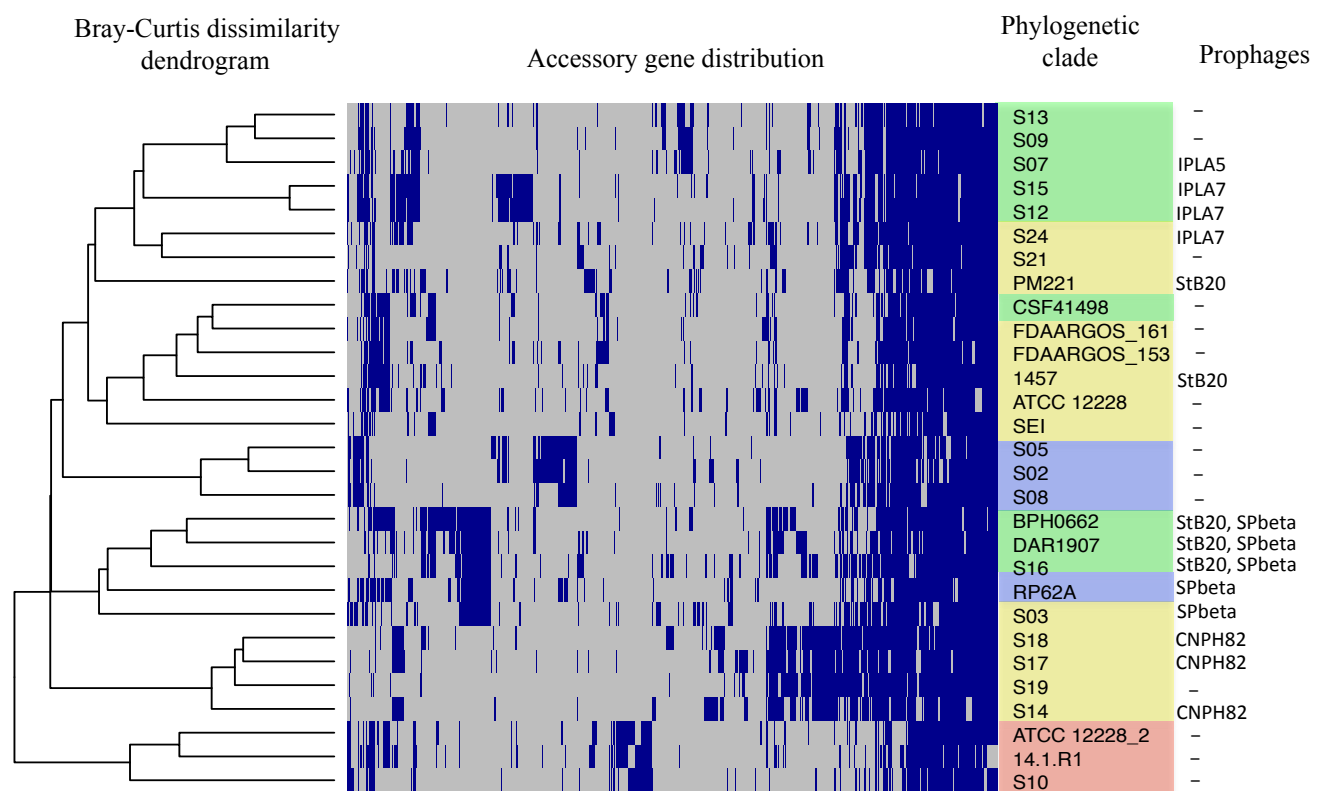


Figure 3(on next page)

Rates of recombination/mutation for *SE* INPer strains compared with sets of *SE* from GenBank.

1. Sixteen strains out 17 *SE* of the INPer collection; 2. Seventeen *SE* strains of the INPer collection. 3- 22, r/m rates for the genes encoding proteins classified in COGs in the 17 *SE* INPer: 3. COG C (energy production). 4. COG D (cell division). 5. COG E (amino acid transport and metabolism); 6. COG F (nucleotide transport and metabolism). 7. COG G (carbohydrate transport and metabolism). 8. COG H (coenzyme transport and metabolism). 9. COG I (lipid transport and metabolism). 10. COG J (translation, ribosomal structure and biogenesis). 11. COG K (transcription). 12. COG L (replication, recombination and repair). 13. COG M (cell wall, membrane, envelope biogenesis). 14. COG O (post-translational modification, protein turnover and chaperones). 15. COG P (inorganic ion transport and metabolism). 16. COG Q (secondary metabolites biosynthesis, transport and catabolism). 17. COG R (general function predicted). 18. COG S (function unknown). 19. COG T (signal transduction mechanisms). 20. COG U (intracellular trafficking, secretion and vesicular transport). 21. COG V (defense mechanism). 22. Unassigned COGs. 23. Antibiotic resistance genes predicted in PATRIC server for the *SE* INPer strains. 24. Virulence genes predicted in PATRIC server for the *SE* INPer strains. 25. Reference set of 12 complete genomes of *SE* strains used through this work. 26-33, subsets of draft *SE* strains from GenBank: 26. Ctr1 (n =36). 27. Ctr2 (n =26). 28. Ctr3 (n =22). 29. Ctr4 (n =35). 30. Ctr-5 (n =36). 31. Ctr6 (n =26). 32. Ctr7 (n =36). 33. Ctr8 (n =36). Descriptions of the *SE* strains included in the control sets and their GenBank accession numbers are in Table S3.

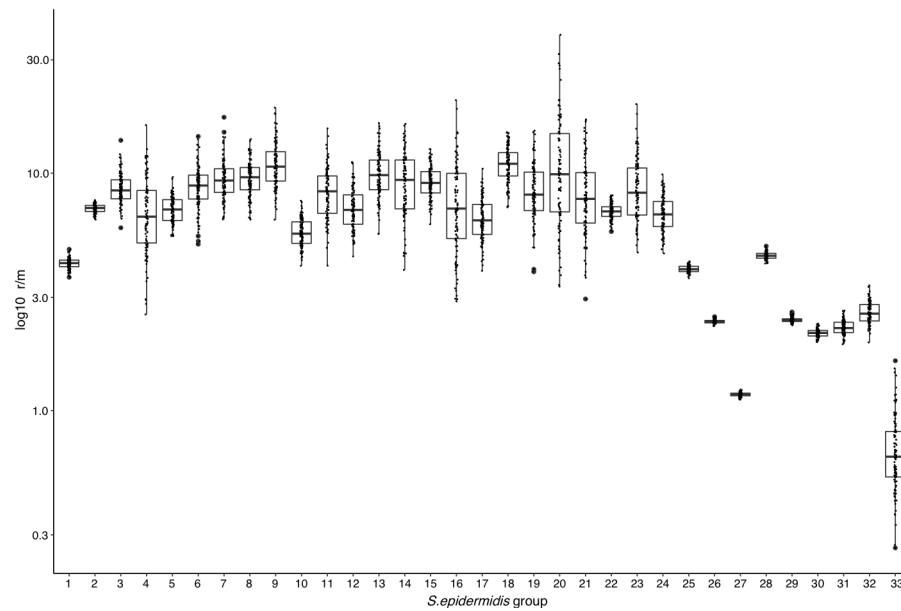


Figure 4 (on next page)

Genome-wide recombination between *SE* strains.

ClonalFrameML program detected several events of recombination along 1,354,455 concatenated genomic regions of 29 *SE* genomes (17 INPer genomes and 12 *SE*GenBank genomes). RaxML nucleotide tree is shown at the left of the scheme. Color dots indicate the corresponding clades in the protein phylogeny (Figure 1). Blue bars indicate recombination events along the concatenated genome segments. White bars indicate non-homoplastic nucleotide substitutions; yellow to red bars are probable homoplastic nucleotide substitutions (Didelot and Wilson, 2015). Red dots line indicates ancestral events of recombination.

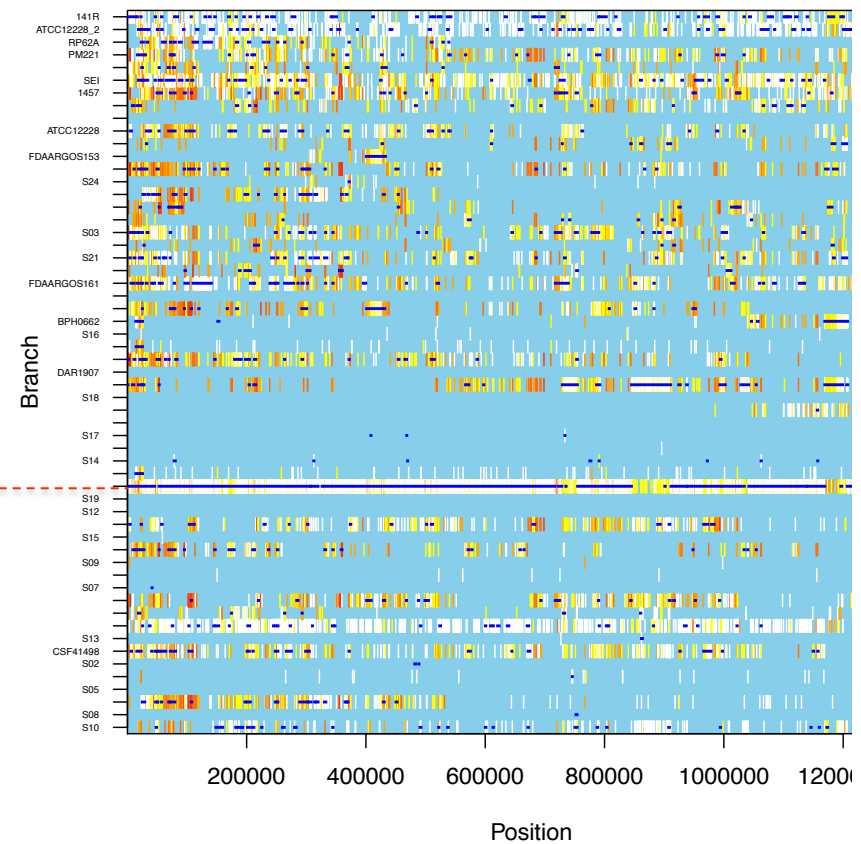
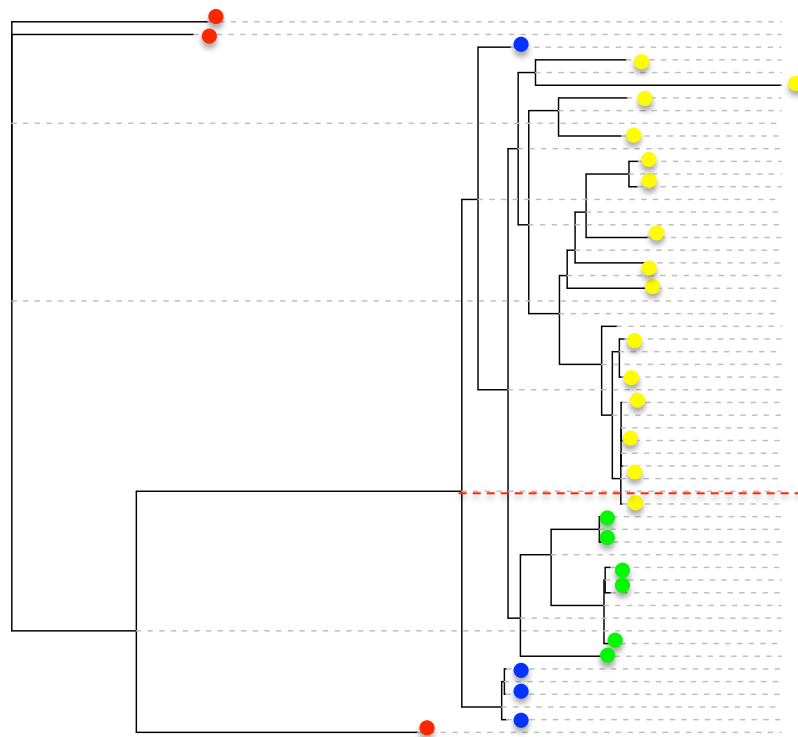


Table 1(on next page)

General features of the genomes of *SE* INPer strains.

^a S05 lack of *icaC*; ^b N prefix indicates strains isolated from neonates; A, isolates from adults; CSF, cerebrospinal fluid; STA, soft tissue aspirate. ^c U-ST, unassigned ST.^d ST profile numbers correspond to alleles of ArcC, AroE, Gtr, MutS, TpiA, and YqiL, according to Thomas et al. (2007). ^e Parentheses indicate the number of complete IS256

1
2

Strain ^a	Isolation year	Origin ^b	ST Type ^c	ST profile ^d	<i>icaADBRC</i>	SCCmec	CRISPR- Cas	Prophages	IS256 ^e
S02	2006	N-Catheter	23	7,1,2,1,3,3,1	+	II, IV	1	–	1 (1)
S03	2006	N-Catheter	89	1,1,2,1,2,1,1	–	IV	0	+	1 (1)
S05	2006	N-Catheter	23	7,1,2,1,3,3,1	+	II, IV	1	–	1 (1)
S07	2007	A-STA	59	2,1,1,1,2,1,1	–	IV	0	+	–
S08	2008	N-Blood	23	7,1,2,1,3,3,1	+	II, IV	1	–	1 (1)
S09	2008	A-Blood culture	59	2,1,1,1,2,1,1	–		1		–
S10	2009	A-Blood	U-ST	65,48,5,5,8,5,11	–	-	0	–	–
S12	2010	N-Blood	5	1,1,1,2,2,1,1	–	II, IV	3	+	1 (1)
S13	2011	N-Blood culture	81	2,17,1,1,2,1,1	–	IV	1	–	1(0)
S14	2011	N-Blood culture	2	7,1,2,2,4,1,1	+	IV	0	+	2 (1)
S15	2011	N-Catheter	5	1,1,1,2,2,1,1	–	II, IV	0	+	1 (1)
S16	2011	N-Blood culture	2	7,1,2,2,4,1,1	+	-	0	+	1 (1)
S17	2012	N-Catheter	2	7,1,2,2,4,1,1	+	IV	0		2 (1)
S18	2012	N-Blood	2	7,1,2,2,4,1,1	+	IV	0	+	2 (1)
S19	2012	N-Catheter	2	7,1,2,2,4,1,1	+	IV	0	–	2 (1)
S21	2012	N-Catheter	35	2,1,2,2,4,1,1	+	IV	1	–	–
S24	2013	N-CSF	5	1,1,1,2,2,1,1	–	IV	3	+	–

Table 2 (on next page)

Antibiotic resistance phenotypes and the genotype profile in *SE* INPer strains

^a Number of strains either resistant or susceptible, and the total tested in the VITEK system (see methods). ^b Parentheses indicate the number of strains that present the most probable gene (s) involved in antibiotic resistance. In *gyrA* the corresponding amino acid substitution in the protein is indicated.

1
2

Antibiotic	Phenotype ^a			Genotype ^b
	Resistant	Susceptible	Tested	
Ampicillin	17	0	17	<i>blaZ</i> (17), <i>mecA</i> (15)
Penicillin	17	0	17	<i>blaZ</i> (17), <i>mecA</i> (15)
Amoxicillin	17	0	17	<i>blaZ</i> (17), <i>mecA</i> (15)
Oxacillin	17	0	17	<i>blaZ</i> (17), <i>mecA</i> (15)
Imipenem	12	0	12	<i>blaZ</i> (17), <i>mecA</i> (15)
Cefalotin	1	0	1	<i>blaZ</i> (17), <i>mecA</i> (15)
Cefazolin	12	0	12	<i>blaZ</i> (17), <i>mecA</i> (15)
Cefoxitin	1	0	1	<i>blaZ</i> (17), <i>mecA</i> (15)
Ceftriaxone	6	0	6	<i>blaZ</i> (17), <i>mecA</i> (15)
Ciprofloxacin	13	4	17	<i>gyrA</i> S84F (10)
Moxifloxacin	3	2	5	<i>gyrA</i> S84F (10)
Levofloxacin	13	3	16	<i>gyrA</i> S84F (10)
Clarithromycin	1	0	1	<i>mphB</i> (3), <i>ermC</i> (11)
Erythromycin	17	0	17	<i>mphB</i> (3), <i>ermC</i> (11)
Clindamycin	16	1	17	<i>linA</i> (1), <i>ermC</i> (11)
Tetracycline	8	9	17	<i>tetL</i> (4)
Gentamycin	17	0	17	<i>aac/aph</i> (15), <i>aph3</i> (2), <i>ant9</i> (3), <i>aadD</i> (11)
Trimethoprim/Sulfamethoxazole	15	2	17	<i>sul3</i> (17)
Chloramphenicol	9	2	11	<i>catB</i> (6)
Vancomycin	1	15	16	<i>vanRI</i> (17)