CHAPTER FIVE

CONCLUSIONS

To the best of my knowledge, this study is the first to characterize at whole genomic level the strains of *S. aureus* between humans and pigs in Africa. The quality control analysis demonstrated the usefulness of whole genome sequencing (WGS) to identify contaminants from the genomes and mixed infections which could have been missed out using conventional typing methods and other laboratory procedures. The results of *in silico* prediction of *spa* types and MLST revealed higher genetic diversity of *S. aureus* in Kiambu County consistent with previous epidemiological findings in Kenya and Africa in general. Phylogeny analyses showed that some isolates of the same *spa* type and MLST were distantly related and were separated by a large number of SNPs demonstrating usefulness of WGS in providing high accuracy and resolution in epidemiological surveillance studies. The observation that humans and pigs share some of the clones and, in some cases, the genomes share similar polymorphisms, justifies further epidemiological investigations to establish the source and the direction of transmission of *S. aureus* strains. The presence of enterotoxin *sec* and human immune evasions gene clusters (IECs) in majority of the strains, even in swine isolates, suggests these strains were recently acquired from humans and were yet to get adapted in pigs. The analyses of enterotoxin gene cluster (*egc*) and Panton-Valentine leucocidin (PVL) genes suggested that they could be associated with ST25 and ST152 lineages respectively. Investigation of Kenyan isolates with global public genomes showed that they were closely related with strains from other African countries suggesting introductions of strains to and/or from Kenya to other African countries. Despite the small datasets analyzed in this study, it provided the basis for further genomic epidemiological surveillance studies and advancement of one health initiative.

5.1 FUTURE PERSPECTIVES OF THE STUDY

5.1.1 High Resolution and Accuracy of Whole Genome Sequencing (WGS)

The quality control analyses of this study provided evidence of the usefulness of WGS to identify presence and proportions of contaminants in the sequence data. In addition, it displayed mixed infections of *S. aureus* strains in some of the genomes. These factors negatively impacted on initial research questions and aims of the study project. However, these could be avoided in future following established purification protocols, choice of selective media for species identification and working under containment laboratory rooms.

5.1.2 Greater Genetic Diversity of Clonal Lineages

Prediction of 9 sequence types (STs) among only 23 strains suggested greater genetic diversity in terms of clonal lineages circulating in Kiambu region. These STs range from specific geographical region clonal lineages such as ST188 that is commonly isolated in Pan-Asian region to globally pandemic lineages such as ST22, ST97 and ST25. However, these genomes were only from small subsets from the collections of samples that were collected in a particular region of approximately 2500 square kilometers (km^2), as opposed to a wider geographic region of Kenya of about 600,000 km^2. Moreover, I did not include clinical isolates to investigate whether these lineages could extend to cause infections in hospitals. Improving power of statistical significance of clonal lineage diversity require wider representation coverage area of the study which will useful in future.

5.1.3 Sharing of Clonal Lineages Between Pigs and Humans

Even though *spa* type and MLST revealed sharing of clonal lineages between humans and pigs, coupled with inter-spread of swine and human strains across the phylogeny and some with similar polymorphisms, we could not be confident enough to suggest possible transmission events. This is because of limited datasets representing small proportions of the cross-sectional study design. However, designing a longitudinal study that covers wider geographic regions and carrying out deep environmental sampling could be ideal in inferring transmission events and evolution of clones in Kenya. This study demonstrated the existence of more than one *S. aureus* strain in a single host therefore sampling multiple sites and deep

sequencing of the colonies of each host are warranted in the future to understand within host diversity. I also realized that epidemiological information of the host such as GPS locations of the farm and the date of isolation, that was lacking in this study could be important in answering transmission questions and estimating the mutation rates of some of the lineages.

5.1.4 Prediction of Antimicrobial Resistance Genes using WGS

Previous studies have demonstrated the usefulness of whole genome sequencing for the prediction of presence of antimicrobial resistance (AMR) genes and the phenotypic expression in routine laboratory antibiotic susceptibility testing (AST) of *S. aureus* strains (Gordon et al., 2014, Aanensen et al., 2016a). Their findings showed WGS could in future replace laborious procedures involved in AST determination in the laboratory although may not be feasible in low- and middle-income countries as of now because of limitations such as expertise knowledge and cost that accompany WGS. Increasing the number of strains sequenced as well as ensuring that the same colonies are used for both AST and DNA extraction and sequencing will sufficiently increase the power of the concordance analysis between *in silico* prediction of AMR and AST. Furthermore, it will be useful in reducing knowledge gaps of the effect of SNPs in AMR genes especially from *S. aureus* strains collected from tropical climatic environments as well as understanding the associations of certain AMR genes profiles with a particular host.

5.1.5 Genetic Relatedness of Kenyan Isolates with Strains of other African Countries

Phylogenetic analyses of human and swine strains of this study with global publicly genomes revealed close genetic relatedness of Kenyan isolates with genomes of other African countries. It will be interesting in the future to carry out epidemiological studies of *S. aureus* in Sub-Saharan Africa that involve large scale genomic analyses of strains sampled longitudinally from diverse countries. This will help in establishing the directions of movement of *S. aureus* strains between countries as well as determining clonal lineages that could be dominant in some regions and temporal signals investigation of certain genes located in mobile genetic elements for adaptations of those dominant lineages.