First Description of Novel Arginine Catabolic Mobile Elements (ACMEs) Types IV and V Harboring a *kdp* Operon in *Staphylococcus epidermidis* Characterized by Whole Genome Sequencing

Aoife M. O'Connor[a], Brenda A. McManus[a], David C. Coleman[a*]

[a]Microbiology Research Unit, Division of Oral Biosciences, Dublin Dental University Hospital, University of Dublin, Trinity College Dublin, Dublin 2, Republic of Ireland

Running title: **Characterization of ACME types IV and V**

*** Correspondence:**
Corresponding Author David Coleman, Mailing address: Microbiology Research Unit, Division of Oral Biosciences, Dublin Dental University Hospital, University of Dublin, Trinity College Dublin, Dublin 2, Republic of Ireland. Phone: +353 1 6127276. Fax: +353 1 6127295. E-mail : david.coleman@dental.tcd.ie.

A.M.O'C. and B.A.McM. contributed equally to this work.

Abbreviations: arginine catabolic mobile element; ACME, whole genome sequencing; WGS, direct repeat sequences; DRs, multilocus sequence typing; MLST, sequence types; STs, Clustered Regularly Interspaced Short Palindromic Repeats; CRISPR, methicillin-resistant *Staphylococcus aureus*; MRSA, staphylococcal chromosomal cassette *mec*; SCC*mec*, coagulase negative staphylococci; CoNS

Abstract

The arginine catabolic mobile element (ACME) was first described in the methicillin-resistant *Staphylococcus aureus* strain USA300 and is thought to facilitate survival on skin. To date three distinct ACME types have been characterized comprehensively in *S. aureus* and/or *Staphylococcus epidermidis*. Type I harbors the *arc* and *opp*3 operons encoding an arginine deaminase pathway and an oligopeptide permease ABC transporter, respectively, type II harbors the *arc* operon only, and type III harbors the *opp3* operon only.

To investigate the diversity and detailed genetic organization of ACME, whole genome sequencing (WGS) was performed on 32 ACME-harboring oro-nasal *S. epidermidis* isolates using MiSeq- and PacBio-based WGS platforms. In nine isolates the ACMEs lacked the *opp3* operon, but harbored a complete *kdp* operon (*kdp*E/D/A/B/C) located a maximum of 2.8 kb upstream of the *arc* operon. The *kdp* operon exhibited 63% DNA sequence identity to the native *S. aureus kdp* operon. These findings identified a novel, previously undescribed ACME type (designated ACME IV), which could be subtyped (IVa and IVb) based on distinct 5′ flanking direct repeat sequences (DRs).

Multilocus sequence typing (MLST) sequences extracted from the WGS data identified the sequence types (STs) of the isolates investigated. Four of the nine ACME IV isolates belonged to ST153, and one to ST17, a single locus variant of ST153.

A tenth isolate, identified as ST5, harbored another novel ACME type (designated ACME V) containing the *kdp*, *arc* and *opp3* operons and flanked by DR_F, and DR_B but lacked any internal DRs. ACME V was colocated with a staphylococcal chromosome cassette *mec* (SCC*mec*) IV element and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) in a 116.9 kb composite island.

The extensive genetic diversity of ACME in *S. epidermidis* has been further elucidated by WGS, revealing two novel ACME types IV and V for the first time.

Keywords: ACME, *Staphylococcus epidermidis*, *kdp* operon, *arc* operon, *opp3* operon, oral cavity.

1 Introduction

The arginine catabolic mobile element (ACME) was first described in the methicillin-resistant *Staphylococcus aureus* (MRSA) strain USA300 and is thought to contribute to the transmission, colonization and persistence of this pathogen on human skin (Diep et al., 2008; Planet et al., 2013). Like the staphylococcal chromosomal cassette *mec* (SCC*mec*) element, ACME integrates into the staphylococcal chromosomal *orfX* locus using the *attB* attachment site and is flanked by direct repeat sequences (DRs) at integration sites. Like SCC*mec*, ACME is thought to have originated in coagulase negative staphylococci (CoNS), specifically *Staphylococcus epidermidis*, in which the prevalence and diversity of both SCC*mec* and ACME is significantly greater than in *S. aureus.* In many cases, these ACMEs also contain internal DRs, indicating that these elements are assembled in a stepwise, modular manner (Thurlow et al., 2013).

To date, three distinct ACME types have been characterized in detail; type I harbors both the *arc* and *opp*3 operons which encode an arginine deaminase pathway and an oligopeptide permease ABC transporter, respectively, type II harbors the *arc* operon only, and type III harbors the *opp3* operon only. ACME types I and II and variants thereof have been described in *S. aureus* (Diep et al., 2006; Rolo et al., 2012; Shore et al., 2011) and all three ACME types and variants thereof have been described in *S. epidermidis* (Barbier et al., 2011; McManus et al., 2017; Miragaia et al., 2009; Onishi et al., 2013; Soroush et al., 2016). ACME variants have been described in *S. epidermidis* based on distinct PCR-based scanning patterns of 30 overlapping segments of DNA sequence, 1-2 kb in size (Miragaia et al., 2009). Additional studies have identified distinct ACME-*arc* and ACME-*opp3* allotypes by PCR-based amplification and DNA sequence analysis of the ACME-*arcA* and -*opp3AB* genes (Barbier et al., 2011), respectively. Other studies identified distinct, truncated variants of ACME type I (designated types Δ1.1-3) and ACME type II (designated type ΔII) in *S. epidermidis* and MRSA, using PCR-profiling and Sanger-based sequencing. These truncated variants of ACME were based on variations in the nucleotide sequence of the regions surrounding the *arc* and *opp3* operons or of the *opp3* operon itself (Onishi et al., 2013; Urushibara et al., 2016).

The importance of *S. epidermidis* as a causative agent of various community acquired diseases and infections associated with indwelling medical devices is being increasingly recognised and in this context, ACME likely plays a significant role in successful host colonization and the potential accumulation and spread of genes encoding antimicrobial resistance. Furthermore, the evolution of ACME in *S. epidermidis* could have important consequences for the epidemiology of *S. aureus* due to the capability of this species to serve as a genetic reservoir for *S. aureus*.

As part of a larger study investigating the prevalence and structural diversity of ACME, 32 oro-nasal *S. epidermidis* isolates recovered from orally healthy patients with or without dental implants, and from patients with periodontal disease or peri-implantitis, in which ACME was detected using ACME-*arc*- and ACME-*opp3*- specific primers, as previously described (McManus et al., 2017) were further characterized using whole genome sequencing (WGS). This was undertaken to elucidate the detailed genetic organization and diversity of these ACMEs, as such investigations may yield new insights into the evolutionary origins and spread of each ACME type. Analysis of the WGS data revealed a structurally unique group of ACMEs which consistently harbored a *kdp* operon encoding a ABC transporter upstream and adjacent to the *arc* operon in 10 *S. epidermidis* isolates, which indicated that these ACMEs represented highly distinct, previously undescribed ACME types. In addition, the

presence of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) were identified downstream of the *kdp* and *arc* operons in one of these isolates.

2 Materials and Methods

2.1 Isolates

The isolates investigated in this study were recovered from nasal swabs, subgingival sites or oral rinse samples taken by qualified Dentists from patients attending the Dublin Dental University Hospital, Ireland. Ethical approval was granted by the Faculty of Health Sciences Ethics Committee of Trinity College Dublin in February 2014.

Subgingival sites were sampled by inserting a PerioPaper™ gingival fluid collection strip (Oroflow Inc., NY, USA) into the subgingival crevice for 30 s. Following sampling the collection strips were placed in sterile 2 ml screw-capped tubes (Sarstedt AG & Co., Numbrecht, Germany) containing 1 ml of nutrient broth (Oxoid Ltd., Hampshire, UK). Oral rinse samples were collected by providing participants with sterile 100 ml polypropylene containers (Sarstedt AG & Co., Wexford, Ireland) containing 25 ml sterile phosphate buffered saline and instructing participants to rinse their mouths for 30 s before returning the fluid to the container. Following sampling, all samples were transported immediately to the microbiology laboratory and processed within 4 h.

Vials containing PerioPaper™ strips suspended in nutrient broth were vortexed at maximum speed for 1 min and 100 µl aliquots of the resulting cell suspension were plated onto mannitol salt agar and Sa*Select* (Bio-Rad, Hertfordshire, United Kingdom) agar. Oral rinse samples were processed by transferring a 1 ml aliquot to a sterile 1.5 ml Eppendorf Safe-lock™ microfuge tube (Eppendorf, Hamburg, Germany) and centrifuged at 20,000 x *g* for 1 min, after which the supernatant was discarded and the pellet resuspended in 200 µl nutrient broth. To isolate staphylococcal colonies, 100 µl aliquots of this cell suspension were plated on mannitol salt agar and Sa*Select*, both of which were incubated at 37°C for 48 h in a static incubator (Gallenkamp, Leicester, United Kingdom).

Bacterial isolates were cultured on Columbia blood agar (Fannin Ltd., Dublin, Republic of Ireland) at 37ºC for 48 h prior to identification by Matrix Assisted Laser Desorption Ionization Time-of-Flight (MALDI-TOF) technology using the VITEK® MS system (bioMérieux, Marcy L'Etoile, France) according to the manufacturer's instructions. All isolates were stored on Microbank™ storage beads (Pro-lab diagnostics, Cheshire, UK) at -80ºC.

2.2 Whole Genome Sequence Analysis

The genome sequence of 32 isolates, selected as representatives of different patients, patient groups, oro-nasal sample sites and each previously described ACME type was determined using a MiSeq sequencer (Illumina, Essex, United Kingdom). One additional isolate was sequenced using a Pacific Biosciences (PacBio) RS sequencing platform (CA, USA) at an average coverage of 302x with subsequent Hierarchal Genome Assembly Process (HGAP.3) analysis (The Genome Analysis Centre [TGAC], Norwich, United Kingdom). Genomic DNA extraction and library construction was performed as previously described (Earls et al., 2017).

For each isolate, reads were checked for quality and then aligned to a selection of ACMEs and SCC*mec* elements previously characterized in *S. aureus* and *S. epidermidis* (Diep et al., 2006; McManus et al., 2017; Zhang et al., 2003) in order to select the most appropriate

reference ACME type to use as a scaffold. This was performed using the Burrows-Wheeler aligner (BWA) tool in SPAdes version 3.6 (http://cab.spbu.ru/software/spades/). Following these analyses, the ACME I sequence from MRSA USA300 strain FPR3757 (GenBank accession number CP000255.1) was selected as the most appropriate scaffold for isolates harboring both the ACME-*arc* and ACME-*opp3* genes and the ACME II sequence from *S. epidermidis* strain ATCC12228 (GenBank accession number AE015929) was selected as the most appropriate scaffold for ACMEs harboring only the ACME-*arc* genes.

For each isolate, contigs were generated by BWA assembly and aligned to the most appropriate reference scaffold. Contigs containing sequences previously associated with SCC*mec* or ACME were selected and annotated using the BioNumerics Genome Analysis Tool (GAT) plug-in version 7.6 (Applied Maths, Sint-Martens-Latem, Belgium). For each isolate investigated, ACME–associated genes were identified on between one and six separate contigs (Table 1). These contigs were organized and reorientated as appropriate using the relevant ACME scaffold and Artemis sequence viewer (Berriman and Rutherford, 2003) and Artemis Comparison Tool (Carver et al., 2005). Further annotation was carried out using BLAST software (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

In order to confirm the genetic organization and orientation of contigs, primers were designed using the Artemis sequence viewer (Berriman and Rutherford, 2003) that targeted a minimum distance of 200 nucleotides from the contig boundaries. The target specificity of primers was confirmed using BLAST software (https://blast.ncbi.nlm.nih.gov/Blast.cgi) (Supplementary Table S1). All primers were supplied by Sigma-Aldrich Ltd. (Wicklow, Republic of Ireland). Amplification products were subjected to Sanger-based sequencing carried out commercially by Source BioScience (Waterford, Republic of Ireland).

2.3 Multilocus Sequence Typing (MLST)

The sequence type (ST) of each isolate was determined by submitting the relevant genomic regions (Thomas et al., 2007) to the *S. epidermidis* MLST online database (https://pubmlst.org/sepidermidis/).

2.4 Nucleotide sequence accession numbers

The nucleotide sequences of the nine ACME type IVs characterized have been submitted to GenBank under the accession numbers MG787414 - MG787422 (Table 1). The nucleotide sequences of the ACME type V characterized has been submitted to GenBank under the accession number MG787423.

3 Results

3.1 Identification of previously described ACME types

The ACMEs harbored by 22 isolates investigated encoded only the *arc* and/or *opp3* operons indicative of the previously described ACME types I, II and III and have not been further described in the present study.

3.2 Identification and molecular characterization of ACME IV

Analysis of WGS data revealed that nine of the 32 *S. epidermidis* isolates sequenced harbored ACMEs encoding a *kdp* operon and the *arc* operon, but lacking the *opp3* operon. All nine isolates lacked the *mecA* gene, however, in six of these isolates (33BR, 120PPC, I9OR1,

I14OR4, PS21NS and PS30PH), additional genes such as *sdr*H, *speG* or SCC-associated genes were identified upstream of ACME (Figure 1). In each of the nine ACMEs characterized, the *kdp* operon was located adjacent to the *arc* operon, separated by a maximum of 2.8 kb. In contrast, the *arc* and *opp3* operons in ACME I are separated by 11.5 kb (Diep et al., 2006). Based on the presence of the *kdp* operon directly adjacent to the *arc* operon, and the lack of the *opp3* operon in these ACMEs, we propose that these novel ACMEs be distinguished as ACME type IV, corresponding to the previously described ACME II, but recognizing the presence of the additional *kdp* operon. The structural organization of the composite island including ACME IV was identical in isolates 120PPC, I9OR1, I14OR1 and these composite islands encoding ACME exhibited >99.9% nucleotide sequence identity to each other (Figure 1).

The ACME IVs could be divided into two distinct subtypes (IVa and IVb) based on the distinct combinations of flanking DRs. In eight of these nine isolates, ACME IVa was demarcated by DR_B and DR_C, and in the remaining isolate, ACME IVb was demarcated by DR_F and DR_C (Table 2 and Figure 1). An internal DR_G was identified within the ACME of all nine isolates (Figure 1), correlating with the presence of DR_G within the reference ACME II previously described in *S. epidermidis* (GenBank accession AE015929) but absent in the reference ACME I from *S. aureus* (GenBank accession CP000255.1).

The ST of each isolate was determined from the WGS data (Table 1). Four of the nine isolates harboring ACME IV belonged to ST153, and one belonged to ST17, which was a single locus variant of ST153 and differed by a single nucleotide in the *arcC* MLST locus, even though the nine isolates were recovered from separate patients from three distinct patient groups.

3.3 Identification and molecular characterization of ACME V

The final ACME-positive isolate investigated (PS19PH) was methicillin-resistant and harbored the *kdp, arc* and *opp3* operons alongside SCC*mec* IV (Table 1 and Figure 2). Similar to ACME IV, the *kdp* operon was located directly adjacent to the *arc* operon, separated by 2.5 kb. The *opp3* operon was located 2.5 kb downstream of the *arc* genes (Figure 2). Based on the presence of the *arc* and *opp3* operons in addition to the *kdp* operon, we propose that this novel ACME be designated as ACME type V (Figure 2). The ACME from PS19PH was part of a 116.9 kb composite island, which consisted of a SCC*mec* IV module, the ACME V module and a CRISPR module which was separated from a *copA* gene and *ars* operon (Figure 2) by DR_G. All of these modules were identified on the same contig following WGS using the PacBio platform and structures were confirmed by PCRs using primers specific for each distinct genomic region (Supplementary Table S1).

The CRISPR sequence was identified downstream of ACME V, separated by two sets of DRs (Figure 2), DR_B and DR_L (Table 2). The CRISPR element identified within isolate PS19PH harbored the caspase-encoding genes *cas1* and *cas2*, and exhibited 99% DNA sequence homology with that previously identified in *S. epidermidis* isolates RP62A (Gill et al., 2005) and SE95 (Genbank accession number CP024437.1) and 92% DNA sequence identity with CRISPRs identified in other CoNS species. Interestingly, DR_G, which is also present within ACME IV, was identified between the CRISPR-encoding module and the module harboring the *copA* and *ars* genes encoding resistance to heavy metals (Figure 2).

3.4 The *kdp* operon

In all ten isolates characterized in detail in the present study, the complete *kdp* operon was consistently detected upstream of the *arc* operon separated by a maximum distance of 2.8 kb. The *kdp* genes detected in ACME IV and V exhibited 75% DNA sequence identity and the same gene orientation and synteny to the corresponding *kdp* operon previously described in SCC*mec* II in MRSA and *S. epidermidis* (Gill et al., 2005; Ito et al., 1999), as well as 63% DNA sequence identity to the native *kdp* operon in *S. aureus* (GenBank accession number CP000253.1 and GenBank accession number BA000017.4). The *kdp* genes harbored by ACME exhibited an average amino acid identity of 61.9% with the corresponding five amino acid sequences previously described for the native *kdp* operon of *S. aureus* (GenBank accession number BA000017.4). Similarly, the *kdp* genes harbored by ACME exhibited an average amino acid identity of 79.0% to those previously described in SCC*mec* II (Gill et al., 2005; Ito et al., 1999).

The frameshift mutation previously identified in the *kdpA* homologue described in SCC*mec* II (Ito et al., 1999), was not detected in any of the isolates harboring ACME IV in the present study, however, a point mutation was identified at position +1360 of the *kdpA* ORF harbored by ACME V, resulting in a premature stop codon and the truncation of the encoded potassium binding and transporting protein. With the exception of this point mutation in the *kdpA* gene harbored by ACME V, the *kdp* operon was well conserved among the *S. epidermidis* isolates investigated here, exhibiting >99.4% nucleotide identity for each *kdp* gene.

A native *kdp* operon could not be detected in the WGS data obtained from the 10 *S. epidermidis* isolates investigated. These investigations were carried out using BLAST searches based on DNA sequences from both the native and SCC*mec kdp* operons of *S. aureus* and the ACME *kdp* operon of *S. epidermidis.*

4 Discussion

This study identified the existence of and characterized the genetic organization of two novel ACME types, designated IV and V, in oro-nasal *S. epidermidis* isolates for the first time using WGS. All novel ACMEs characterized in this investigation harbored the *kdp* operon (*kdp*E/D/A/B/C) which was consistently located a maximum distance of 2.8 kb upstream of the *arc* gene cluster (Figures 1 and 2). The *kdp* operon encodes a potassium transporter system that is composed of the KdpDE two-component system and the transporter components KdpABC. In *S. aureus*, the native *kdp* operon has been identified in addition to other potassium uptake systems and determined to be fully functional (Price-Whelan et al., 2013; Xue et al., 2011). It encodes a high-affinity multicomponent transporter that is strongly induced in conditions of high osmolarity and likely contributes to the osmotolerance of *S. aureus*, enabling this species to survive and grow on human skin. Furthermore, previous research suggests that in *S. aureus, kdpE* encodes a DNA-binding protein that likely acts as a method of global transcriptional regulation for a multitude of virulence genes, thus potentially contributing to the pathogenesis of staphylococcal infection (Xue et al., 2011). A second *kdp* operon exhibiting approximately 63% nucleotide sequence identity to the native *S. aureus kdp* operon has been observed within SCC*mec* type II elements in both MRSA and *S. epidermidis*, and has therefore been previously associated with mobile genetic elements integrated at *orfX* (Gill et al., 2005; International Working Group on the Classification of Staphylococcal Cassette Chromosome, 2009). The *kdp* operon harbored by ACME exhibited an average amino acid sequence identity of 61.9% and 79.0% with the native *S. aureus kdp*

operon and the *kdp* operon harbored by SCC*mec* II, respectively. Based on this relatively low level of amino acid homology, it is unlikely that the ACME *kdp* operon is derived from either of these two potential sources. We suggest that the true origin of the *kdp* operon described in the present study is likely another coagulase negative staphylococcal species.

Previous studies have also revealed the colocation of SCC*mec* IV with ACME I in the MRSA_USA300 strain FPR3757 (Diep et al., 2006), however the DR that separated SCC*mec* IV from ACME I (DR_I, Table 2) in the USA300 MRSA strain FPR3757 differed from the DR identified in the present study in *S. epidermidis* PS19PH between SCC*mec* IV and ACME V (DR_F, Table 2). These findings suggest that the colocation of SCC*mec* IV and ACME V in *S. epidermidis* isolate PS19PH arose from a distinct process to that of SCC*mec* IV and ACME I in MRSA USA300.

In the present investigation, we identified two distinct ACME IV subtypes, IVa and IVb, based on the detection of distinct DRs flanking these ACMEs. As DRs play a crucial role in the stepwise, modular-based assembly and evolution of ACMEs, we believe that assigning subtypes based on these DRs is appropriate and comparable to the use of joining regions in the definition of SCC*mec* subtypes (International Working Group on the Classification of Staphylococcal Cassette Chromosome, 2009).

The *copA* and *ars* genes have been identified adjacent to DR_G upstream of ACME III in previously characterized composite islands in *S. epidermidis* (McManus et al., 2017), further highlighting the potential ability of ACME to accumulate antimicrobial resistance-encoding genes (Diep et al., 2006; McManus et al., 2017). Furthermore, the arrangement of these genes downstream of ACME V and CRISPR in the composite island harbored by isolate PS19PH to the alternative location previously reported in ACME III is in agreement with previous investigations that suggested ACME typically evolves in a stepwise, modular-based method (Thurlow et al., 2013).

The CRISPR element is an array of multiple short DRs separated by comparatively short segments of DNA accompanied by CRISPR-associated (*cas*) genes encoding caspases. In combination, this constitutes a prokaryotic defense mechanism against foreign DNA. The prevalence of CRISPR in *S. epidermidis* is rare, detected in less than 10% of isolates (Rossi et al., 2017), however it has previously been detected downstream of an SCC*mec* II element in *S. epidermidis* (Gill et al., 2005), and associated with composite islands in specific lineages of MRSA (Kinnevey et al., 2013).

Interestingly, five of the nine isolates that harbored ACME IV were identified as ST153 or as single locus variants of this ST (Table 1). To date, only three other ST153 isolates have been identified in the *S. epidermidis* MLST database, one of which was another oral isolate from a patient in Ireland (MLST database accessed 6th February 2018). It is possible that the origin(s) of ACME IV is linked with this lineage, reflecting the findings of a previous study that indicated that the origin of ACME III is possibly associated with ST329 (McManus et al., 2017) in *S. epidermidis*. Similarly, the enrichment of these STs could reflect the fact that all of the isolates recovered in the current investigation were recovered from the oro-nasal cavities of individuals in Ireland. However, other investigators have previously revealed the association of a particular truncated ACME I variant with methicillin-resistant *S. epidermidis* isolates belonging to ST5 recovered from a variety of clinical specimens (Onishi et al., 2013). Previous studies mainly relied on PCR-based identification of the *arcA* gene and the *opp3AB* genes previously described in ACME types I-III (Barbier et al., 2011; Onishi et al., 2013; Urushibara et al., 2016). Based on this approach, the potential presence of the *kdp* operon

identified in the present study within these ACMEs would not have been detected. This highlights the considerable advantage and importance of using WGS in the characterization of mobile genetic elements such as ACME and SCC*mec*. Future WGS-based investigations of such elements in staphylococci will likely reveal further novel SCC*mec* and ACME types.

Previous research suggested that the constitutive expression of the ACME-*arc* genes confers a selective advantage, facilitating the survival of staphylococci under acidic conditions such as lactic acid on human skin and mucous membranes (Lindgren et al., 2014; Planet et al., 2013). The role and potential benefit conferred by the *opp3* genes is less apparent, however, these genes encode a permease likely involved in the transport of a multitude of metabolites. The *kdp* operon encodes a potassium transport system that is upregulated in conditions of high-osmolarity, enabling cells to increase intracellular potassium concentrations for maintaining intracellular pH homeostasis, cell physiology and metabolic processes (Price-Whelan et al., 2013). The concentration of K^+ ions has been found to be at least four-fold higher than other cations such as Na^+ and Ca^{++} in the fluid portion of dental plaque (Margolis and Moreno, 1994) and thus it is likely that the *kdp* operon harbored by ACME IV and V confers a selective advantage to *S. epidermidis* in the oral cavity. Recent research has revealed the importance of potassium homeostasis for biofilm formation, stress tolerance and survival of the opportunistic oral pathogen, *Streptococcus mutans*, in dental plaque (Binepal et al., 2016).

The present investigation used WGS to reveal the existence of two novel ACME types containing the *kdp* operon and harbored by *S. epidermidis* for the first time. These were designated as ACME types IV and V and highlights the extensive genetic diversity present among ACME in *S. epidermidis* and the genetic reservoir that exists for potential spread into *S. aureus.* It is highly likely that future WGS-based studies will reveal the presence of additional ACME types and subtypes in staphylococci.

5 Declaration of Interest

None

6 Author Contributions

AOC conceived and designed the study, performed the WGS data analysis and drafted the manuscript. BMcM conceived the study and helped with the study co-ordination, WGS data analysis and wrote the manuscript. DC conceived the study, purchased the required materials, assisted with data analysis and drafted the manuscript. All authors read and approved the final manuscript.

7 Funding

This work was supported by the Microbiology Research Unit, Dublin Dental University Hospital. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

8 Acknowledgments

We thank Peter Slickers at the InfectoGnostics Research Campus, Jena, Germany for technical assistance with *de novo* assemblies using SPAdes software.

365 Table 1. *Staphylococcus epidermidis* isolates investigated harboring novel ACME types

Isolate	**Patient group**[a]	**Isolate origin**[a]	**ACME-*kdp*/*arc*/*opp3* operons present**	**Number of contigs**[b]	**ACME type**	**ACME GenBank accession number**	**ST**
P8OR3	PD	OR	*kdp* and *arc*	2	IVb	MG787414	210
I9OR1	HWI	OR	*kdp* and *arc*	3	IVa	MG787415	153
I14OR1	HWI	OR	*kdp* and *arc*	4	IVa	MG787416	153
120PPC	OH	SG	*kdp* and *arc*	4	IVa	MG787417	153
218PP361	OH	SG	*kdp* and *arc*	1	IVa	MG787418	130
33BR	OH	OR	*kdp* and *arc*	6	IVa	MG787419	17
PS21NS	PI	NS	*kdp* and *arc*	2	IVa	MG787420	297
PS30PH	PI	SG	*kdp* and *arc*	5	IVa	MG787421	153
PS36PD	PI	SG	*kdp* and *arc*	3	IVa	MG787422	432
PS19PH[c]	PI	SG	*kdp, arc* and *opp3*	1	V	MG787423	5

366 [a]Abbreviations: PD; periodontal disease, HWI; healthy patients with implants, OH; orally
367 healthy, PI; peri-implantitis, OR; oral rinse, SG; subgingival site, NS; nasal swab, ST;
368 sequence type.

369 [b]Number of assembled MiSeq-generated WGS contigs containing sequences previously
370 associated with ACME or SCC elements

371 [c]Isolate PS19PH was subjected to WGS using the PacBio RS sequencing platform.

372

373 Table 2. Direct repeat sequences (DRs) identified among ACME types investigated

374

DR	**Sequence (5'-3')**
DR_A	GAAGCATATCATAAATGA
DR_B	GAAGCGTATCACAAATAA
DR_C	GAAGCGTATCGTAAGTGA
DR_F	GAAAGTTATCATAAGTGA
DR_G	GAAGCGTATAATAAGTAA
DR_J	GAGGCGTATCATAAGTAA
DR_L	GAAGCATATCATAAGTGA
DR_N	GAAGCGTATCACAAATGA
DR_P	GAAGCTTATCATAAATGA

9 FIGURE LEGENDS

FIGURE 1 Schematic diagram showing the genetic organization of novel ACME type IV elements characterized in nine oral *S. epidermidis* isolates

Schematic diagram showing the genetic organization of previously described ACME type II (A) in *S. epidermidis* (GenBank accession number AE015929) and the comparative organization of ACME type IV identified in nine distinct oro-nasal methicillin-susceptibile *S. epidermidis* isolates, defined according to the presence of the ACME-*arc* and ACME-*kdp* operons and identified by whole genome sequencing. Two distinct ACME IV subtypes (IVa and IVb) were defined according to the distinct combinations of flanking DRs, DR_B and DR_C (B - G) and DR_F and DR_C (H), respectively. Arrows indicate the position and orientation of open reading frames. Genes commonly associated with antimicrobial resistance, SCC or ACME are shaded in color; ACME-*arc* (red), ACME-*kdp* (purple), *speG* (dark grey), *copA* (lime green), *pbp* (dark green), *ccr* (navy) and *tetR* (mustard). For each ACME, *orfX* is indicated in black and specific direct repeat sequences (DRs) identified are indicated. The sequences of each DR are shown in Table 2.

FIGURE 2 Schematic diagram showing the genetic organization of novel ACME type V colocated with SCC*mec* IV and CRISPR in a composite island harbored by an oral *S. epidermidis* isolate.

Schematic diagram showing the genetic organization of previously described ACME type I in MRSA USA300 strain FPR3757 (GenBank accession number CP000255.1) (A) and the comparative organization of the distinct composite island harboring SCC*mec* IV, ACME V, a CRISPR-encoding region and a region encoding the heavy metal resistance *ars* operon and *copA* gene (B) identified in *S. epidermidis* oral isolate PS19PH in this study. Arrows indicate the position and orientation of open reading frames. Genes commonly associated with antimicrobial resistance, SCC or ACME are shaded in color; ACME-*arc* (red), ACME-*kdp* (purple), ACME-*opp3* (blue), *copA* (lime green), *ars* (yellow), *pbp* (dark green), *ccr* (navy) and CRISPR (turquoise). A point mutation was identified at position +1360 of the *kdpA* ORF, resulting in a premature stop codon and the truncation of the encoded potassium binding and transporting protein. The *orfX* is indicated in black and each specific direct repeat sequence (DR) identified as separating each region are indicated. The sequences of each DR are shown in Table 2. The dashed line indicates the adjacent position of the genomic regions harboring the CRISPR and *copA* and *ars* genes directly downstream of ACME V.

References

Barbier, F., Lebeaux, D., Hernandez, D., Delannoy, A.S., Caro, V., François, P., Schrenzel, J., Ruppé, E., Gaillard, K., Wolff, M., Brisse, S., Andremont, A., Ruimy, R., 2011. High prevalence of the arginine catabolic mobile element in carriage isolates of methicillin-resistant *Staphylococcus epidermidis*. J. Antimicrob. Chemother. 66, 29–36. https://doi.org/10.1093/jac/dkq410

Berriman, M., Rutherford, K., 2003. Viewing and annotating sequence data with Artemis. Br. Bioinform 4, 124–132. https://doi.org/NO_DOI

Binepal, G., Gill, K., Crowley, P., Cordova, M., Brady, L.J., Senadheera, D.B., Cvitkovitch, D.G., 2016. Trk2 potassium transport system in *Streptococcus mutans* and its role in potassium homeostasis, biofilm formation, and stress tolerance. J. Bacteriol. 198, 1087–1100. https://doi.org/10.1128/JB.00813-15

Carver, T.J., Rutherford, K.M., Berriman, M., Rajandream, M.A., Barrell, B.G., Parkhill, J., 2005. ACT: the Artemis Comparison Tool. Bioinformatics 21, 3422–3423. https://doi.org/10.1093/bioinformatics/bti553

Diep, B.A., Gill, S.R., Chang, R.F., Phan, T.H., Chen, J.H., Davidson, M.G., Lin, F., Lin, J., Carleton, H.A., Mongodin, E.F., Sensabaugh, G.F., Perdreau-Remington, F., 2006. Complete genome sequence of USA300, an epidemic clone of community-acquired methicillin-resistant *Staphylococcus aureus*. Lancet 367, 731–739. https://doi.org/10.1016/S0140-6736(06)68231-7

Diep, B.A., Stone, G.G., Basuino, L., Graber, C.J., Miller, A., Etages, S. des, Jones, A., Palazzolo-Ballance, A.M., Perdreau-Remington, F., Sensabaugh, G.F., DeLeo, F.R., Chambers, H.F., 2008. The arginine catabolic mobile element and staphylococcal chromosomal cassette *mec* linkage: convergence of virulence and resistance in the USA300 clone of methicillin-resistant *Staphylococcus aureus*. J. Infect. Dis. 197, 1523–1530. https://doi.org/10.1086/587907

Earls, M.R., Kinnevey, P.M., Brennan, G.I., Lazaris, A., Skally, M., O'Connell, B., Humphreys, H., Shore, A.C., Coleman, D.C., 2017. The recent emergence in hospitals of multidrug-resistant community-associated sequence type 1 and *spa* type t127 methicillin-resistant *Staphylococcus aureus* investigated by whole-genome sequencing: Implications for screening. PLoS One 12, e0175542. https://doi.org/10.1371/journal.pone.0175542

Gill, S.R., Fouts, D.E., Archer, G.L., Mongodin, E.F., Deboy, R.T., Ravel, J., Paulsen, I.T., Kolonay, J.F., Brinkac, L., Beanan, M., Dodson, R.J., Daugherty, S.C., Madupu, R., Angiuoli, S. V, Durkin, A.S., Haft, D.H., Vamathevan, J., Khouri, H., Utterback, T., Lee, C., Dimitrov, G., Jiang, L., Qin, H., Weidman, J., Tran, K., Kang, K., Hance, I.R., Nelson, K.E., Fraser, C.M., 2005. Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a biofilm-producing methicillin-resistant *Staphylococcus epidermidis* strain. J. Bacteriol. 187, 2426–2438. https://doi.org/10.1128/JB.187.7.2426-2438.2005

International Working Group on the Classification of Staphylococcal Cassette Chromosome, E., 2009. Classification of staphylococcal cassette chromosome *mec* (SCC*mec*): guidelines for reporting novel SCC*mec* elements. Antimicrob. Agents Chemother. 53,

4961–4967. https://doi.org/10.1128/AAC.00579-09

Ito, T., Katayama, Y., Hiramatsu, K., 1999. Cloning and nucleotide sequence determination of the entire *mec* DNA of pre-methicillin-resistant *Staphylococcus aureus* N315. Antimicrob. Agents Chemother. 43, 1449–1458.

Kinnevey, P.M., Shore, A.C., Brennan, G.I., Sullivan, D.J., Ehricht, R., Monecke, S., Slickers, P., Coleman, D.C., 2013. Emergence of sequence type 779 methicillin-resistant *Staphylococcus aureus* harboring a novel pseudo staphylococcal cassette chromosome *mec* (SCC*mec*)-SCC-SCCCRISPR composite element in Irish hospitals. Antimicrob. Agents Chemother. 57, 524–531. https://doi.org/10.1128/AAC.01689-12

Lindgren, J.K., Thomas, V.C., Olson, M.E., Chaudhari, S.S., Nuxoll, A.S., Schaeffer, C.R., Lindgren, K.E., Jones, J., Zimmerman, M.C., Dunman, P.M., Bayles, K.W., Fey, P.D., 2014. Arginine deiminase in *Staphylococcus epidermidis* functions to augment biofilm maturation through pH homeostasis. J. Bacteriol. 196, 2277–2289. https://doi.org/10.1128/jb.00051-14

Margolis, H.C., Moreno, E.C., 1994. Composition and cariogenic potential of dental plaque fluid. Crit. Rev. Oral. Biol. Med 5, 1–25.

McManus, B.A., O'Connor, A.M., Kinnevey, P.M., O'Sullivan, M., Polyzois, I., Coleman, D.C., 2017. First detailed genetic characterization of the structural organization of type III arginine catabolic mobile elements harbored by *Staphylococcus epidermidis* by using whole-genome sequencing. Antimicrob. Agents Chemother. 61. https://doi.org/10.1128/AAC.01216-17

Miragaia, M., de Lencastre, H., Perdreau-Remington, F., Chambers, H.F., Higashi, J., Sullam, P.M., Lin, J., Wong, K.I., King, K.A., Otto, M., Sensabaugh, G.F., Diep, B.A., 2009. Genetic diversity of arginine catabolic mobile element in *Staphylococcus epidermidis*. PLoS One 4. https://doi.org/10.1371/journal.pone.0007722

Onishi, M., Urushibara, N., Kawaguchiya, M., Ghosh, S., Shinagawa, M., Watanabe, N., Kobayashi, N., 2013. Prevalence and genetic diversity of arginine catabolic mobile element (ACME) in clinical isolates of coagulase-negative staphylococci: Identification of ACME type I variants in *Staphylococcus epidermidis*. Infect. Genet. Evol. 20, 381–388. https://doi.org/10.1016/j.meegid.2013.09.018

Planet, P.J., LaRussa, S.J., Dana, A., Smith, H., Xu, A., Ryan, C., Uhlemann, A.C., Boundy, S., Goldberg, J., Narechania, A., Kulkarni, R., Ratner, A.J., Geoghegan, J.A., Kolokotronis, S.O., Prince, A., 2013. Emergence of the epidemic methicillin-resistant *Staphylococcus aureus* strain USA300 coincides with horizontal transfer of the arginine catabolic mobile element and *speG*-mediated adaptations for survival on skin. MBio 4, e00889-13. https://doi.org/10.1128/mBio.00889-13

Price-Whelan, A., Poon, C.K., Benson, M.A., Eidem, T.T., Roux, C.M., Boyd, J.M., Dunman, P.M., Torres, V.J., Krulwich, T.A., 2013. Transcriptional profiling of *Staphylococcus aureus* during growth in 2 M NaCl leads to clarification of physiological roles for Kdp and Ktr K+ uptake systems. MBio 4. https://doi.org/10.1128/mBio.00407-13

Rolo, J., Miragaia, M., Turlej-Rogacka, A., Empel, J., Bouchami, O., Faria, N.A., Tavares, A., Hryniewicz, W., Fluit, A.C., de Lencastre, H., Group, C.W., 2012. High genetic

496 diversity among community-associated *Staphylococcus aureus* in Europe: results from a
497 multicenter study. PLoS One 7, e34768. https://doi.org/10.1371/journal.pone.0034768

498 Rossi, C.C., Souza-Silva, T., Araujo-Alves, A. V, Giambiagi-deMarval, M., 2017. CRISPR-
499 *cas* systems features and the gene-reservoir role of coagulase-negative staphylococci.
500 Front. Microbiol. 8, 1545. https://doi.org/10.3389/fmicb.2017.01545

501 Shore, A.C., Rossney, A.S., Brennan, O.M., Kinnevey, P.M., Humphreys, H., Sullivan, D.J.,
502 Goering, R. V., Ehricht, R., Monecke, S., Coleman, D.C., 2011. Characterization of a
503 novel arginine catabolic mobile element (ACME) and staphylococcal chromosomal
504 cassette *mec* composite island with significant homology to *Staphylococcus epidermidis*
505 ACME type II in methicillin-resistant *Staphylococcus aureus* genotype . Antimicrob.
506 Agents Chemother. 55, 1896–1905. https://doi.org/10.1128/AAC.01756-10

507 Soroush, S., Jabalameli, F., Taherikalani, M., Amirmozafari, N., Imani Fooladi, A.A.,
508 Asadollahi, K., Beigverdi, R., Emaneini, M., 2016. Investigation of biofilm formation
509 ability, antimicrobial resistance and the staphylococcal cassette chromosome *mec*
510 patterns of methicillin resistant *Staphylococcus epidermidis* with different sequence
511 types isolated from children. Microb. Pathog. 93, 126–130.
512 https://doi.org/10.1016/j.micpath.2016.01.018

513 Thomas, J.C., Vargas, M.R., Miragaia, M., Peacock, S.J., Archer, G.L., Enright, M.C., 2007.
514 Improved multilocus sequence typing scheme for *Staphylococcus epidermidis*. J. Clin.
515 Microbiol. 45, 616–619. https://doi.org/10.1128/jcm.01934-06

516 Thurlow, L.R., Joshi, G.S., Clark, J.R., Spontak, J.S., Neely, C.J., Maile, R., Richardson,
517 A.R., 2013. Functional modularity of the arginine catabolic mobile element contributes
518 to the success of USA300 methicillin-resistant *Staphylococcus aureus*. Cell Host
519 Microbe 13, 100–107. https://doi.org/10.1016/j.chom.2012.11.012

520 Urushibara, N., Kawaguchiya, M., Onishi, M., Mise, K., Aung, M.S., Kobayashi, N., 2016.
521 Novel structures and temporal changes of arginine catabolic mobile elements in
522 methicillin-resistant *Staphylococcus aureus* genotypes ST5-MRSA-II and ST764-
523 MRSA-II in Japan. Antimicrob. Agents Chemother. 60, 3119–3122.
524 https://doi.org/10.1128/AAC.02356-15

525 Xue, T., You, Y., Hong, D., Sun, H., Sun, B., 2011. The *Staphylococcus aureus* KdpDE two-
526 component system couples extracellular K+ sensing and Agr signaling to infection
527 programming. Infect. Immun. 79, 2154–2167. https://doi.org/10.1128/IAI.01180-10

528 Zhang, Y.Q., Ren, S.X., Li, H.L., Wang, Y.X., Fu, G., Yang, J., Qin, Z.Q., Miao, Y.G.,
529 Wang, W.Y., Chen, R.S., Shen, Y., Chen, Z., Yuan, Z.H., Zhao, G.P., Qu, D., Danchin,
530 A., Wen, Y.M., 2003. Genome-based analysis of virulence genes in a non-biofilm-
531 forming *Staphylococcus epidermidis* strain (ATCC 12228). Mol. Microbiol. 49, 1577–
532 1593.

Figure 1

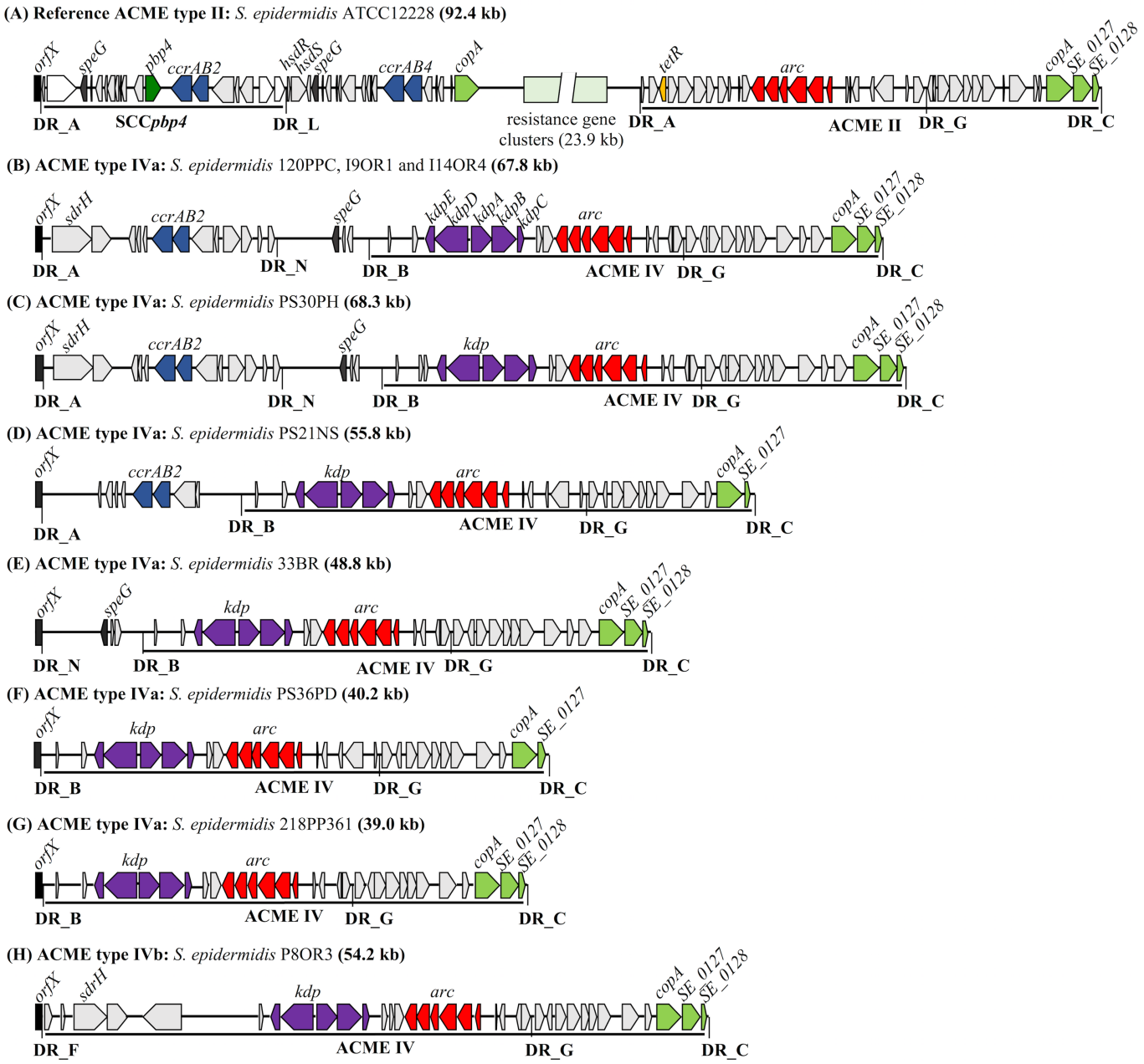
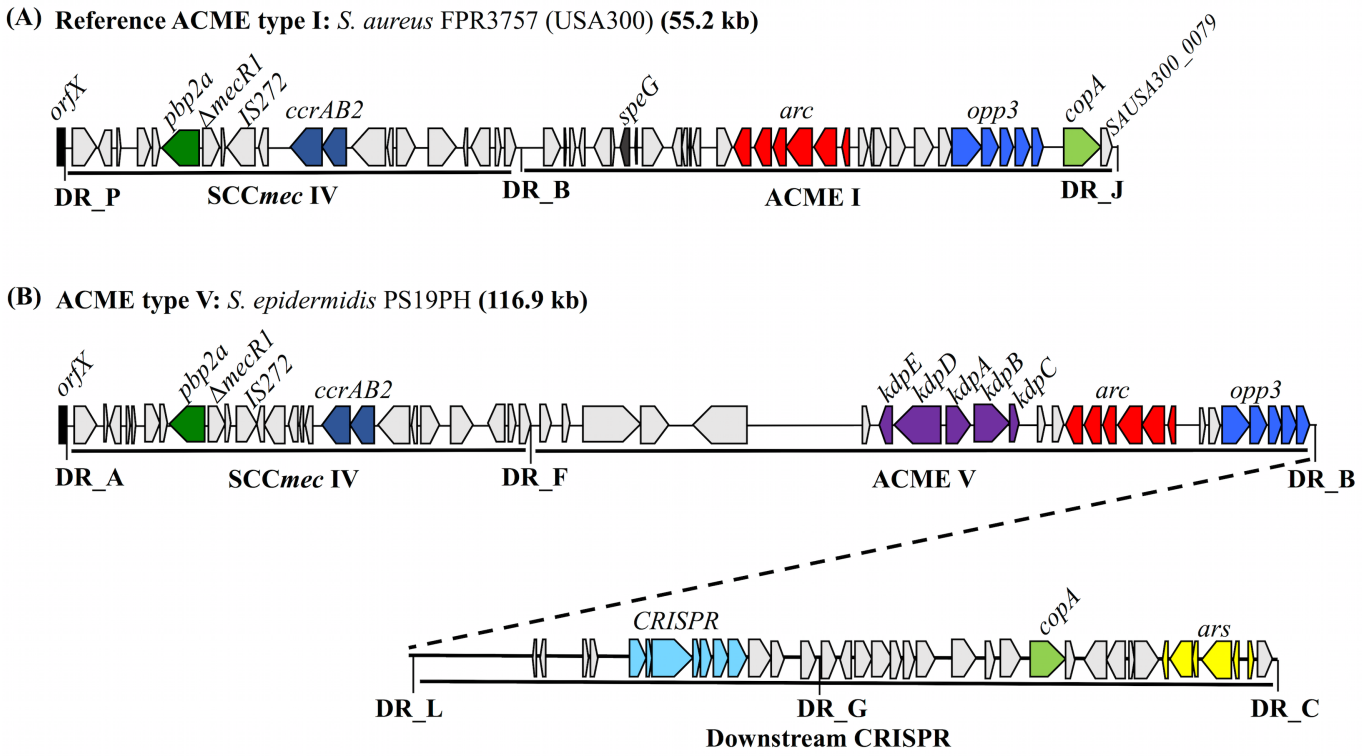

(A) Reference ACME type II: S. epidermidis ATCC12228 (92.4 kb)
orfX
speG
pbp4
ccrAB2
hsdR
hsdS
speG
ccrAB4
copA
tetR
arc
copA
SE_0127
SE_0128
DR_A
SCCpbp4
DR_L
resistance gene clusters (23.9 kb)
DR_A
ACME II
DR_G
DR_C
(B) ACME type IVa: S. epidermidis 120PPC, I9OR1 and I14OR4 (67.8 kb)
orfX
sdrH
ccrAB2
speG
kdpE
kdpD
kdpA
kdpB
kdpC
arc
copA
SE_0127
SE_0128
DR_A
DR_N
DR_B
ACME IV
DR_G
DR_C
(C) ACME type IVa: S. epidermidis PS30PH (68.3 kb)
orfX
sdrH
ccrAB2
speG
kdp
arc
copA
SE_0127
SE_0128
DR_A
DR_N
DR_B
ACME IV
DR_G
DR_C
(D) ACME type IVa: S. epidermidis PS21NS (55.8 kb)
orfX
ccrAB2
kdp
arc
copA
SE_0127
DR_A
DR_B
ACME IV
DR_G
DR_C
(E) ACME type IVa: S. epidermidis 33BR (48.8 kb)
orfX
speG
kdp
arc
copA
SE_0127
SE_0128
DR_N
DR_B
ACME IV
DR_G
DR_C
(F) ACME type IVa: S. epidermidis PS36PD (40.2 kb)
orfX
kdp
arc
copA
SE_0127
DR_B
ACME IV
DR_G
DR_C
(G) ACME type IVa: S. epidermidis 218PP361 (39.0 kb)
orfX
kdp
arc
copA
SE_0127
SE_0128
DR_B
ACME IV
DR_G
DR_C
(H) ACME type IVb: S. epidermidis P8OR3 (54.2 kb)
orfX
sdrH
kdp
arc
copA
SE_0127
SE_0128
DR_F
ACME IV
DR_G
DR_C

Figure 2


(A) Reference ACME type I: S. aureus FPR3757 (USA300) (55.2 kb)
orfX
pbp2a
ΔmecR1
IS272
ccrAB2
speG
arc
opp3
copA
SAUSA300_0079
DR_P
SCCmec IV
DR_B
ACME I
DR_J
(B) ACME type V: S. epidermidis PS19PH (116.9 kb)
orfX
pbp2a
ΔmecR1
IS272
ccrAB2
kdpE
kdpD
kdpA
kdpB
kdpC
arc
opp3
DR_A
SCCmec IV
DR_F
ACME V
DR_B
CRISPR
copA
ars
DR_L
DR_G
DR_C
Downstream CRISPR

Supplementary Table 1. Primers used for confirmation of contig position, orientation, and ACME structures

Isolate	Primer name	Primer sequences (5′-3′)	Amplimer size (bp)
120PPC	120C6iF	GAGAGGCGAAGCATATC	1600
	120C17R	CATAGCGAGGATAATATTGTG	
	120C17F	GATCTGATAGACTGACCCC	1000
	200C18iR	CTATTTTACCGTCTAAAGCG	
	120C20F	CTACATCTACATCAGCATGG	1700
	120C18iR	GTAGGAAGACGAGGCTG	
I9OR1	I9C5iF	GTTGGGATGCCTCAG	1200
	200C18iR	CTATTTTACCGTCTAAAGCG	
	I9C22iF	CATGGGGCAAAGAATATAC	1800
	I9C20iR	GAGTGTATTGTCATGCGATAG	
I14OR4	I14C6F2	GCAGCAGAAAAGAATCAAG	600
	I14C16R2	GCACAGACAATTCGACTTC	
	I14C16F2	GTTATGAGGTTGGGATGC	1100
	I14C19R2	GTTCAGTGCCCTAGGATTATAG	
	I14C19F	CATTAAAGGACAAATCATTAGTG	1700
	I14C17R	CAATTTGCTTTTCTAGACCTAC	
P8OR3	P8C11F2	CGTAGATCTGATAGACTGACC	1300
	200C18iR	CTATTTTACCGTCTAAAGCG	
33BR	33C3iF2	GTTATGAAGCTAGATTAATGGC	1000
	33C38R	GACACAGCCCAAGAAAG	
	33C38F	GACTGACCCCAATTAGTG	1100
	33C47R	CTAATCCTGCTAGAGATGTAATC	
	33C47F	CTCCAAAATGTCTTGCC	4200

	33C56R2	GCAATATCATTGATAAGGGG	
	33C56F	GTTAAATGACCAACAAATTTC	1100
	33C51R2	GTGCAAAGTGTCATGACTAC	
	33C51F	GGGGCAAAGAATATACG	2000
	33C28R	CTAATGTAGGACGTGGAGAC	
PS21NS	P8C11F2	CGTAGATCTGATAGACTGACC	1100
	200C18iR	CTATTTTACCGTCTAAAGCG	
PS30PH	orfX2	CTTACAACGCAGCAACTATG	2000
	I23C17R	CCAGAGGTTGATTCCG	
	P8C11F2	CGTAGATCTGATAGACTGACC	1100
	200C18iR	CTATTTTACCGTCTAAAGCG	
	368C22F	CATGGGGCAAAGAATATACG	2000
	368C17R	CATCGATGACAAGGTCTAATG	
PS36PD	P8C11F2	CGTAGATCTGATAGACTGACC	1100
	200C18iR	CTATTTTACCGTCTAAAGCG	
	166C26F2	GGGACAGAACTTCTTTTAGC	600
	166C15R	GATTGACGTCGACTGAAG	
PS19PH	100-1F	CATTTCTACTTCACCATTATCG	4800
	100-2R	GATAACAACTGGTCGCTTC	
	P16-3F	GATGGAAGTCACAGTATTCTTTG	4200
	100-3R	CCATTTATAAATGAAGAACAATTG	
	100-4F	GTCGTAGCCTAGTGATTGTAGC	4300
	100-52R	CAGTAGTCAAGTCTCCCATCC	
	100-6F	GATTGGCCAAGTGATATTC	3000
	200C18iR	CTATTTTACCGTCTAAAGCG	

100-7F	CTTGTAAGTCACGAGAAATAGTTG	3100
100-8R	GTCGTTGTTAAAAATGAGCAC	
100-9F	GATGCAGGTTGGAAGTAAAC	4000
100-10R	CTGTATGTTTATCAAAAGGCTC	
100-11F	CGAAAGTTCCTAGCTATTCAG	3200
100-12R	CGAGAGATATGTCAGTAATGTTC	
100-13F	GGACTATGCCAATATAGAAAATC	3200
100-14R	CTGATTTTGAAGATGCTTATATG	
100-15F	CACAACCATTTACATTATTAACTG	3500
100-16R	CTTACCTTTTCAATTCACATTTC	