# CHAPTER FOUR

## DISCUSSION

This is the first study to characterize the whole genome of *Staphylococcus aureus* between humans and pigs in Africa. Moreover, the previous studies that involved characterization of *S. aureus* in Kenya are skewed towards only human clinical isolates and importantly all were carried out using conventional typing tools (Omuse et al., 2016, Aiken et al., 2014). Investigation of Multi-locus sequence type (MLST) and *spa* type revealed a significant degree of genetic diversity of clonal lineages circulating in humans and pigs that were sampled from Kiambu, Kenya between October 2015 and August 2016. Furthermore, analysis showed that humans and swine strains were inter-spread across the phylogeny with some having homogenous genetic polymorphisms. In addition, there was no clear distinctions of antimicrobial resistance and virulence genes between the hosts providing evidence for the need for future genomic surveillance and advancement of one health initiative.

The results of determination of population structure clones provided evidence of a considerable degree of genetic variation of *S. aureus* in Kiambu, Kenya. This was comparable with the recent findings of Omuse et al. (Omuse et al., 2016), who identified 40 *spa* types among 55 clinical strains from Aga Khan University hospital and KEMRI in Nairobi, which are located approximately 30 kilometers (Km) from Kiambu county. These *spa* types included t223, t355 and t091 that were also identified in this study (Fig. 6), suggesting that these strains could represent isolates causing hospital infections in Kenya. Investigation of MLST, identified ST188 as the dominant clone, which was present in both swine and human strains (Fig. 5). This ST188 clone, is frequently isolated in hospitals and community acquired infections in Pan Asian regions (Wang et al., 2018), and has been reported to be of high prevalence among healthy humans and various animals in Africa, including chimpanzee and lemurs in Uganda and Madagascar, respectively (Schaumburg et al., 2013), and goats, cats and cattle in Tunisia (Gharsa et al., 2015) and has been associated with disease outbreak among laboratory mice (Sung et al., 2008, McCarthy et al., 2012). This suggest that this clone has potential to colonize different types of hosts which corresponds with the finding of this study. In addition, the strains of ST188 had been

documented to have low level of antibiotic resistance genes. For instance, none of these African isolates, including Kenyan isolates of this study, were MRSA. However, Wang et al. demonstrated during a follow-up study of this lineage between 2012 and 2014 in China that this clone was gaining antibiotic resistance over time (Wang et al., 2018), hence the need for future continuous surveillance of this lineage.

There was no evidence of antimicrobial resistance genes profiles that were associated with either pigs or humans, as expected considering the small number of the remaining good quality genomes in this study. This is in contrast with recent findings by Richardson et al. where they identified specific classes of antimicrobial resistance genes being significantly associated with some host species suggesting important role of antibiotic selection pressure in humans and animals in evolution and emergence of clones with antimicrobial resistance genes (Richardson et al., 2018). Notably, the high prevalence of *blaZ* gene (Fig. 9) that confers resistance to penicillin and ampicillin was expected in this study because penicillin has been commonly used for treatment of infections in humans and animals since 1940s (Gundogan et al., 2005), and in agreement with previous studies in other African countries (Gitau et al., 2018, Katakweba et al., 2016, Akanbi et al., 2017).

The investigation of virulence determinants identified many important genetic markers (Fig. 10) that had been associated previously with diseases and host adaptations. This include *lukF/S* genes that were found in ST152 and ST15 which have been associated with carriage of PVL genes (Schaumburg et al., 2014b). Furthermore, the ST152 clone is frequently isolated in both community and hospital settings in Sub-Saharan Africa (over 40% of the strains) (Ouedraogo et al., 2016). The analysis of PVL in ST152 lineage showed identical distribution of nucleotides sequences of its encoding genes, *lukF-PV* and *lukS-PV* (Fig. 19) except in two strains of the distant clade (Fig. 18). Although PVL is highly specific for human neutrophils, its detection in a swine isolate is consistent with recent reports in Algeria where the investigators identified 4/19 sheep isolates belonging to ST152 being positive for this gene (Agabou et al., 2017) suggesting that PVL could be conserved in this lineage irrespective of the hosts.

Studies have demonstrated that the type of enterotoxins produced by *S. aureus* could be used to indicate the host origin of the strains. The classical enterotoxins SEA and SEC have been shown to be associated with strains that originated from humans and SED from animals

(Gonzalez et al., 2017, Jones et al., 2002). The identification of swine strains with SEC genes in this study could point to a human origin. The co-existence of newly described enterotoxins *sem, sei, seo sen* and *seg* was unsurprising since they belong to enterotoxin gene clusters (*egc*) that are found in the same genomic islands *vSaß* (Yan et al., 2012). Moreover, loci of these *egc* in the *vSaß* had been shown previously to be highly correlated with the clonal lineages (van Belkum et al., 2006, Chao et al., 2015, Song et al., 2016), and were found in this study in ST25 and ST22 lineages. Furthermore, co-detection of *tsst-1*, causing toxic shock syndrome, with *egc* had also been documented in other previous studies (Song et al., 2016).

Typically, ß-hemolysin converting bacteriophages that harbor immune evasion clusters (IECs) genes *sak*, *scn*, *chp* and *sea* are strongly associated with humans (Verkaik et al., 2011a) and its acquisition or loss had been considered specific genetic markers for host adaptations. The presence of these genes in swine strains of the clonal lineages that were shared with the humans indicate that the isolates are likely to have been transferred recently to swine. Proper surveillance needs to be established to monitor strains carrying these genes since they have the potential to cause infections in humans. In addition, there were two strains with IECs phage type H (associated with livestock adaptations) that belonged to ST6 and ST22 lineages which have been demonstrated in previous studies to colonize animals and get adapted by losing these IECs genes. For instances, in monkeys in Gambia (Senghore et al., 2016), cats and dogs in United Kingdom (Harrison et al., 2014), Guinea pigs, rabbits in Germany (Walther et al., 2008).

Studies in Africa have demonstrated that the major clonal lineages circulating in animals are also present in humans (Lozano et al., 2016a, Schaumburg et al., 2015) and this is also a general observation of this study. In addition to using MLST and *spa* typing to investigate the possible sharing of clones between humans and swine in Kiambu, I also looked for evidence of transmission within and across homesteads based on core genome SNPs. Surprisingly, the strains of swine and humans of rare clones ST789 and ST188, yet to be described as global pandemic lineages for hospitals infections, or livestock associated lineages, were highly similar in terms of SNPs even though some were sampled from different homesteads. This suggests that the strains within each of these clones recently shared a common ancestor and spread across homesteads. Alternatively, the clones could be

stable within the environments. However, a previous study of *S. aureus* strains belonging to ST188 from primates, environment and personnel working in a sanctuary center in the USA reported highly similar polymorphisms of seven strains although the isolates were sampled at different times with different antibiotic resistance profiles (Soge et al., 2016). This could indicate this lineage could be having different lower mutation rates compared to the mutation rates of $1.2 \times 10^{-6}$ to $2.0 \times 10^{-6}$ reported for other *S. aureus* lineages (Fitzgerald and Holden, 2016) or $3.3 \times 10^{-6}$ in ST239 lineage (Harris et al., 2010b).

The livestock-associated ST398 lineage has gained a lot of relevance in public health because of its broad host tropisms, and to cause clinical infections. Therefore, the genetic relatedness of ST398 and its double variant ST580 is of great relevance in epidemiological surveillance in the light of one health concept. To the best of my knowledge, this is the first documentation to use whole genome sequencing data to infer relatedness of these two lineages. The earliest documentation of ST580 in MLST database (http://saureus.mlst.net/), is DCC1185 which was methicillin susceptible and was isolated from human colonization at Lisbon, Portugal in 1997. This current study demonstrated that ST580 (one human and two swine) strains were distantly related to the ST398 reference separated by large number of SNPs with long deep branch in the phylogeny (Fig. 20). Notably, they were closer to the strains of basal human associated clade in ST398 phylogeny suggesting that these strains of ST580 belong to human adapted lineages. Previous studies have associated mobile genetic elements found in the accessory genomes of *S. aureus* with the hosts adaptations (Lindsay, 2010). To further validate this hypothesis, the accessory genomes of strains in ST580 and ST398 lineages were compared. Interestingly, strains of ST580 co-segregated together with isolates of human-adapted ST398 lineage (Fig. 21) based on the distribution of the accessory genes, thus supporting the importance of mobile genetic elements in hosts adaptations.

This study has several limitations. First, the elimination of 71 genomes from the analysis mainly due to contamination impaired the analysis of transmission of *S. aureus* between humans and swine within and across homesteads. Even in the absence of contamination, the selection of only a small number of swine and farmers for sampling in each homestead, and the isolation of a single *S. aureus* colony per sampled subject (thus ignoring within-host diversity) could have resulted in uncertainties in the inference of possible transmission

events between subjects/hosts. Furthermore, lack of information such as sampling date for the isolates, and global positioning system (GPS) location of the homesteads as a result of ethical inconsideration at the time of sampling, could have been useful in suggesting recent transmission events of the strains with similar polymorphisms. Secondly, the cross-sectional design of this study relied on assumptions that the farmers and the pigs were persistent carriers of the pathogen. A longitudinal study of *S. aureus* involving deep sequencing of samples, multiple sampling sites of the host, sampling of other livestock in the farm as well as the environments, and screening of pathogen in the hospitals would have constituted an ideal scenario. However, this is not always possible due to operational and budget constraints. Thirdly, the public genomes that were included in the study were skewed towards the human host and this significantly reduced the power of investigating antibiotic resistance genes that could be associated with a particular host, in addition to examining virulence factors that could be essential in hosts adaptations. Finally, clinical isolates that were sampled in Kiambu county hospital (where most farmers get treatment) between 2013 and 2016 were excluded from the whole-genome sequencing project, and their inclusion could have been of great importance in understanding genetic relatedness of clonal lineages that are present in healthy carriers in the community with those of hospital infections.