# Comparative genomics analysis demonstrated a link between staphylococci isolated from different sources: a possible public health risk

**Cave, Rory J., Misra, Raju, Chen, Jiazhen, Wang, Shiyong and Mkrtchyan, Hermine (2021) Comparative genomics analysis demonstrated a link between staphylococci isolated from different sources: a possible public health risk. Frontiers in Microbiology, 12.**

**http://dx.doi.org/10.3389/fmicb.2021.576696**

**This is the Accepted Version of the final output.**

# Comparative genomics analysis demonstrated a link between staphylococci isolated from different sources: A Possible public health risk.

Rory Cave[1] Raju Misra[2]; Jiazhen Chen[3]; Shiyong Wang[3]; Hermine V Mkrtchyan[4,*]

[1]University of East London, School of Health, Sport and Bioscience, Water Lane, London,

[2]Natural History Museum, Core Research Laboratories, Molecular Biology, Cromwell Rd, London

[3]Department of Infectious Disease, Huashan Hospital, Fudan University, 12 Middle Wulumuqi Rd., Shanghai 200040, China

[4]School of Biomedical Sciences, University of West London, St Mary's Rd, London, UK, W5 5RF

*** Correspondence:**
Hermine V Mkrtchyan
hermine.mkrchytan@uwl.ac.uk

**Public settings, WGS, Phylogenetic, Staphylococci, public health**

## Abstract

Coagulase-negative staphylococci (CoNS) have been recovered from different ecological niches, however, little is known about the genetic relatedness of these isolates. In this study, we used whole genome sequencing to compare *mecA* positive (*mecA*$^+$) *Staphylococcus epidermidis*, *Staphylococcus haemolyticus* and *Staphylococcus hominis* isolates recovered from hand-touched surfaces from general public settings in East and West London with data of isolates deposited to European Nucleotide Archive (ENA) by other research groups. These included isolates associated with hospital settings (including those recovered from patients), healthy humans, livestock, pets, plants and natural and other public environments. Using core and accessory phylogenetic analyses we were able to identify that the *mecA*$^+$ *S. epidermidis* and *S. haemolyticus* isolates recovered from general public settings were genetically related to isolates recovered from the bloodstream, urinary tract and eye infections. *S. epidermidis* isolates recovered in our study were also shown to be genetically related to isolates previously recovered from livestock/livestock housing, whereas *S. haemolyticus* isolates were genetically related to isolates recovered from a dog and kefir (fermented cow milk drink). *MecA*$^+$ *S. hominis* isolates were not genetically related to any isolates recovered from clinical samples but were genetically related to isolates recovered from mosquitoes, air samples (residential areas) and kefir. All three species showed to have genetic relatedness to isolates recovered from healthy humans. These results show that CoNS isolates in this study share genetic similarities with those of different lineages and that *mecA*$^+$ *S. epidermidis* and *S. haemolyticus* isolates found in general public settings in this study may pose a risk to public health.

## Introduction

Hand touched surfaces in public areas act as an intermediate for human to human transmission of pathogenic bacteria (Lei et al., 2017). Many bacteria responsible for various infections are known to have originated from humans or animals' sources which later transmitted across the species barrier (Fitzgerald, 2012; Argudín et al., 2015b) However, little is known of the genetic lineages of coagulase-negative staphylococci (CoNS) recovered from hand touched surfaces in public settings and the threat they pose to public health.

CoNS are the most common commensal group of bacteria found on human skin and frequently found on surfaces in hospitals, hotel rooms, libraries, university campus and public transport (Becker et al., 2014; Seng et al., 2017; Xu et al., 2018b; Cave et al., 2019). Unlike coagulase-positive *Staphylococcus aureus* they lack many key virulent factors; however, *Staphylococcus epidermidis*, *Staphylococcus haemolyticus* and *Staphylococcus hominis* have been identified as significant pathogens associated with nosocomial infections and medical devices (Chaves et al., 2005; Cherifi et al., 2013; Czekaj et al., 2015). *S. epidermidis* accounts for 22% of bloodstream infections in intensive care unit patients in the USA, whereas CoNS accounts for 23.1% and 12.7% of bloodstream infections in Israel and China respectively (Otto, 2009; Abu-Saleh et al., 2018; Cui et al., 2019). In addition, CoNS have also been reported to be associated with community and animal/livestock infections (Nanoukon et al., 2017; Schoenfelder et al., 2017). Many CoNS infections are challenging to treat due to them being resistant to multiple antibiotics (Lee et al., 2018; Cui et al., 2019; Pain et al., 2019). One antibiotic resistance gene in particular which makes infections caused by staphylococci challenging to treat is *mecA*. This gene confers resistance to all beta-lactam antibiotics, including carbapenem, cephalosporin, penam, cephamycin and monobactam and is associated with the "superbug" methicillin resistant *Staphylococcus aureus* (MRSA), a bacterium that has caused severe infections in healthcare settings, community and livestock worldwide.

London is the most densely populated city in Europe, with a population of 8.7 million people (2016) and a population density of 5,590 per km$^2$ (2016) (Land Area and Population Density, Ward and Borough - London Datastore). In 2018, it was estimated that there were 60,788 cases of antibiotic-resistant severe infections in England with London having the highest rate of bloodstream infections caused by antibiotic resistant bacteria (42.9 per 100,000 population) (English Surveillance Programme for Antimicrobial Utilisation and Resistance (ESPAUR), 2019). In addition, CoNS can be commonly isolated from environmental sites in London; with one study showing that 94% of the recovered CoNS isolates were antibiotic-resistant and 11% carrying the *mecA* gene (Xu et al., 2018b).Currently, due to the improvement of whole genome sequencing (WGS) technology, it is possible to determine the genetic relationship of less frequently studied bacteria from different settings (Quainoo et al., 2017).

In this study, we genetically compared *mecA* positive (*mecA*$^+$) *S. epidermidis*, S. *haemolyticus* and *S. hominis* isolates recovered from high-frequency touched surfaces of general public settings in the community and public areas in hospitals with complete and draft genomes (obtained from the European Nucleotide Archive (ENA)) of the isolates recovered from different sources, including bloodstream, urinary tract, and eye infections, healthy humans, livestock, pets, plants, fermented milk drink, natural and other public environmental sites. This comparative genomics analysis helped to assess whether *mecA*$^+$ CoNS isolates recovered in our study were genetically similar to isolates recovered from different sources (obtained from the ENA), including those that have previously been reported to cause infections.

## Materials and Methods

### Sample collection

Staphylococcal isolates were recovered between November 2016 to September 2017 from high-frequency hand touched surfaces of inanimate objects from two locations in East London and two locations in West London. These locations included one area considered as public settings (shopping centres and train stations) and another area within hospitals where the general public had easy access, without being a patient (reception area, public washrooms, corridors, lifts). Fifty sites from each location were randomly sampled using COPAN dry swabs (Copan Diagnostics Inc., USA). Six hundred staphylococcal isolates were recovered of which 224 were from East London and 376 from West London. One hundred and eighty-two of the isolates were from the community area and 418 from hospital areas. Ninety-seven of the isolates were from East London community area; 85 from West London Community area; 127 from East London Hospital (ELH) and 376 from West London Hospital (WLH) (Cave et al., 2019).

### Isolation of staphylococci

All samples were directly inoculated onto mannitol salt agar (MSA, Oxoid Basingstoke, UK) within 1-3 hours of recovery and incubated aerobically for 24-72 hours at 37°C. To prevent bias up to 10 colonies from each plate were picked each having different colony morphology or if there are less than 10 different colony morphologies an equal amount of different colony morphologies was selected. These isolates were screened for potential staphylococci characteristics, including performing catalase and coagulase tests. Prolex™ staph latex kit (ProLab Diagnostics, Neston, UK) was used to distinguish *S. aureus* and coagulase-negative staphylococci (Cave et al., 2019).

### Identification of staphylococci recovered from high-frequency hand touch areas

Staphylococcal isolates were initially identified by gram staining and catalase test. The presumptive staphylococcal isolates were further identified at species level using Matrix-assisted laser desorption ionisation-time of flight mass spectroscopy (MALDI-TOF-MS, Microflex LT, Bruker Daltonics, Coventry, UK) in a positive linear mode (2000–20,000 m/z range). Samples were prepared using a full protein extraction method as described previously. (Cave et al., 2019).

### Detection of *the mecA* gene by PCR

Forty three bacterial isolates DNA was extracted via boil lysis and used to perform PCR for detection of the *mecA* gene as previously described (Cave et al., 2019). Using the Met1 and Met2 primers (Eurofins, Germany) PCR reactions were performed in a 20 µl volume (Hanssen et al., 2004), consisting of 10 µl of Phusion Master Mix;1 µl of forward primer, 1 µl of reverse primer, 7 µl of sterile distilled water and 1 µl of isolates DNA template. The PCR condition for this reaction was 94 °C for 5 minutes followed by 35 cycles of denaturation at 94 °C for 30 seconds, annealing at 52 °C for 30 seconds and extension at 72°C for 1 minute with a final extension at 72°C for 10 minutes.

### Genome sequencing and assembly

Whole genome sequencing was performed for 43 *mecA*$^+$ staphylococci isolates using Illumina HiSeq platform. Seven out of 43 isolates were whole genome sequenced by MicrobesNG (Birmingham, UK) and the remaining isolates were sequenced at Fudan University, Shanghai, China (Supplementary Table S1).

Genomic DNA was extracted using TIANamp Bacteria DNA kit (Tiangen, China) and paired-end sequencing libraries were constructed using Nextera XT DNA Sample Preparation kits or TruSeq DNA HT Sample Prep Kit (Illumina, USA) following manufacturer's instruction.

Quality of reads was assessed using FASTQC and trimmed using trimmomatic (Version 0.35), default settings, specifying a Phred cutoff of Q20 (Andrews, 2011; Bolger et al., 2014). Trimmed reads were *de novo* assembly by SPAdes 3.11 and contig assembly was analysed by QUAST and contigs that were ≤ 500bp were removed (Bankevich et al., 2012; Gurevich et al., 2013). The species of these isolates were confirmed by 16s RNA sequencing from the assembled genomes by using the barrnap software (https://github.com/tseemann/barrnap) and searched against a database of known 16S ribosomal RNA sequences using NCBI BLAST tool with a cutoff for species identity of 95% similarity (Altschul et al., 1990; Janda and Abbott, 2007).

## Phylogenetic analyses

A core SNP Maximal likelihood tree was constructed using isolates previously recovered from different sources (data obtained from the ENA, accessed in July 2019; Supplementary Table S2). The main selection criteria of the isolates obtained from ENA database was their isolation sources, which have been recorded in the ENA and/or reported in a peer-reviewed literature. The selection criteria for isolates recovered in this study was the presence of clinically important *mecA* gene. SMALT version 0.5.8 (https://www.sanger.ac.uk/science/tools) was used to map short reads against reference genomes. Reference genomes used to map each staphylococcal species included *S. epidermidis* ATTC 11228; *S. haemolyticus* JCSC 1435 and *S. hominis* K1. SNP calling was done in parallel with all samples of the same species using VarScan version 2.3.9 (Koboldt et al., 2009). VCF file was converted to multi-FASTA alignment file using the python script vcf2phylip (https://github.com/edgardomortiz/vcf2phylip). Recombination was detected and removed from the genome with Gubbins (Croucher et al., 2015). A maximal likelihood tree was constructed using RAxML version 8 using the generalised time reversible model (GTR) model with GAMMA method of correction for site rate variation and 100 bootstrap replications (Stamatakis, 2014). The phylogenetic tree was visualised and annotated using ITOL (Letunic and Bork, 2016).

The distance of the accessory genome for each isolate was determined using the variable-length *k*-mer comparisons to distinguish isolates' divergence in shared sequence and gene content using the POPpunk pipeline (Lees et al., 2019). The number of mixture components was adjusted for each species for obtaining a low-density score (proportion of edges in the network), high transitivity score and high overall score (network score based on density and transitivity). Accessory genome distance was determined by t-SNE with the perplexity (the number of close neighbours each point has) was adjusted for each species to provide the clearest picture of clustering and visualised using Microreact (Argimón et al., 2016).

## Identifying the unique accessory genes that were only present in the ENA *S. hominis* clinical isolates

To determine the unique accessory genes in *S. hominis* hospital associated isolates obtained from the ENA, pangenome analysis was performed using Roary pipeline version 3.4.2. We then compared the differences in genes between all *S. hominis* hospital associated isolates obtained from the ENA with the rest of the *S. hominis* isolates used in phylogenetic analyses (Page et al., 2015).

## Results

### CoNS *mecA* from public settings

The *mecA* gene was identified in 43 isolates that were recovered from high frequency touched surfaces from general public settings from the community and public areas in hospitals (Table 1). This included: *S. epidermidis* (n=17); *S. haemolyticus* (n=10), *S. hominis* (n=10), S. *cohnii (n=3)* and *S. warneri* (n=3).

### Phylogenetic analysis of *S. epidermidis*

Phylogenetic analysis was performed to determine the relatedness of environmental isolates in this study with those recovered from other sources, including isolates recovered from infections (ENA). In this study, 17 *S. epidermidis* *mecA*$^+$ isolates recovered from East and West London were compared to those obtained from the ENA that have previously been recovered from infections (n=34); healthy humans (n=9), livestock (n=13) rodents (n=2), plants (n=4), hospital environment (n=7), animal housing environments (n=2) and natural environments (n=2). Core single nucleotide polymorphism (SNP) phylogenetic tree analysis identified two distinctive clades of which 59 out of 90 were *mecA*$^+$ (Figure 1). Four isolates from East London Hospital (ELH) (321, 327, 329 and 355) belonged to clade A, whereas 1 isolate (407) recovered from ELH was in clade B together with all (n= 12) isolates recovered from West London Hospital (WLH). Interestingly, all ENA isolates recovered from infections, except for VCU128 which was recovered from human airways, were found to be in clade B. ENA isolates recovered from healthy humans, and animals were found within both clades whereas those recovered from plants (all rice seeds) were found in clade A only. Isolate 355 recovered form public settings in our study was genetically related to those ENA isolates that have been recovered from healthy humans (MRSE 52-2 and NIHLM057); isolate 407 was genetically related to those isolates that have been recovered from cow (Y24), pig (PR246B0) and animal housing (M01 and M025), whereas isolates 435, 436, 465, 475, 631, 673, 711, 712, 713, 715 and 716 from WLH were genetically related to isolates previously recovered (ENA) from bloodstream (B45679, 764 SEPI, FDARGOS 153, FDARGOS 83, VCU045, SH06 17, SH03 17, SH03 17) and an endotracheal tube biofilm of a mechanically ventilated patient (ET-0240). 321, 327 and 329 recovered from ELH uniquely showed no relatedness to any other isolate.

PopPUNK analyses revealed that *S. epidermidis* isolates can be combined into 31 groups by their combined core and accessory genome. The accessory genome t -SNE analyses, set at the perplexity of 20, showed that there were 5 distinct clusters (Figure 2). Two of these groups had a mixture of isolates belonging to different combined clusters. None of the clusters includes isolates belonging to a single sequence type. In this study, the *mecA*$^+$ *S. epidermidis* isolates recovered from East and West London were found in different clusters. The accessory genome of *mecA*$^+$ *S. epidermidis* isolates recovered from public settings was related to those ENA isolates that have been previously recovered from bloodstream infections (B45679-10, FDAARGOS 83 and FDAARGOS 161), infected airways (VCU45 and VCU128) and cerebrospinal fluid (CSF41498), endotracheal tube biofilm of a mechanically ventilated patient (ET-024), central venous catheter (1457); from healthy human skin (14.1.R1, mucosa (ATCC 12228) and human airways (MRSE 52-2); from livestock (cows (SNUC 901, SNUC 3608, SNUC 75, SNUC, y24, PM221 and NW32), (a pig (PR246B0 and a sheep (AG42)), a mouse (SCL25); plants (SE2.9, SE4.8, SE4.7 and SE4.6), and from a natural environment (SNUT). In addition, we found that cluster 1 included ENA isolates that were recovered from hospital environments in medical wards and isolates recovered from bloodstream infections, whereas cluster 4 included isolates all recovered from the bloodstream and only a single isolate from healthy human skin (M008).

## Phylogenetic analyses of *S. haemolyticus*

Ten *S. haemolyticus mecA*$^+$ isolates recovered from public settings in East and West London were phylogenetically compared with isolates recovered from other sources, including those recovered from different infections (n=48) and central venous catheter (n=2), commensal isolates recovered from human skin, nares and eyes (n=7); livestock (cow; n=7), kefir (fermented cow milk drink; n=1), companion animals (dog; n=1), public settings (from the surface area of a building and tropical air sample (n=2)), natural environments (n=3), and plant-associated isolates (n=2). The core SNP maximum likelihood phylogenetic tree revealed two distinctive clades (Clade A and Clade B), of which 65 out of 83 isolates were identified as *mecA*$^+$ (Figure 3). Clade A consisted of ENA isolates included those recovered from the bloodstream, vagina and sputum, livestock (cows), a companion animal (dog), groundwater and healthy human eyes. Clade B consisted of ENA isolates recovered from the bloodstream, eye infection, colon infection, central venous catheter, healthy human skin, kefir, willow tree, livestock (cows), tropical air samples, copper alloy coin, surface area of a building and waste and the hygiene compartment of the International Space Station. In addition, all *mecA*$^+$ S. *haemolyticus* isolates recovered from public areas in London in our study were found in clade B, except for the isolate 492, which was recovered from (WLH). We found that isolates we recovered from public areas in ELH and WLH (373, 445 and 538) and isolates recovered from the public settings in East London (1, 93, 99 and 105) were genetically related to ENA isolates that have previously been recovered from an eye infection (SH1572), bloodstream (M-176), and central venous catheter (95671). In addition, one isolate recovered from public areas in WLH (492) in this study were genetically related to an ENA isolate that has previously been recovered from a dog (SW007), whereas one isolate (445) was genetically similar to an isolate that was recovered from kefir (OG2; obtained from the ENA). Moreover, we found four isolates obtained from the ENA, two livestock-associated isolates from cows (BC05211 and NW19), one isolate recovered from kefir (OG2) and one isolate recovered from a willow plant (RIT283) were genetically related to ENA isolates which were recovered from the bloodstream. Interestingly, using the ENA, we were unable to find *S. haemolyticus* isolates that belonged to the same genetic lineage of isolate 506, which was recovered in public areas in the hospital in West London in this study.

PopPUNK analyses revealed that there were 38 combined core and accessory gene clusters predicted, of which the accessory genome was found within 11 clusters (Figure 4). Five of these clusters included isolates that were identified belonging to the same combined cluster and 6 clusters included isolates that were only composed of those ENA isolates that have previously been recovered from clinical samples (blood). Interestingly, isolates recovered from the East London Community and the ELH were found together in the same cluster despite not always possessing the same combined core and accessory cluster (cluster 3). Isolates recovered from West London were found in different clusters (clusters 2, 3 and 6). We observed genetic relatedness of the accessory genomes of all isolates recovered from East London with those ENA isolates that have previously been recovered from an eye infection ( SH1572); venous catheter (95671); and environmental isolates (a copper alloy coin; R1P1), whereas genetic relatedness of the accessory genomes of isolates recovered from West London was observed with those ENA isolates that were recovered from eye infections (SH1572 and SH1574), colon (1HT3), bloodstream ( FDAARGOS-148), vagina (DNF00585), and sputum (C10F); healthy humans (JCSC1435, MTCC 3383, SHN3) and plants (RIT283 and 167); livestock (NW19) and a companion animal (SW007).

## Phylogenetic analysis of *S. hominis*

To determine the relatedness of *S. hominis* isolates in this study with those in the ENA, ten *mecA*$^+$ *S. hominis* isolates recovered from public settings in East and West London in our study were phylogenetically compared with *S. hominis* isolates recovered from different sources that have previously been submitted to the ENA. This included isolates recovered from bloodstreams (n=11), healthy human skin (n=6); livestock (cows; n=11), kefir (n=1) (mosquitos (n=3), natural environments (ancient permafrost and an air sample from residential areas; n=2) and a plant isolate (rice seeds; n=1). SNP core phylogenetic tree of *S. hominis* isolates contained 2 distinct clades: A and B (Figure 2). Clade A consisted of isolates obtained from the ENA and previously recovered from livestock (cows), healthy human skin, the air in residential areas, whereas clade B consisted of ENA isolates previously recovered from clinical bloodstream infections, healthy human skin, kefir and rice seeds, mosquitos and ancient permafrost. Among *mecA*$^+$ *S. hominis* isolates recovered in this study, only the isolate 385 from public areas in hospitals was found in clade A, whereas the remaining of the isolates were found together in clade B. Isolates 387, 386, 620, 623 and 372 recovered from hospitals in East and West London were found in the same subclade and were genetically related to ENA isolate that have been previously recovered from healthy human skin (ZBW5). Isolates 207, 208 and 209 recovered from West London public areas in the community have been found in the same subclade and were genetically related to an isolate recovered from a skin of a healthy human (UMB022), environmental isolates (ancient permafrost in Russia; MMP2), Asian Malaria Mosquito bodies (AS1, AS2 and AS3) and kefir (KR) (Hughes et al., 2016; Kashuba et al., 2017). We did, however, observe that isolates 207, 208 and 209 in this subclade had divergence in their genetic relationship with the ENA isolates. Isolate 385 recovered from East London was genetically related to ENA isolates previously recovered from healthy humans (Hudgins) and air samples from residential areas (H69). Interestingly, isolate 479 recovered from public areas in a WLH was not genetically related to other isolates. All *S. hominis* ENA isolates that were genetically related to isolates from public settings in London except the ENA isolate recovered from mosquito and the permafrost have been shown to harbour the *mecA* gene. All *S. hominis* ENA isolates that were previously recovered from clinical samples and 8 out of 10 ENA isolates recovered from livestock (SNUC 2444, SNUC 5746, SNUC 3403, SNUC 5852, SNUC 4474, SNUC 2620, SNUC 5336 and SNUC 3870) were not genetically related to isolates that we recovered from public settings or other ENA isolates we analysed.

PopPUNK analyses identified 23 combined core and accessory gene clusters, 5 of which were distinct clusters based on accessory genomes (Figure 6). Clinical isolates accessory genomes were found to be clustered together (cluster 5) distinct from those that were recovered from different sources. Additionally, 8 of the 11 *S. hominis* ENA isolates that have been previously recovered from livestock (SNUC 2444, SNUC 5746, SNUC 3403, SNUC 5852, SNUC 4474, SNUC 2620, SNUC 5336 and SNUC 3870) were clustered together (cluster 2). All but one (385) *mecA*$^+$ isolates recovered from public settings in hospitals in this study, were in the same accessory genome cluster (cluster 4). This cluster includes isolate 479 from WLH, which was shown not to be phylogenetically related to other isolates by its core genome. The isolates in cluster 4 in this study were recovered from public settings and were related to those ENA isolates that have previously been recovered from healthy human skin (ZBW5), rice seed (RE2.10) and air samples from residential areas (H69) by their accessory genome.

### Accessory gene analysis of ENA of the *S. hominis* clinical isolates

Core genome phylogenetic analyses showed that *S. hominis* ENA isolates that were previously recovered from the hospital associated infections belonged to a single subclade and formed a single cluster by accessory genome analyses (PopPunk pipeline). Due to these observations, we decided to investigate these isolates further to identify whether they harboured any unique genes ubiquities to them/their accessory genome. We found 22 unique genes that were ubiquities to these *S. hominis* hospital associated infections ENA isolates. Fifteen of these genes were identified as being hypothetical (Supplementary information Table S3). Those that were identified as non-hypothetical genes included the chromosome recombinase gene *ccrA3/B3*; transposition regulatory protein allele *tnpB*; cadmium resistance genes *cadA*, *cadX and cadD* allele; putative DNA repair protein, *radC* and copper-sensing transcriptional repressor *ricR*.

## Discussion

Genomic comparative studies are vital to aid our understanding of the relatedness of pathogenic bacteria recovered from different ecological niches and the transmission of these bacteria between human, livestock and the environment (Hartfield et al., 2014). To date, the majority of phylogenetic studies of CoNS has focused on studying bacteria recovered from clinical settings (Cavanagh et al., 2014a; Post et al., 2017). In addition, there is a limited number of studies reporting on community-associated and livestock-associated CoNS, but little is known about the genetic lineages of CoNS recovered from public settings (Conlan et al., 2012; Chaudhry and Patil, 2016). Using a comparative genome approach, we aimed to determine whether the *mecA*$^+$ CoNS recovered from public settings in East and West London in our study were genetically related to those isolates previously submitted to the ENA by others, including isolates recovered from different sources and those associated with infections (McEwen and Collignon, 2018).

In our study, we found that *S. epidermidis*, *S. haemolyticus* and *S. hominis* were the most common species of CoNS recovered from public settings that were *mecA* positive. This is consistent with reports that previously showed that these species were the most common isolates recovered from nosocomial infections and public settings (Xu et al., 2015; Asaad et al., 2016; Seng et al., 2017). Given that there is a limited number of WGS data available in the European Nucleotide Archive which has the record of their isolation source (https://www.ebi.ac.uk/ena) for *S. epidermidis*, *S. haemolyticus* and *S. hominis* isolated from different sources, we carried out phylogenetic analyses of these isolates only.

Core phylogenetic analysis of all three species revealed that *mecA* isolates recovered from these setting were genetically diverse and span across different clades. For *S. epidermidis the* majority of the isolates recovered from public settings in WLH (711, 712, 713, 715 and 716) were phylogenetically similar to ENA clinical isolates (bloodstream) by their core genome as well as their sequence types (ST2). The discovery of ST2 in public areas hospital areas is not surprising as it is the most common sequence types found in hospital-acquired infections (Deplano et al., 2016). However, it does suggest a route in which well-known hospital-acquired strains can reach the community from a public area. We also found that *mecA*$^+$ isolates 435, 475, 631 recovered from the public areas in WLH were genetically related to clinical isolates recovered from bloodstream; isolates 436 recovered from West London was genetically related to isolate recovered from urine tract infection (FDAARGOS-83), whereas isolate 465 was genetically related to an isolate recovered from endotracheal tube biofilm of a mechanically ventilated patient (ET-024). Additionally, ENA isolates Y24, PR246B0 M01 and M025 previously recovered from livestock and their housing was phylogenetically related to an isolate in this study that was recovered from public areas in hospitals (407).These isolates belonged to the same sequence type (ST59), which has previously been associated with isolates recovered from both livestock and humans (Argudín et al., 2015b; Xu et al., 2018a).These findings indicate that isolates found in public areas in hospitals are genetically related to those that have previously been reported to be associated with infections in humans and livestock. Other studies have shown *mecA*$^+$ *S. epidermidis* as a common cause in bovine mastitis as well as has been recovered from cows milk (Feßler et al., 2010; Fernandes Dos Santos et al., 2016). In addition, pigs have also been shown to be a reservoir of *mecA*$^+$ *S. epidermidis* that had similar virulence and antibiotic resistance gene profiles as isolates recovered from humans, largely indicating the transmission between humans and pigs (Tulinski et al., 2012; Argudín et al., 2015a). These reports, combined with our data, suggest that *S. epidermidis* may represent zoonoses and that livestock-associated *mecA*$^+$ *S. epidermidis* isolates belong to the same genetic lineages as the isolates that have been shown to cause infections in humans. Some known lineages that cause infections in humans may have originated in animals and have been transferred to humans and their associated environments either via direct contact of farmers with animals or via food.

For *S. haemolyticus* core genome phylogenetic analysis all but two isolates recovered from public settings in this study were genetically related to isolates recovered from an eye conjunctivitis (SH1572), bloodstream infection (M-176) and central venous catheter (95671) obtained from the ENA. These results show that the isolates recovered from public settings in this study may potentially pose a public health risk as they belong to the same genetic lineages that have been shown to cause eye conjunctivitis and bloodstream infections (Cavanagh et al., 2014b; Panda and Singh, 2016). This includes the 4 *mecA*$^+$ isolates that were recovered from general public settings in East London (1, 93, 99 and 105), suggesting that isolates that cause bloodstream infections are not only present in public areas in hospitals but can also be found in general public settings. Moreover, we also identified that isolates recovered from public areas in hospitals were genetically related to those isolates that have previously been recovered from a dog (SW007; ENA) and kefir, that as a fermented milk product commonly contains different species of CoNS (OG2; ENA), which have been shown to carry the *mecA* gene (Prado et al., 2015). Previous reports have demonstrated that companion animals are potential reservoirs for the *mecA* gene, which can be transmitted to humans via contact as well as via food products (Ruzauskas et al., 2014). In this study, we were not able to determine whether those *S. haemolyticus* isolates that were genetically related to an isolate recovered from a dog could pose a potential risk to public health as no previous studies have linked the genetic lineages of *S. haemolyticus* isolates recovered from companion animals to that of isolates that have caused infections in humans. To the best of our knowledge, there are no reports that *S. haemolyticus* recovered from livestock belongs to the same genetic lineages known to cause infections in humans. In addition, we found that an isolate in the ENA recovered from kefir (OG2), was also related to the isolate in the ENA that has been recovered from clinical bloodstream infection (M-176).

Core genome phylogenetic analysis of *S. hominis* showed that the ENA livestock (cows) and ENA clinical isolates were genetically different to each other with none being related to the isolates recovered from public settings in East and West London suggesting that that the isolates from livestock and clinical samples have evolved separately and not crossed over into other niches. We did, however, find that our isolates from public setting to be phylogenetically related to isolates recovered from healthy humans skin (Hudgins and ZBW5), air samples from residential areas (H69), mosquitos' bodies (As1, As2 and As3), ancient permafrost (MMP2) and kefir (KR) (Hughes et al., 2016; Rivera-Perez et al., 2016). This suggests that mosquitoes could be possible vectors for transmitting *S. hominis* while feeding on their host and that genetically these *mecA*$^{+}$ *S. hominis* isolates recovered from general public settings have not evolved much since ancient times (Hughes et al., 2016). Mosquitos are vectors for viruses, protozoa and parasites that can spread and cause disease in humans and animals but currently, it is unreported if they can transfer and initiate bacterial infections (Torres-Guerrero et al., 2017; Assaid et al., 2020; Multini et al., 2020). The ENA isolates recovered from mosquitos (As1, As2 and As3) were genetically related to an ENA isolate recovered from healthy human skin (UMB0272) (Figure 5). Therefore, we can conclude that the ENA isolates recovered from mosquitos belong to the same genetic lineages as those recovered from humans. In addition, the findings that the *mecA*$^{+}$ *S. hominis* isolates (385) in our study belonged to the same genetic lineage as the isolate recovered from an air sample in a residential area (H69) suggests that *S. hominis* can be transmitted through the air from humans to high-frequency touched surfaces or vice versa (Lymperopoulou et al., 2017).

Accessory genome is important for bacterial adaption and survival in different environments with studies of *Vibrio vulnificus, Legionella pneumophila* and *Pseudomonas aeruginosa* showing that clinical and environmental isolates can be distinguished by their accessory genomes (Kung et al., 2010; Koton et al., 2015; Mercante et al., 2018, 1). However, studies looking at accessory genomes of CoNS species recovered from different ecological niches are lacking. In the t-SNE plots, we observed that all the clusters generated in *S. epidermidis* analyses had a mixture of isolates from different sources, whereas *S. haemolyticus* and *S. hominis* analysis had clusters that were generated purely on isolates obtained from the ENA, which have been recovered from clinical bloodstream infection samples. In addition, there were two clusters of *S. hominis* ENA isolates that were only recovered from livestock (cows). This suggests that *S. epidermidis* isolates recovered from infections have previously been recovered from other niches. This also coincides with a previously published study which showed that S. *epidermidis* infections are derived from a diverse genetic backgrounds that possess k-mers (infection-associated genetic elements) associated with pathogenicity traits (Méric et al., 2018). These genes have likely originated from bacterial species associated with a particular niche and been transmitted to a new niche via humans or food. We did, however, observed clusters in the accessory genome that only contain *S. haemolyticus* and *S. hominis* ENA clinical associated isolates. This suggests that these isolates might have originated in nosocomial environments but have not spread to other non-hospital niches. This is further supported by the fact that *S. hominis* isolates in this study shared 22 unique genes with those isolates that have previously been recovered from different infections. We found that of these twenty-two genes, in particular those responsible for cadmium and copper resistance were the most interesting. Previous studies have shown the importance of cadmium resistance in *Helicobacter pylori and Listeria monocytogenes* virulence, whereas copper resistance is important for *S. aureus* survival within macrophages (Stähler et al., 2006; Purves et al., 2018). Moreover, it has been shown that metal resistance genes (cadmium, arsenic and zinc) are extensively exchanged between clinical associated *S. epidermidis* and *S. aureus* isolates and may play a role in their survival within hospital settings (Méric et al., 2015). However, it is currently unknown whether these genes contribute to the virulence or survival of *S. hominis or S. haemolyticus* within clinical settings. We found that *S. haemolyticus* isolates recovered from public settings in East

London occupied the same cluster, whereas all but one *S. hominis* isolate from public areas in hospitals were found within the same cluster. This indicates that these species in these areas possess a similar pool of genes that are horizontally transferred due to the similarities in the microbiome in that geographical area or environment and/or adaption required to survive in these niches (Segerman, 2012). In addition, we found that *S. epidermidis* and *S. haemolyticus* recovered from public settings in this study had similar accessory genomes with those isolates that have previously been recovered from different infections (*S. epidermidis:* bloodstream, urinary tract infections and cerebrospinal fluid; *S. haemolyticus:* bloodstream and eye infections) healthy humans, animals (*S. epidermidis* and *S. haemolyticus:* livestock; *S. haemolyticus:* companion animals), plants and wider environments. Both core and accessory genomes phylogenetic analyses suggested that *mecA*$^{+}$ isolates in this study originated from different ecological niches. Therefore, we hypothesis that these isolates contain genes associated with bacterial species across many different genera from different environments. These genes can horizontally be transferred to other CoNS isolates in public settings.

This study implies that CoNS bacteria from high frequency touched surfaces may be a potential risk to public health. Although, they lack virulence determinants, many cases of CoNS infections have been reported inside and outside of hospital settings (Rogers et al., 2009). In addition, CoNS harbour many types of antibiotic resistance genes which can horizontally be transferred across to more virulent pathogens (Xu et al., 2018b; Lee et al., 2019). It also plausible that the CoNS strains we recovered from public settings may have originated from animals either from livestock or companion animals. Animals can be a reservoir of antibiotic-resistant CoNS isolates which can be transmitted to humans furthering the spread of antibiotic-resistance bacteria (Bhargava and Zhang, 2012; Argudín et al., 2015a). The results of this study can be used to review the hand hygiene practices adapted by the general public; the design of the build environments to reduce infections and the use of appropriate disinfectants to prevent further transmission of potentially pathogenic bacteria in these areas.

## Conclusion

*MecA*$^{+}$ *S. epidermidis* and *S. haemolyticus* isolates recovered from public settings in the community and hospitals in this study may pose a potential health risk as we showed that they belong to the same genetic lineages of those isolates that have previously been recovered from different infections (data obtained from the ENA). However, in this study we did not find similar features for the *S. hominis* *mecA*$^{+}$ isolates recovered from public settings. In addition, we showed the relatedness of the *mecA*$^{+}$ staphylococcal isolates (all three species) in this study with isolates recovered from livestock and healthy humans (ENA). Further studies are warranted to aid our understanding of whether the isolates recovered in this study that were genetically related to those recovered from livestock (ENA) could potentially cause infections in humans. The results from this study should be used to instigate a review on how the public wash their hands, the design of public areas (including ventilation practices) and how disinfectants are used to improve the hygiene practices in these areas.

## Acknowledgment

Rory Cave is funded by the UEL PhD studentship awarded to HVM.

## Reference

Abu-Saleh, R., Nitzan, O., Saliba, W., Colodner, R., Keness, Y., Yanovskay, A., et al. (2018). Bloodstream Infections Caused by Contaminants: Epidemiology and Risk Factors: A 10-Year Surveillance. *Isr. Med. Assoc. J.* 20, 433–437.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi:10.1016/S0022-2836(05)80360-2.

Andrews, S. (2011). *FastQC A Quality Control tool for High Throughput Sequence Data*. Cambridge, UK: Babraham Institute Available at: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ [Accessed September 23, 2019].

Argimón, S., Abudahab, K., Goater, R. J. E., Fedosejev, A., Bhai, J., Glasner, C., et al. (2016). Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. *Microb Genom* 2. doi:10.1099/mgen.0.000093.

Argudín, M. A., Vanderhaeghen, W., and Butaye, P. (2015a). Antimicrobial resistance and population structure of *Staphylococcus epidermidis* recovered from pig farms in Belgium. *Vet. J.* 203, 302–308. doi:10.1016/j.tvjl.2015.01.008.

Argudín, M. A., Vanderhaeghen, W., Vandendriessche, S., Vandecandelaere, I., André, F.-X., Denis, O., et al. (2015b). Antimicrobial resistance and population structure of *Staphylococcus epidermidis r*ecovered from animals and humans. *Vet. Microbiol.* 178, 105–113. doi:10.1016/j.vetmic.2015.04.019.

Asaad, A. M., Ansar Qureshi, M., and Mujeeb Hasan, S. (2016). Clinical significance of coagulase-negative staphylococci isolates from nosocomial bloodstream infections. *Infect Dis (Lond)* 48, 356–360. doi:10.3109/23744235.2015.1122833.

Assaid, N., Mousson, L., Moutailler, S., Arich, S., Akarid, K., Monier, M., et al. (2020). Evidence of circulation of West Nile virus in Culex pipiens mosquitoes and horses in Morocco. *Acta Trop.*, 105414. doi:10.1016/j.actatropica.2020.105414.

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. doi:10.1089/cmb.2012.0021.

Becker, K., Heilmann, C., and Peters, G. (2014). Coagulase-Negative Staphylococci. *Clinical Microbiology Reviews* 27, 870–926. doi:10.1128/CMR.00109-13.

Bhargava, K., and Zhang, Y. (2012). Multidrug-resistant coagulase-negative Staphylococci in food animals. *Journal of Applied Microbiology* 113, 1027–1036. doi:10.1111/j.1365-2672.2012.05410.x.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi:10.1093/bioinformatics/btu170.

Cavanagh, J. P., Hjerde, E., Holden, M. T. G., Kahlke, T., Klingenberg, C., Flægstad, T., et al. (2014a). Whole-genome sequencing reveals clonal expansion of multiresistant *Staphylococcus haemolyticus* in European hospitals. *J Antimicrob Chemother* 69, 2920–2927. doi:10.1093/jac/dku271.

Cavanagh, J. P., Hjerde, E., Holden, M. T., Kahlke, T., Klingenberg, C., Flægstad, T., et al. (2014b). Whole-genome sequencing reveals clonal expansion of multiresistant *Staphylococcus* *haemolyticus* in European hospitals. *J Antimicrob Chemother* 69, 2920–2927. doi:10.1093/jac/dku271.

Cave, R., Misra, R., Chen, J., Wang, S., and Mkrtchyan, H. V. (2019). Whole genome sequencing revealed new molecular characteristics in multidrug resistant staphylococci recovered from high frequency touched surfaces in London. *Sci Rep* 9, 1–13. doi:10.1038/s41598-019-45886-6.

Chaudhry, V., and Patil, P. B. (2016). Genomic investigation reveals evolution and lifestyle adaptation of endophytic *Staphylococcus epidermidis*. *Scientific Reports* 6, 19263. doi:10.1038/srep19263.

Chaves, F., Garcia-Alvarez, M., Sanz, F., Alba, C., and Otero, J. R. (2005). Nosocomial Spread of a *Staphylococcus hominis* subsp. novobiosepticus Strain Causing Sepsis in a Neonatal Intensive Care Unit. *Journal of Clinical Microbiology* 43, 4877–4879. doi:10.1128/JCM.43.9.4877-4879.2005.

Cherifi, S., Byl, B., Deplano, A., Nonhoff, C., Denis, O., and Hallin, M. (2013). Comparative Epidemiology of *Staphylococcus epidermidis* Isolates from Patients with Catheter-Related Bacteremia and from Healthy Volunteers. *Journal of Clinical Microbiology* 51, 1541–1547. doi:10.1128/JCM.03378-12.

Conlan, S., Mijares, L. A., Becker, J., Blakesley, R. W., Bouffard, G. G., Brooks, S., et al. (2012). *Staphylococcus epidermidis* pan-genome sequence analysis reveals diversity of skin commensal and hospital infection-associated isolates. *Genome Biol* 13, R64. doi:10.1186/gb-2012-13-7-r64.

Croucher, N. J., Page, A. J., Connor, T. R., Delaney, A. J., Keane, J. A., Bentley, S. D., et al. (2015). Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* 43, e15. doi:10.1093/nar/gku1196.

Cui, J., Liang, Z., Mo, Z., and Zhang, J. (2019). The species distribution, antimicrobial resistance and risk factors for poor outcome of coagulase-negative staphylococci bacteraemia in China. *Antimicrobial Resistance & Infection Control* 8, 65. doi:10.1186/s13756-019-0523-5.

Czekaj, T., Ciszewski, M., and Szewczyk, E. M. (2015). Staphylococcus haemolyticus – an emerging threat in the twilight of the antibiotics age. *Microbiology* 161, 2061–2068. doi:10.1099/mic.0.000178.

Deplano, A., Vandendriessche, S., Nonhoff, C., Dodémont, M., Roisin, S., and Denis, O. (2016). National surveillance of *Staphylococcus epidermidis* recovered from bloodstream infections in Belgian hospitals. *J. Antimicrob. Chemother.* 71, 1815–1819. doi:10.1093/jac/dkw086.

English Surveillance Programme for Antimicrobial Utilisation and Resistance (ESPAUR) (2019). Public Health England Wellington House 133-155 Waterloo Road London SE1 8UG: Public Health England Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/843129/English_Surveillance_Programme_for_Antimicrobial_Utilisation_and_Resistance_2019.pdf.

Fernandes Dos Santos, F., Mendonça, L. C., Reis, D. R. de L., Guimarães, A. de S., Lange, C. C., Ribeiro, J. B., et al. (2016). Presence of mecA-positive multidrug-resistant *Staphylococcus epidermidis* in bovine milk samples in Brazil. *J. Dairy Sci.* 99, 1374–1382. doi:10.3168/jds.2015-9931.

Feßler, A. T., Billerbeck, C., Kadlec, K., and Schwarz, S. (2010). Identification and characterization of methicillin-resistant coagulase-negative staphylococci from bovine mastitis. *J Antimicrob Chemother* 65, 1576–1582. doi:10.1093/jac/dkq172.

Fitzgerald, J. R. (2012). Livestock-associated Staphylococcus aureus: origin, evolution and public health threat. *Trends in Microbiology* 20, 192–198. doi:10.1016/j.tim.2012.01.006.

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi:10.1093/bioinformatics/btt086.

Hanssen, A.-M., Kjeldsen, G., and Sollid, J. U. E. (2004). Local variants of Staphylococcal cassette chromosome mec in sporadic methicillin-resistant *Staphylococcus aureus* and methicillin-resistant coagulase-negative Staphylococci: evidence of horizontal gene transfer? *Antimicrob. Agents Chemother.* 48, 285–296. doi:10.1128/aac.48.1.285-296.2004.

Hartfield, M., Murall, C. L., and Alizon, S. (2014). Clinical applications of pathogen phylogenies. *Trends in Molecular Medicine* 20, 394–404. doi:10.1016/j.molmed.2014.04.002.

Hughes, G. L., Raygoza, J. G., Koundal, V., Rasgon, J. L., and Mwangi, M. M. (2016). Genome Sequences of *Staphylococcus hominis* Strains ShAs1, ShAs2, and ShAs3, Isolated from the Asian Malaria Mosquito Anopheles stephensi. *Genome Announc* 4. doi:10.1128/genomeA.00085-16.

Janda, J. M., and Abbott, S. L. (2007). 16S rRNA Gene Sequencing for Bacterial Identification in the Diagnostic Laboratory: Pluses, Perils, and Pitfalls. *Journal of Clinical Microbiology* 45, 2761–2764. doi:10.1128/JCM.01228-07.

Kashuba, E., Dmitriev, A. A., Kamal, S. M., Melefors, O., Griva, G., Römling, U., et al. (2017). Ancient permafrost staphylococci carry antibiotic resistance genes. *Microb Ecol Health Dis* 28. doi:10.1080/16512235.2017.1345574.

Koboldt, D. C., Chen, K., Wylie, T., Larson, D. E., McLellan, M. D., Mardis, E. R., et al. (2009). VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics* 25, 2283–2285. doi:10.1093/bioinformatics/btp373.

Koton, Y., Gordon, M., Chalifa-Caspi, V., and Bisharat, N. (2015). Comparative genomic analysis of clinical and environmental *Vibrio vulnificus* isolates revealed biotype 3 evolutionary relationships. *Front. Microbiol.* 5. doi:10.3389/fmicb.2014.00803.

Kung, V. L., Ozer, E. A., and Hauser, A. R. (2010). The Accessory Genome of *Pseudomonas aeruginosa*. *Microbiol. Mol. Biol. Rev.* 74, 621–641. doi:10.1128/MMBR.00027-10.

Land Area and Population Density, Ward and Borough - London Datastore Available at: https://data.london.gov.uk/dataset/land-area-and-population-density-ward-and-borough [Accessed May 11, 2020].

Lee, J. Y. H., Monk, I. R., Gonçalves 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. doi:10.1038/s41564-018-0230-7.

Lee, J.-H., Heo, S., Jeong, M., and Jeong, D.-W. (2019). Transfer of a mobile *Staphylococcus saprophyticus* plasmid isolated from fermented seafood that confers tetracycline resistance. *PLOS ONE* 14, e0213289. doi:10.1371/journal.pone.0213289.

Lees, J. A., Harris, S. R., Tonkin-Hill, G., Gladstone, R. A., Lo, S. W., Weiser, J. N., et al. (2019). Fast and flexible bacterial genomic epidemiology with PopPUNK. *Genome Res.* 29, 304–316. doi:10.1101/gr.241455.118.

Lei, H., Li, Y., Xiao, S., Yang, X., Lin, C., Norris, S. L., et al. (2017). Logistic growth of a surface contamination network and its role in disease spread. *Scientific Reports* 7, 1–10. doi:10.1038/s41598-017-13840-z.

Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44, W242-245. doi:10.1093/nar/gkw290.

Lymperopoulou, D. S., Coil, D. A., Schichnes, D., Lindow, S. E., Jospin, G., Eisen, J. A., et al. (2017). Draft genome sequences of eight bacteria isolated from the indoor environment: *Staphylococcus capitis* strain H36, *S. capitis* strain H65, *S. cohnii* strain H62, *S. hominis* strain H69, *Microbacterium* sp. strain H83, *Mycobacterium iranicum* strain H39, *Plantibacter* sp. strain H53, and *Pseudomonas oryzihabitans* strain H72. *Stand Genomic Sci* 12. doi:10.1186/s40793-017-0223-9.

McEwen, S. A., and Collignon, P. J. (2018). Antimicrobial Resistance: a One Health Perspective. *Microbiol Spectr* 6. doi:10.1128/microbiolspec.ARBA-0009-2017.

Mercante, J. W., Caravas, J. A., Ishaq, M. K., Kozak-Muiznieks, N. A., Raphael, B. H., and Winchell, J. M. (2018). Genomic heterogeneity differentiates clinical and environmental subgroups of *Legionella pneumophila* sequence type 1. *PLOS ONE* 13, e0206110. doi:10.1371/journal.pone.0206110.

Méric, 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. doi:10.1038/s41467-018-07368-7.

Méric, G., Miragaia, M., de Been, M., Yahara, K., Pascoe, B., Mageiros, L., et al. (2015). Ecological Overlap and Horizontal Gene Transfer in S*taphylococcus aureus* and *Staphylococcus epidermidis*. *Genome Biol Evol* 7, 1313–1328. doi:10.1093/gbe/evv066.

Multini, L. C., Wilke, A. B. B., and Marrelli, M. T. (2020). Neotropical Anopheles (Kerteszia) mosquitoes associated with bromeliad-malaria transmission in a changing world. *Acta Trop.*, 105413. doi:10.1016/j.actatropica.2020.105413.

Nanoukon, C., Argemi, X., Sogbo, F., Orekan, J., Keller, D., Affolabi, D., et al. (2017). Pathogenic features of clinically significant coagulase-negative staphylococci in hospital and community infections in Benin. *International Journal of Medical Microbiology* 307, 75–82. doi:10.1016/j.ijmm.2016.11.001.

Otto, M. (2009). *Staphylococcus epidermidis* – the “accidental” pathogen. *Nat Rev Microbiol* 7, 555–567. doi:10.1038/nrmicro2182.

Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T. G., et al. (2015). Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31, 3691–3693. doi:10.1093/bioinformatics/btv421.

Pain, M., Hjerde, E., Klingenberg, C., and Cavanagh, J. P. (2019). Comparative Genomic Analysis of *Staphylococcus haemolyticus* Reveals Key to Hospital Adaptation and Pathogenicity. *Front. Microbiol.* 10. doi:10.3389/fmicb.2019.02096.

Panda, S., and Singh, D. V. (2016). Whole-Genome Sequences of Staphylococcus haemolyticus Isolated from Infected Eyes and Healthy Conjunctiva in Bhubaneswar, India. *Genome Announc* 4. doi:10.1128/genomeA.00099-16.

Post, V., Harris, L. G., Morgenstern, M., Mageiros, L., Hitchings, M. D., Méric, G., et al. (2017). Comparative Genomics Study of Staphylococcus epidermidis Isolates from Orthopedic-Device-Related Infections Correlated with Patient Outcome. *Journal of Clinical Microbiology* 55, 3089–3103. doi:10.1128/JCM.00881-17.

Prado, M. R., Blandón, L. M., Vandenberghe, L. P. S., Rodrigues, C., Castro, G. R., Thomaz-Soccol, V., et al. (2015). Milk kefir: composition, microbial cultures, biological activities, and related products. *Front. Microbiol.* 6. doi:10.3389/fmicb.2015.01177.

Purves, J., Thomas, J., Riboldi, G. P., Zapotoczna, M., Tarrant, E., Andrew, P. W., et al. (2018). A horizontally gene transferred copper resistance locus confers hyper-resistance to antibacterial copper toxicity and enables survival of community acquired methicillin resistant *Staphylococcus aureus* USA300 in macrophages. *Environ Microbiol* 20, 1576–1589. doi:10.1111/1462-2920.14088.

Quainoo, S., Coolen, J. P. M., Hijum, S. A. F. T. van, Huynen, M. A., Melchers, W. J. G., Schaik, W. van, et al. (2017). Whole-Genome Sequencing of Bacterial Pathogens: the Future of Nosocomial Outbreak Analysis. *Clinical Microbiology Reviews* 30, 1015–1063. doi:10.1128/CMR.00016-17.

RIVERA-PEREZ, J. I., SANTIAGO-RODRIGUEZ, T. M., and TORANZOS, G. A. (2016). Paleomicrobiology: a Snapshot of Ancient Microbes and Approaches to Forensic Microbiology. *Microbiol Spectr* 4. doi:10.1128/microbiolspec.EMF-0006-2015.

Rogers, K. L., Fey, P. D., and Rupp, M. E. (2009). Coagulase-Negative Staphylococcal Infections. *Infectious Disease Clinics of North America* 23, 73–98. doi:10.1016/j.idc.2008.10.001.

Ruzauskas, M., Siugzdiniene, R., Klimiene, I., Virgailis, M., Mockeliunas, R., Vaskeviciute, L., et al. (2014). Prevalence of methicillin-resistant *Staphylococcus haemolyticus* in companion animals: a cross-sectional study. *Ann. Clin. Microbiol. Antimicrob.* 13, 56. doi:10.1186/s12941-014-0056-y.

Schoenfelder, S. M. K., Dong, Y., Feßler, A. T., Schwarz, S., Schoen, C., Köck, R., et al. (2017). Antibiotic resistance profiles of coagulase-negative staphylococci in livestock environments. *Vet. Microbiol.* 200, 79–87. doi:10.1016/j.vetmic.2016.04.019.

Segerman, B. (2012). The genetic integrity of bacterial species: the core genome and the accessory genome, two different stories. *Front Cell Infect Microbiol* 2. doi:10.3389/fcimb.2012.00116.

Seng, R., Leungtongkam, U., Thummeepak, R., Chatdumrong, W., and Sitthisak, S. (2017). High prevalence of methicillin-resistant coagulase-negative staphylococci isolated from a university environment in Thailand. *Int. Microbiol.* 20, 65–73. doi:10.2436/20.1501.01.286.

Stähler, F. N., Odenbreit, S., Haas, R., Wilrich, J., Vliet, A. H. M. V., Kusters, J. G., et al. (2006). The Novel Helicobacter pylori CznABC Metal Efflux Pump Is Required for Cadmium, Zinc, and Nickel Resistance, Urease Modulation, and Gastric Colonization. *Infect Immun* 74, 3845–3852. doi:10.1128/IAI.02025-05.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi:10.1093/bioinformatics/btu033.

Torres-Guerrero, E., Quintanilla-Cedillo, M. R., Ruiz-Esmenjaud, J., and Arenas, R. (2017). Leishmaniasis: a review. *F1000Res* 6. doi:10.12688/f1000research.11120.1.

Tulinski, P., Fluit, A. C., Wagenaar, J. A., Mevius, D., Vijver, L. van de, and Duim, B. (2012). Methicillin-Resistant Coagulase-Negative Staphylococci on Pig Farms as a Reservoir of Heterogeneous Staphylococcal Cassette Chromosome mec Elements. *Appl. Environ. Microbiol.* 78, 299–304. doi:10.1128/AEM.05594-11.

Xu, Z., Misra, R., Jamrozy, D., Paterson, G. K., Cutler, R. R., Holmes, M. A., et al. (2018a). Whole Genome Sequence and Comparative Genomics Analysis of Multi-drug Resistant Environmental Staphylococcus epidermidis ST59. *G3: Genes, Genomes, Genetics* 8, 2225–2230. doi:10.1534/g3.118.200314.

Xu, Z., Mkrtchyan, H. V., and Cutler, R. R. (2015). Antibiotic resistance and *mecA* characterization of coagulase-negative staphylococci isolated from three hotels in London, UK. *Front Microbiol* 6. doi:10.3389/fmicb.2015.00947.

Xu, Z., Shah, H. N., Misra, R., Chen, J., Zhang, W., Liu, Y., et al. (2018b). The prevalence, antibiotic resistance and *mecA* characterization of coagulase negative staphylococci recovered from non-healthcare settings in London, UK. *Antimicrobial Resistance & Infection Control* 7, 73. doi:10.1186/s13756-018-0367-4.

## Tables

**Table 1: *mecA*$^+$ coagulase negative staphylococcal isolates recovered from East and West London public settings**

| Isolate ID | Species (*mecA*$^+$) | Public settings in London |
|---|---|---|
| 1 | *S. haemolyticus* | ELC |
| 93 | *S. haemolyticus* | ELC |
| 99 | *S. haemolyticus* | ELC |
| 105 | *S. haemolyticus* | ELC |
| 207 | *S. hominis* | WLC |
| 208 | *S. hominis* | WLC |
| 209 | *S. hominis* | WLC |
| 211 | *S. cohnii* | WLC |
| 321 | *S. epidermidis* | ELH |
| 327 | *S. epidermidis* | ELH |
| 329 | *S. epidermidis* | ELH |
| 343 | *S. cohnii* | ELH |
| 349 | *S. cohnii* | ELH |
| 355 | *S. epidermidis* | ELH |
| 361 | *S. haemolyticus* | ELH |
| 372 | *S. hominis* | ELH |
| 373 | *S. haemolyticus* | ELH |
| 385 | *S. hominis* | ELH |
| 386 | *S. hominis* | ELH |
| 387 | *S. hominis* | ELH |
| 407 | *S. epidermidis* | ELH |

| | | |
|---|---|---|
| 435 | *S. epidermidis* | WLH |
| 436 | *S. epidermidis* | WLH |
| 445 | *S. haemolyticus* | WLH |
| 465 | *S. epidermidis* | WLH |
| 475 | *S. epidermidis* | WLH |
| 479 | *S. hominis* | WLH |
| 492 | *S. haemolyticus* | WLH |
| 506 | *S. haemolyticus* | WLH |
| 538 | *S. haemolyticus* | WLH |
| 620 | *S. hominis* | WLH |
| 623 | *S. hominis* | WLH |
| 631 | *S. epidermidis* | WLH |
| 664 | *S. epidermidis* | WLH |
| 673 | *S. epidermidis* | WLH |
| 699 | *S. warneri* | WLH |
| 700 | *S. warneri* | WLH |
| 702 | *S. warneri* | WLH |
| 711 | *S. epidermidis* | WLH |
| 712 | *S. epidermidis* | WLH |
| 713 | *S. epidermidis* | WLH |
| 715 | *S. epidermidis* | WLH |
| 716 | *S. epidermidis* | WLH |

ELC= East London Community; WLC= West London Community ELH= East London Hospital; WLH=West London Hospital