BIUST Research and Innovation Symposium 2017 (RDAIS 2017)
Botswana International University of Science and Technology
Palapye, Botswana, 12-14 June 2017

ISSN: 2521-229X
ISBN: 978-99968-0-608-7

Correlative Analysis Of Plasmid Mobility And Antibiotic Resistance Within ColE1 Plasmids

Monamodi. L. Kesamang
Department of Biology and Biotechnological sciences,
College of Science.
Botswana International University of Science and Technology, Palapye, Botswana.
Email: monamodi.kesamang@studentmail.biust.ac.bw

Teddie. O. Rahube
Department of Biology and Biotechnological sciences,
College of Science.
Botswana International University of Science and Technology, Palapye, Botswana.
Email: rahubet@biust.ac.bw

Abstract*— Antibiotic resistance in bacteria remains a major global health challenge in the control of infectious diseases. Antibiotic resistance genes are frequently transferred across bacterial communities by mobile genetic elements such as plasmids. The majority of plasmids possess genes that allow them to replicate and transfer efficiently in various bacterial hosts. The plasmids in the ColE1 family have a series of mobility genes (mbeA, mbeB, mbeC, mbeD, and mbeE), which makes the plasmid mobilizable. However, the relationship between this mobilizable plasmid group and antibiotic resistance genes is not well understood. The mbeC gene, is the biggest and most important of the mobility genes of the ColE1 plasmids, this gene has been shown to display variation of occurrence (presence or absence) in antibiotic resistance plasmids and non-antibiotic resistance plasmids of the same family of plasmids. Using a bioinformatics approach, the differences in DNA sequences available in various online databases can be analysed by comparing the several genes associated with plasmid mobility. The innovative aspect of this research is that it provides the opportunity for development of robust molecular detection tools for tracking of specific types of plasmids with antibiotic resistance genotypes in clinical and environmental samples. This bioinformatics analysis approach is applicable anywhere and particularly important to low income countries as it does not require one to conduct in vivo experiments such as DNA sequencing for the study of plasmids or other genetic elements.

Key words - mbeC-gene, ColE1, Plasmid mobility, Antibiotic resistance, Bioinformatics Introduction

I. Background

Antibiotic resistance genes can be expected to have a correlative relationship with plasmid mobility on the basis that they display a high selective pressure when the environment demands it. In such an instance, plasmids will possibly have a high occurrence of genes that will allow them to spread efficiently by horizontal gene transfer. However in ColE1 plasmids that are mobilizable, antibiotic resistance is surprisingly found to have a negative correlation with plasmid mobility as confirmed by comparative sequence analysis methods. The reasons for this negative correlation is subject for further research, however it is possible that since ColE1 plasmids are small, they may not have tolerance for maintaining the coexistence of many genes that will make them larger.

A. Plasmid Mobility

Conjugation is one of the mechanisms through which horizontal gene transfer occurs in bacteria, and it is probably the most relevant. The appearance of multi-drug resistant bacteria brought horizontal gene transfer to the attention of scientists worldwide; this is because the dissemination of antibiotic resistance genes by conjugative plasmids poses serious health problems. ColE1 is mobilized by a wide range of conjugative plasmids and is the example of a family that encompasses mobilizable plasmids found in gram-negative and gram-positive bacteria [1].

Plasmids can be classified into three categories with regards to mobility. These are conjugative, mobilizable, and non-mobilizable plasmids. The distinction between these type plasmids is the presence or absence of the type of genes that allow the conjugation or mobility to occur. In a scheme described by Smillie and colleagues (Fig 1), there are three types of genes and one recognition site that allow plasmid mobility to occur by conjugation, [2]. The three types of genes are a relaxase protein, a type 4 coupling protein (T4CP) and a type 4 secretion system (T4SS). One of these 3 genes, the relaxase protein is associated with a recognition site called the origin of transfer (OriT).

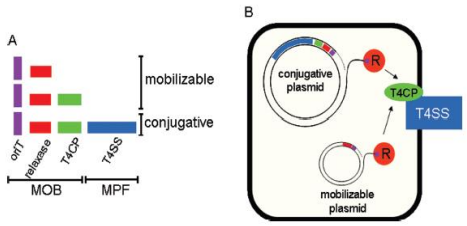


Fig 1. (A) A schematic interpretation of the genetic constitution of transmissible ColE1 like plasmids. (B) A representation of the functions that each gene has during the process of conjugation [2].

The *OriT* is recognized by a relaxase protein which nicks (forms a single stranded break) the plasmid at the particular location. The relaxase protein then attaches to the 3'end of the

ISSN: 2521-229X
ISBN: 978-99968-0-608-7

DNA and unwinds it, it then develops affinity to the type 4 coupling protein (T4CP) which has been coupled to the type 4 secretion system (T4SS). TheT4SS is responsible for extending the fertility pilus between two bacterial cells [2].

B. Mobility in ColE1 plasmids

ColE1 are a family of plasmids that have been described by the circumstance that the earliest of these plasmids was found to carry a gene for colicin E1 also called the *cea* gene [3]. These plasmids are not conjugative, that is they cannot move from one cell to another independently, rather they are mobilizable. They only move from one cell to another if there is a conjugative plasmid within the host cell. ColE1 specifically has been reported to be mobilized by conjugative members of various incompatibility groups, including IncIa, IncFI, IncW and IncP. There are five genes of mobilization in ColE1 plasmids, these are (*mbeA*, *mbeB*, *mbeC*, *mbeD*, and *mbeE*). Two of them (*mbeB* and *mbeD*) are entirely overlapping *mbeA*, seen in the schematic;(Fig2).

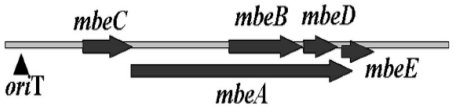


Fig 2. The mobilization region of ColE1 plasmids [1].

The genes *mbeA* and *mbeC* are found conserved in all ColE1 family plasmids whereas the genes *mbeB*, *mbeD*, and *mbeE* are not conserved. From this standing, researchers denote that the common ColE1 ancestral plasmid must have had a mobilization region containing only the *mbeA* and *mbeC* genes [1]. These genes are the most important where mobility is concerned because *mbeC* serves as the functional component of the nicking accessory proteins in the MOBP5 family (previously MOB_{HEN}) that are associated with mobilizable ColE1 plasmids and *mbeA* functions as the central component of the 'relaxation complex' [1], [2]. Initially, in around the 1970s, only 3 relaxase genes had been described for ColE1 plasmids, these were *mbeA*, *mbeB* and *mbeC*, with subsequent Genetic research: [4], [5], [6] 2 more, *mbeD* and *mbeE* were characterised. Of these *mbeE* was considered non-essential.

C. Antibiotic Resistance

Antibiotic resistance, in the general sense refers to the ability of bacteria to evade the bactericidal effects that antibiotics have against them. However, different studies have adopted different definitions for the term "Antibiotic resistance" depending on the objectives of each study. Martinez and colleagues described three general definitions for which the term can be characterized, these are; the clinical definition, which explains it as the MIC (minimum inhibitory concentration) breakpoints denoting therapeutic failure in human patients of a given antibiotic concentration. The epidemiological definition, defines it as the MIC value that corresponds to the upper limit of the wild-type population of a given species. The operational definition defines resistance as the pairwise comparison of a non-resistant strain against a mutant strain that has acquired resistance through HGT [7]. Characteristics of resistance are transmitted through bacterial populations through genes. Antibiotic resistant genes (ARG) have been around along with the prevalence of antibiotics in nature. Subsequent research has since the discovery of the first antibiotics witnessed the documentation of a collection of ARGs from community of bacteria in various environments that are now referred to as the resistome [8], [9]. Antibiotic resistance genes are found in plasmids and this occurs by mechanisms of horizontal gene transfer, these genes transfer from one environment to another depending on the selective pressure. When the selective pressure is high the genetic mechanisms of horizontal gene transfer are elevated including plasmid mobility [10].

II. MATERIALS AND METHODS

1. Data mining

The plasmid sequences are publicly available as part of the NCBI database resources. They were downloaded as a compressed tape-archive gunzip (tar.gz) file from the NCBI ftp directory, index of; (ftp://ftp.ncbi.nlm.nih.gov/genomes/Plasmids/). From this link the plasmids were contained in the file named "plasmids.all.fna.tar.gz". For this research, the file was retrieved in November 15 2015. It comprised a total of 6,090 plasmid sequences in fna (fasta nucleic acid) format and a sequence identification file "Plasmid.ids". The identification file contained a list of 6,197 plasmids, this was 107 more than the total number fna plasmid sequences the compressed file actually contained. The lists pecified the plasmid host bacteria and corresponsive NCBI accession and gi numbers. The above link is out-dated and has since been updated to be accessible at the index of; (ftp://ftp.ncbi.nlm.nih.gov/refseq/release/plasmid/) last accessed October 5 2016. From this new link the plasmids are contained in the file named "plasmid.1.1.genomic.fna.gz" and "plasmid.2.1.genomic.fna.gz" respectively.

2. Preliminary data sorting

Plasmids were sorted to remove irrelevant data to the research. From this sorting procedure, they were distributed amongst the three domains of life, in which those from domain bacteria were selected, these constituted (97%) 5951 plasmids. They were then sorted according to whether the plasmid was submitted into the database completely or incompletely sequenced, these constituted (95%) 5598 plasmids. From the completely sequenced plasmids, plasmids that contained antibiotic resistance genes were selected from the completely sequenced plasmids, these constituted (24.7%) 1406 plasmids. Lastly they were sorted according to whether they were circular or linear plasmids, for this research only circular

ISSN: 2521-229X
ISBN: 978-99968-0-608-7

plasmids were required, these constituted (96%) 1347 circular resistance plasmids.

3. Plasmid Analyses

a) Analysis of ColE1 plasmids

ColE1 plasmids have a sequence specific replicon that has been described by García-Fernández and colleagues 2009. This replicon is a generic sequence that extracts all ColE1 plasmids within a given population of any size using the recently developed plasmid based replicon typing (PBRT) technique, as in the case of this research. For this research an amplicon derived *in-silico* from the ColE generic sequence (accession number AM746977) using replicon typing primers oricolE FW (5'-GTT CGT GCA TAC AGT CCA-3') and oricolE RV (5'-GGC GAA ACC CGA CAG GAC T-3'), [11]. This amplicon was used in the Meg-Align pro utility of the DNA-star software using the mauve alignment tool. All positive hit alignments with a consensus sequence determined by default settings of the mauve alignment tool were considered ColE1 plasmids. These were selected from the population of 1347 circular resistance plasmids and 4276 non-resistance plasmids. From this analysis approach, 83 plasmids out of the 1347 plasmids were found to have the ColE replicon within the population of resistance plasmids, and 142 plasmids out of the 4276 plasmids were found to have the ColE replicon within the population of non-resistance plasmids. A representative over-view of plasmids with the ColE1 replicon as a fraction of all (resistant and non-resistant) plasmids is indicated in (Chart 1.1).

b) Analysis of the mbeC mobility gene within ColE like plasmids

The *mbeC* gene is a part of an operon of 4 genes (*mbeC*, mbeA, mbeB, and mbeD) that form the MOB_{P5}. The operon is available on NCBI (accession number X15873.1). It is the only gene that does not overlap with other genes in this operon, it is also considered along with mbeA the most conserved since they are found in all mobilizable ColE family of plasmids. Within the operon, the *mbeC* open reading frame is found at nucleotides position 97-420. This region was extracted within the 83 circular resistant ColE1 plasmids and the 142 circular non-resistant ColE1 plasmids. From the 83 circular resistant ColE1 plasmids 4 contained the *mbeC* gene, (Chart 1.5) and from the 142 circular non-resistant ColE1 plasmids 72 contained the *mbeC* gene, (Chart 1.4).

c) Phylogenetic tree analysis *of mbeC genes in resistance and non-resistance genes*

Multi-sequence alignments (MSA) of the *mbeC* genes found in non- resistance, resistance and all CoE1 plasmids were created on the DNA star software (Lasergene 12.0) using the Clustal Omega function. It was made to determine the diversity of *mbeC* amongst non- resistance and resistance plasmids and determine if there was a notable difference considering lineage associations. Phylogenetic trees were made to map these relationships. A comparative protein multi-sequence alignment of resistant, non-resistant and all ColE1 plasmid *mbeC* gene products were generated to compare with the nucleic acid alignments, trees were generated from these alignments as well. This was done by first using an online tool to reverse transcribe all the inverted sequences so that all sequences read from the leading strand (in the 5' to 3'direction), online, available at; <http://www.bioinformatics.org/sms/rev_comp.html>. The sequences were then translated to amino acid sequences using the EMBL-EBI nucleotide sequence translation tool; Transeq (EMBOSS), online, available at; <https://www.ebi.ac.uk/Tools/st/emboss_transeq/>. These sequences were aligned on the MegAlign Pro tool on the DNA star software using the Clustal Omega function. All trees were used to infer lineages were any were relevant to the aim of this study.

d) Correlative analysis for plasmid resistance and mbeC gene presence by comparative genomics

Plasmids were annotated to determine the arrangement that the mobility gene *mbeC* has within the general structure of the plasmid relative to other genes in the vicinity; synteny. This was done for both resistant and non-resistant plasmids with the ColE1 replicon. Antibiotic resistance genes were then depicted relative to the *mbeC* gene and other associative genes in the vicinity of the *mbeC* for resistance plasmids. This exercise was done solely to determine any form of structural interference or other mutation events amongst resistance genes and mobility genes. This was done using subsequent pairwise alignments using the mauve function of the DNAstar software against all plasmid sequences sequentially. The alignments were then graphically mapped to represent the evolutionary events (from a single locus) that all the plasmids have in common; the ColE1 replicon.

Along with the DNA star software mauve alignment function, an online tool (Pathosystems Resource Integration Center) PATRIC 3.3.38beta / RAST (Rapid Annotations using subsystems technology) was used to annotate resistance genes collectively for both resistance and non-resistance CoE1 plasmids. Online available at: < https://www.beta.patricbrc.org/>. Annotations were made under parameters: {Taxonomy name; [superkingdom] Bacteria, Taxonomy ID; [2]}. Annotations compared resistance and non-resistance ColE1 plasmids collectively in the circular viewer tool.

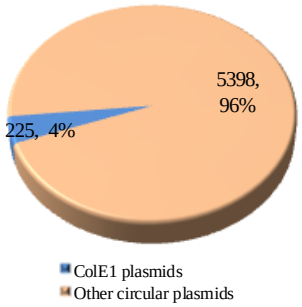
ISSN: 2521-229X
ISBN: 978-99968-0-608-7

III. Results

1. *Preliminary data sorting*

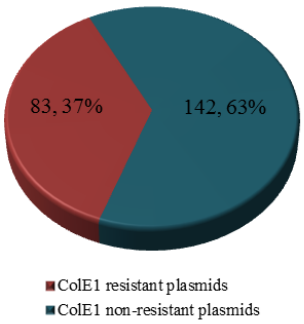
a) Determination of plasmids with the ColE1 replicon.

Chart 1.1. Plasmids with ColE1 replicon with other plasmids



b) Analysis of ColE1 plasmids in resistant and non-resistant plasmids

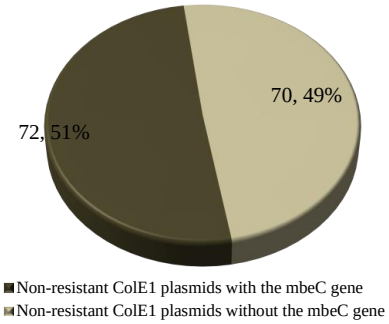
Chart 1.2. Resistant ColE1 plasmids within non-resistant ColE1 plasmids



2. *Analysis of the mbeC mobility gene within (resistant and non-resistant) ColE1 plasmids*

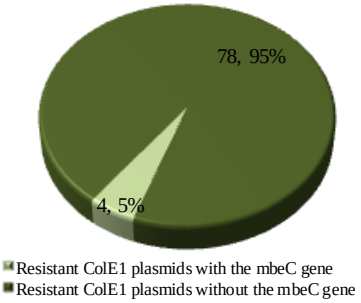
a) Non-resistant ColE1 plasmids

Chart 1.3. Non-resistant ColE1 plasmids with the mbeC gene against those without the mbeC gene



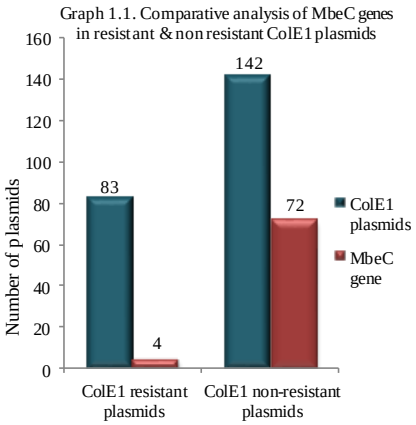
b) Resistant ColE1 plasmids

Chart 1.4. Resistant ColE1 plasmids with the mbeC gene against those without the mbeC gene



3. *Total collection of MbeC mobility genes in resistant ColE1 like plasmids compared to non- resistant ColE1 like plasmids.*

Graph 1.1. Comparative analysis of MbeC genes in resistant & non resistant ColE1 plasmids



BIUST Research and Innovation Symposium 2017 (RDAIS 2017)
Botswana International University of Science and Technology
Palapye, Botswana, 12-14 June 2017
ISSN: 2521-229X
ISBN: 978-99968-0-608-7

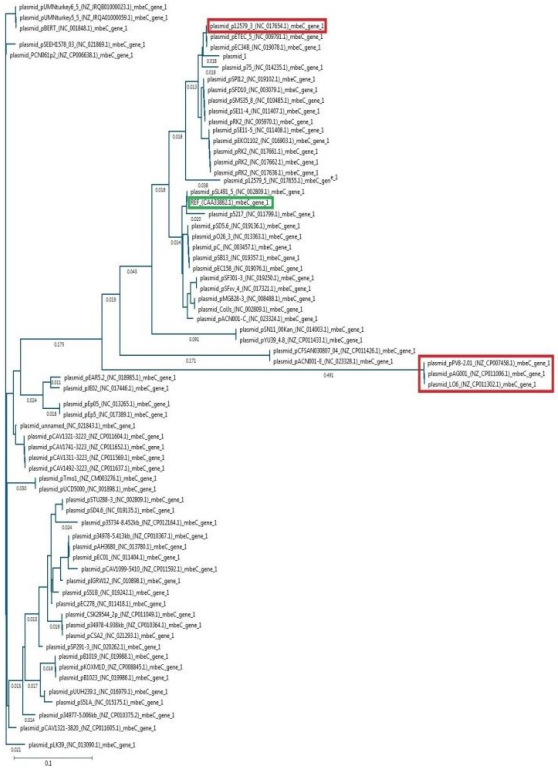


Fig 3. Phylogenetic tree analysis of *mbeC* genes in both resistant and non-resistant ColE1 plasmids, protein sequence phylogenetic tree * Red boxes indicate *mbeC* proteins from resistant plasmids, The Green box indicates the reference outgroup *mbeC* gene.

4. Annotation of antibiotic resistance in ColE1 plasmids.

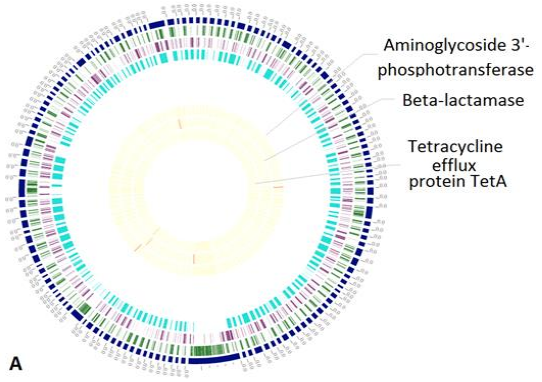


Fig 4. PATRIC 3.2.38beta RAST comprehensive antibiotic resistance annotation, circular viewer. A) CDS annotation of 6 loosely associated genes (These genes are not functional) by CARD on 142 **non-resistant** ColE1 plasmids.

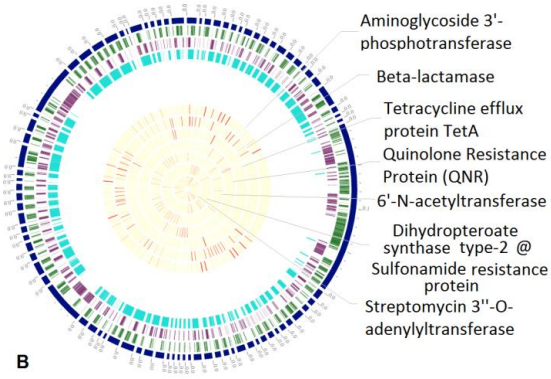


Fig 5. PATRIC 3.2.38beta RAST comprehensive antibiotic resistance annotation, circular viewer. B) CDS annotation of 178 strictly associated genes by CARD on 83 **resistant** ColE1 plasmids.

*NB: (Fig 4 & Fig 5) red slits along a yellow background ring show an annotated antibiotic resistance gene ARG [each ring represent a class of ARG as labelled]. Dark blue (outer) ring = complete plasmid DNA sequences, Dark green slits = forward strand annotated genes/ORF, Purple slits = reverse strand annotated genes/ORF. Light blue ring = non-coding sequences.

IV. DISCUSSION

The correlation of plasmid mobility to antibiotic resistance in ColE1 plasmids set out to be determined by this study was expected to be positive but it has proven otherwise. This disproved assumption was on account of ARGs usually being expected to have a positive selection pressure that would as a consequence incline them to occur in mobilizable plasmids as a matter of accumulation by random chances through mechanisms of MGEs. However this study has shown that the trend followed is actually contrary to that expected. Using the essential MOB_{P5} mobility gene associated with ColE1 mobile plasmids; the *mbeC* gene, the chart 1.3, chart 1.4 and graph 1.1 shows that there is more of the *mbeC* gene found in non-resistant ColE1 plasmids (51%) than in resistant ColE1 plasmids (4%). Upon interrogating the cause of this circumstance by phylogenetic analyses of the gene to infer a possible fitness lineage preference, the results were inconclusive due to fewer *mbeC* genes extracted from resistance plasmids. By ratio, this is 4 genes from resistance plasmids to 72 from non-resistant plasmids. Nonetheless, from these results it was determined that 3 of the 4 *mbeC* genes in resistance plasmids are closely related to one another with the remaining being more divergent from the rest (MSA from the *MbeC* gene ORF amino acid sequences), Phylogenetic tree Fig 3. These results somehow point out the probable fact that the inheritance of resistance genes is not secluded to certain variants of *mbeC*, it is instead random. Despite this, the *mbeC* outgroup lineage (Fig 3., red box) seen in the resistant ColE1 plasmids may be indicative of the possibility that *MbeC* genes may be getting adapted to coexist with resistance genes within

ISSN: 2521-229X
ISBN: 978-99968-0-608-7

the plasmid. In any case, the DNA sequence discrepancies observed are the source of the molecular markers that this research aims to highlight. This is to have a potential for development into molecular detection tools in a subsequent applied research. These markers will be useful for tracking of specific types of ColE1 plasmids with antibiotic resistance genotypes in clinical and environmental samples.

As for the diversity of the *mbeC* genes themselves, it is observed that there are more diverse, since they have developed 3 lineage clusters. This as indicative from the phylogenetic trees generated of the non-resistance *mbeC* counterpart genes. There are 3 different clusters deduced from the tree of non-resistance *mbeC* genes in ColE1 plasmids. The first cluster (top in Fig 3) is stagnant or occurs in a few plasmids and is more closely related to the MOB_p family of relaxes that intersect with the ColE1- relaxase super-family (MOB_{P5}). This cluster somehow links both super-families to a common origin, [1]. The second cluster (bottom in Fig 3) is an intermediate lineage of ColE1 *mbeC* genes whilst the third cluster (middle in Fig 3) is a more diversified lineage of ColE1 *mbeC* genes that are indicative of true ColE1 plasmids. According to the distance scale, this cluster has a branch length of more than 0.3 from both the first and second cluster. A slight discrepancy may occur between the protein and DNA sequences lineages but the general representation is matching, as both DNA phylogenetic trees (not shown) and protein phylogenetic trees represent three clusters. When comparing *mbeC* genes of both resistance and non-resistance plasmids, genes are allocated randomly within different clusters and no conclusive judgement can be validated about selective preference of resistance genes for mobile genes. This is due in part because there are fewer *mbeC* genes for plasmids with resistance; results in chart 1.3, chart 1.4 and graph 1.1.

The comparative genomics approach to determine the relevance of *mbeC* with antibiotic resistance was useful in showing relative comprehensive antibiotic presence in the two groups of ColE1plasmids(resistant & non-resistant). Most of the *mbeC* genes were randomly distributed along the plasmid genomes as deduced by mauve alignments. As expected, the genome annotation by PATRIC 3.2.38beta RAST showed greater antibiotic resistance genes in resistant ColE1 plasmids than in non-resistant ColE1 plasmids (Fig 4 & Fig 5). The 6 loosely associated genes observed in the CDS annotation circular viewer, Fig 4 are not functional from the CARD analysis report. The correlation between the two datasets (resistant & non-resistant ColE1plasmids) based on synteny was not clearly represented by both software due to large data processing limitations, and further analyses are yet to be established and described in other subsequent studies.

ACKNOWLEDGMENT

This work was conducted as part of the Bioinformatics Graduate Project initiated by TOR. We acknowledge Botswana International University of Science and Technology for the postgraduate funding to MK.

REFERENCES

[1] Francia, M. V., A. Varsaki, M. P.,Garcilla´n-Barcia, A.,Latorre, C. D., and F. de la Cruz. (2004). A classification scheme for mobilization regions of bacterial plasmids. FEMS Microbiol. Rev. 28:79–100.

[2] Smillie, C.,Garcilla´n-Barcia, M. P.,Francia, M.V., Rocha, E.P.C and De la Cruz, F. (2010). Mobility of Plasmids. Micro& Mol. Bio Rev. 3: 434 – 452

[3] Hershfield, V., Boyer, H.W., Yanofsky,C., Lovett,M. A. and Helinski, D. R. (1974). Plasmid ColE1 as a Molecular Vehicle for Cloning and Amplification of DNA. Proceedings of the National Academy of Sciences. 71 (9): 3455–59

[4] Blair, D.G. and Helinski, D.R. (1975) Relaxation complexes of plasmid DNA and protein. I. Strand-specific association of protein and DNA in the relaxed complexes of plasmids ColE1and ColE2. J. Biol. Chem. 250, 8785–8789.

[5] Lovett, M.A. and Helinski, D.R. (1975) Relaxation complexes of plasmid DNA and protein. II. Characterization of the proteins associated with the unrelaxed and relaxed complexes of plasmid ColE1. J. Biol. Chem. 250, 8790–8795.

[6] Boyd, A.C., Archer, J.A. and Sherratt, D.J. (1989) Characterization of the ColE1 mobilization region and its protein products. Mol. Gen. Genet. 217, 488–498.

[7] Martinez , J. L., Coque, T. M., and Baquero, F., (2015). What is a resistance gene? Ranking risk in resistomes Nature Reviews Microbiology 13, 116–123.

[8] D'costa, V.M., Mcgrann, K.M., Hughes, D.W., & Wright, G.D. (2006) Sampling the antibiotic resistome. Science 311, 374-377

[9] Wright, G.D. (2007).The antibiotic resistome: The nexus of chemical and genetic diversity. Nature Reviews Microbiology 5, 175-186.

[10] Bennett, P.M. (2008) Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. British Journal of Pharmacology 153, S347-S357.

[11] García-Fernández , A., Fortini, D., Veldman, K., Mevius, D., Carattoli, A.,(2009). Characterization of plasmids harbouring qnrS1, qnrB2 and qnrB19 genes in Salmonella. Journal of Antimicrobial Chemotherapy 63, 274–281