5-8-2020

Whole genome sequence analysis of a transmissible multidrug-resistance plasmid captured without cultivation from poultry litter

Emma C. Eisemann
James Madison University

Whole genome sequence analysis of a transmissible multidrug-resistance plasmid captured without cultivation from poultry litter

An Honors College Project Presented to

the Faculty of the Undergraduate

College of Science and Math

James Madison University

by Emma Carolyn Eisemann

Spring 2020

FACULTY COMMITTEE:

Project Advisor: Dr. James B. Herrick, Ph.D.
Reader: Dr. Kyle Seifert, Ph.D.
Reader: Dr. Steven G. Cresawn, Ph.D.

Acknowledgements

The completion of this Honors thesis would have been impossible without the advising and support from Dr. James Herrick. Dr. Herrick has mentored me as a student and, throughout my two and a half years in his research lab, he has tremendously helped me mature as a scientist. I have also had incredible support from past and current peers in the Herrick lab. Specifically Selena Hise & Liam Godbold, who helped with hours of data collection. I'm grateful for the Herrick Lab graduate students Curtis Kapsak, Noah Greenman, and Max Maza, for their willingness to answer my many questions and lend a helping hand. Specifically, Noah Greenman for his sequencing expertise and assistance. In addition I would like to thank the JMU Honors College and my Capstone Committee, Dr. Kyle Seifert and Dr. Steve Cresawn, for teaching me foundational coursework as well as providing feedback on this project from the developing stages. Lastly, a large thank you to my family and friends for supporting me through the completion of this project and providing moral support. I will forever cherish my time at JMU and in the Herrick Lab.

Table of Contents

List of Tables

List of Figures

Abstract

Use of antibiotics in the agricultural industry introduces selective pressure and, consequently, could increase the presence of antibiotic resistant organisms in surrounding environments. One such environment is litter (manure and bedding) produced during large-scale poultry production in the Shenandoah Valley. Litter, with its microorganisms, is commonly applied to fields within the Shenandoah River watershed. Antibiotic resistance (AR) and virulence genes are potentially transmissible between organisms through horizontal gene transfer of genetic mobile elements, for which poultry litter could be a reservoir. The typical, culture-based approach to detecting and analyzing AR plasmids and other mobile genetic elements is limited due to the inability to culture, isolate, and analyze all bacteria in nearly all environments. In addition, the expense and time of extracting and sequencing plasmids from culturable isolates is great. The goals of this study were (i) to use a non-culture-dependent plasmid isolation method to isolate AR plasmids directly from poultry litter, (ii) to sequence and assemble the whole genome of the plasmid capture strain *E. coli* LA61^{RifR}, and (iii) use a combination of short- and long-read sequencing and computational methods to assemble and annotate one of the captured plasmids. It was also wished to determine the antibiotic susceptibility of the captured plasmids. An exogenous plasmid capture method was used to isolate tetracycline-resistance plasmids EH1-12, some of which conferred phenotypic resistance to a range of late-generation, clinically-significant antibiotics. Of the 12 transconjugants, 11 conferred resistance to more than one antibiotic (excluding tetracycline), the most common were resistances to piperacillin and piperacillin/tazobactam. Perhaps most striking was the resistance conferred by plasmid EH11 to aztreonam, a monobactam antibiotic effective against gram negative aerobic organisms, which has rarely been observed. Other surprising resistance phenotypes included ceftazidime and ciproflaxocin which are members of the cephalosporin and quinolone drug classes, respectively. The whole genomes of both the plasmid capture strain LA61^{RifR} and one of the multidrug resistant transconjugants, LA61^{RifR}::pEH11, were sequenced. SPAdes and Canu were used to assemble the genomes of LA61^{RifR} and of LA61^{RifR}::pEH11, respectively. Ninety-seven contigs assembled from short-read sequencing data comprised the LA61^{RifR} genome and 5 contigs assembled from long-read data comprised the LA61^{RifR}::pEH11 genome. One contig of LA61^{RifR}::pEH11was identified as plasmid EH11. Genes encoding antibiotic resistance, bacteriocins, and aerobactin siderophore systems were annotated with ARGannot, RAST, and Prokka . Eight repeat regions, 47 transposase genes, and two regions responsible for plasmid replication and transfer were also identified. Overall this study, through phenotypic and genotypic analyses, demonstrated that poultry litter can act as reservoir for transmissible multidrug-resistant plasmids. Genome analysis also demonstrated the potential to transfer genes that contribute to a host's virulence. Such resistances and virulence genes, encoded on transmissible plasmids, provide advantages to infectious agents and enable their survival in poultry litter and other environments, thus possibly complicating treatment of resulting infections.

Introduction

Background & Significance

Antibiotic resistance is a topic of great concern in clinical settings, where infections are commonly caused by the large reservoir of resistant organisms in hospitals and other health facilities. Many patients also enter healthcare facilities with an antibiotic resistant infection and are in need of effective antibiotics. Consequently, there has been much research on nosocomial antibiotic resistant infections, as well as resistant "superbugs" that directly impact human health (1). Less considered, antibiotics are commonly used in the agriculture industry as growth promoters and to prevent infection (2), and are present in runoff from agriculture fields and poultry farms. According to a study in 2013, approximately 70,000 kg and over 13,500,000 kg of antibiotics were applied annually to crops and livestock, respectively, in the United States (3). These antibiotics may affect the reservoirs of antibiotic resistance genes (ARGs) in native soils and streams. These genes are commonly found on plasmids, genetic mobile elements that can potentially be transmitted across species and even higher taxa via conjugation. Plasmids can carry and transmit antibacterial resistance genes in the Central Shenandoah Valley (4), where the agriculture and poultry industries are prominent. A high concentration of transmissible plasmids conferring antibiotic resistance could indicate a reservoir for these genes in local environments.

Horizontal Gene Transfer

Horizontal gene transfer (HGT) is the exchange of genetic material from one organism to another laterally, without the production of offspring. There are three modes of HGT:

transformation, transduction, and conjugation. Transformation and transduction are the transfer of genetic material to another organism as naked DNA from the environment or using a viral vector, respectively. Conjugation is the transfer of plasmids or conjugative transposons from one bacterium to another by means of a pilus. By this mode of HGT, plasmids, carrying ARGs, can be replicated and given to nearby cells in environmental reservoirs.

Mobile Genetic Elements

A mobile genetic element (MGE) is genetic material that can relocate to another part of a genome or transfer to another organism. These elements may contribute to an organism's virulence because they enable the sharing of ARGs and even virulence genes themselves across species or genera. There are several types of mobile genetic elements. Plasmids are relatively small, extrachromosomal, circular pieces of DNA that bacteria may carry and that can transfer into other bacteria. Plasmids carry "accessory genes", genes that are not directly needed for the organism's survival but contribute to its pathogenicity or ability to survive in a greater variety of environments.

Transposons are intracellular MGEs that allow an organism to move genes from one position in the genome to another using a cut and paste mechanism. This transfer can take place between chromosomal and plasmid DNA, between plasmids, or between different loci in the chromosome of the bacterium. Insertion sequences contain only a transposase gene between two inverted repeats (5). Transposons carry accessory genes, like ARGs, in addition to transposase genes which each remain sandwiched between two inverted repeat sequences (5).

Like transposons and plasmids, integrons can also contribute to an organism's antibiotic resistance (6). Integrons are classified as MGEs, however, they are not actually mobile. Rather,

integrons result in the excision and/or integration of entire gene cassettes across or between genomes (7). Each integron has three essential components: an integrase, attachment/recombination site, and a promoter sequence. The integrase is a tyrosine recombinase enzyme that removes and relocates gene cassettes, the attachment site marks the site of excision or integration of a gene cassette, and the promoter sequence controls the gene expression of the gene cassettes. The expression level of gene cassettes changes as a result of recombination and the promoter sequence to which the cassette is adjacent. Therefore, if an AR gene cassette undergoing recombination is shuffled adjacent to a promoter with greater gene expression, the AR genes will be more frequently transcribed. Plasmid-borne integrons can then be transferred horizontally, contributing to the reservoir of transferable antibiotic resistant genes.

Additional MGEs include integrative conjugative elements (ICEs), conjugative transposons, and genomic islands. ICEs are self-transmissible, chromosomally-integrated MGEs that contain a set of core genes for excision and transmission. ICEs may also contain non-essential, accessory genes which may be beneficial by providing a selective advantage to the host (8). Conjugative transposons, which differ from transposons, have a circularized intermediate that is conjugatable, like plasmids, but not replicable (9). During transfer, one DNA strand of the circular intermediate is linearized and transferred into an adjacent bacteria cell. These elements are most common in gram-positive cocci, however, can also be present in gram-negative bacteria.

Lastly, genomic islands (GIs) are regions of consecutive genes that originated by means of HGT, are distinguishable by their differing GC content relative to the host chromosome, and have a high density of virulence and AR genes (10). These regions can be composed of other MGEs, like integrons, transposons, and ICEs (10). Genomic islands are particularly interesting

because they provide insight into microbial evolution. Genomic islands often contain other MGEs such as transposons, integrons, or ICEs.

Tetracyclines and tetracycline resistance

Tetracyclines are a broad-range class of antibiotics effective against both gram negative and gram positive organisms through prevention of attachment of tRNA to the acceptor (A) site of the ribosome (11). Resistance to tetracyclines is common, and is largely due to plasmid transmitted resistance genes (12). The prevalent use of tetracycline as a growth promoter and to prevent infection in livestock makes it a logical antibiotic to select for in environmental samples. Previously, it has been found that plasmids conferring resistance to tetracycline also confer resistance to additional antibiotics (13).

Exogenous Plasmid Capture

Plasmids in natural environments are typically studied by "endogenous" isolation, in which plasmid DNA is extracted from isolated bacteria. This requires the individual culture of bacterial isolates followed by plasmid DNA extractions. In the "exogenous" or plasmid capture method (13) used in our laboratory, on the other hand, 'donor' or native cells are prepared for conjugation by releasing them (using a sodium pyrophosphate solution) from an environmental sample such as soil or litter, enabling the capture of plasmids from uncultured bacteria. Prior to conjugation, the rifampicin resistant capture strain contains no plasmids. The prepared environmental donor cells are plated on control plates to confirm the presence of cells conferring resistance to tetracycline and that all cells were susceptible to rifampicin. Growth on rifampicin

and tetracycline media indicates that transconjugants are the rifampicin-resistant capture strain that received a tetracycline resistance plasmid (Figure 1).

Previous work in the Herrick lab (13) has focused on transmissible plasmids collected from environmental samples, by capturing plasmids conferring tetracycline resistance into rifampicin-resistant recipient *Pseudomonas* or *E. coli*. Theoretically, the resulting transconjugants are the rifampicin-resistant recipient *E.coli* containing a plasmid that confers tetracycline resistance.

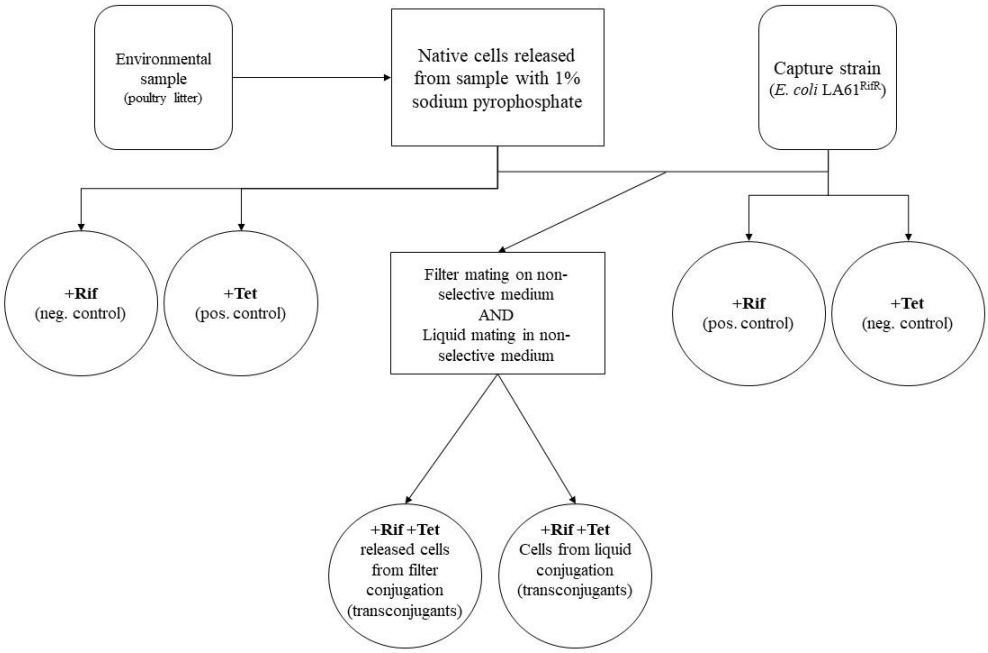

Figure 1. Schematic of exogenous plasmid capture with controls.

Illumina™ and Oxford Nanopore™ Sequencing

The Illumina MiSeq™ is a short read DNA sequencer that produces raw reads with a low error rate. The Oxford Nanopore MinION™ is a portable, handheld DNA sequencer that generates reads within hours, a fraction of the time required by Illumina short read DNA sequencers (14). Read lengths are also much longer (typically greater than 10 kb vs 150 to 250

bp from Illumina). Despite its increased error rate, the MinION is the fastest and most cost effective option for the identification of specific genes, such as those encoding antibiotic resistance or virulence.

De novo Plasmid Genome Assembly

De novo assemblers take the reads produced by sequencers and generate an assembly without a reference genome. The fewer contigs there are in the final assembly, the better the assembly quality is considered to be (15). A perfect chromosomal or plasmid genome assembly would result in a single, circular contig. There are many considerations when assembling plasmid genomes from raw whole genome sequencing reads. The first consideration is the trade-off between short and long read data. Short read data have a lower error rate, however, the shorter reads, when overlapped, produce many contigs. Most genomes from short-read data are actually "draft" genomes. Longer reads produce fewer contigs; however, the reads have a higher error rate. This trade-off can be ameliorated by producing a hybrid assembly using both types of reads in a hybrid assembler such as Unicycler (16). However, a hybrid assembly involves sequencing using two separate platforms, which is less cost and time effective than sequencing using only one method.

Methods

Sample Collection & Selection

Strains used

LA61^{RifR} is a strain of *E. coli* isolated by Dr. Elizabeth Alm from a beach sand sample taken on June 9, 2002 at Lakeport Campground at Lake Huron. The sample was taken at a depth of 11-15 cm. Strain LA61^{RifR} is plasmid-free, tetracycline-sensitive, and has been successfully used in a number of plasmid capture experiments in the Herrick laboratory (4, 13, 17).

The original rifampicin-sensitive wild strain was manipulated using a gradient TSA+rifampicin plate to select for rifampicin resistance (Figure 2) (18). Of note, all strains, unless otherwise noted, were cultivated at 37°C for 24 hours. All H_2O used was treated with a MilliQ water filtration system and autoclaved for sterilization.

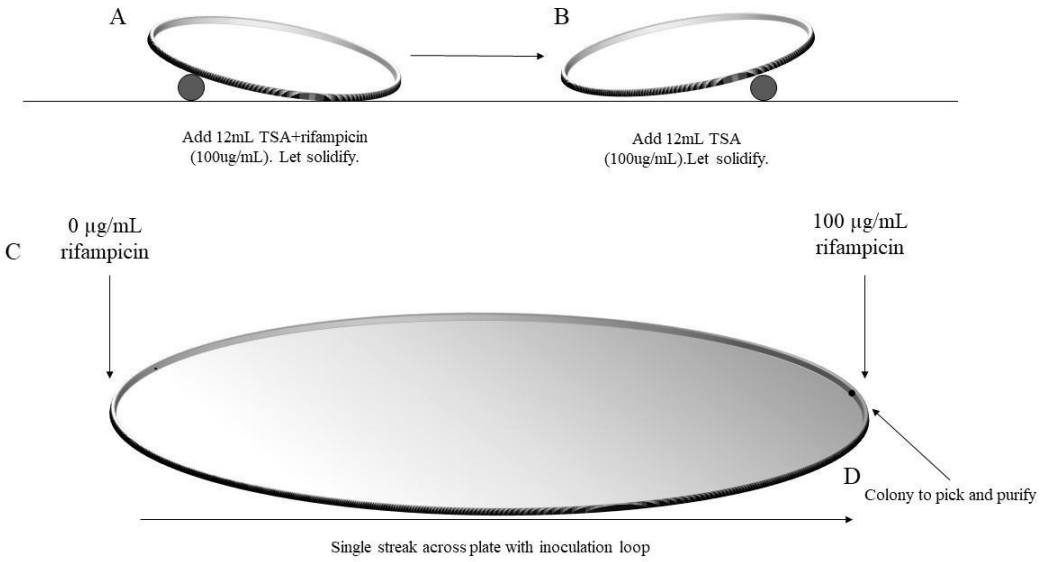

Figure 2. Rifampicin gradient plate preparation and results. A) When preparing the plate, pour 12 mL of TSA+rifampicin (100µg/mL), prop up plate so that the media accumulates on one side and let solidify. B) Prop up the other side of the plate and pour another 12 mL of TSA (no rifampicin) on top of existing agar. Let solidify. C) Make one streak of culture across the plate, using an inoculation loop, and incubate. D) Pick the colony that

grows at the highest rifampicin concentration, purify on TSA+rifampicin (100 μg/mL) plate and store in glycerol stock as a capture strain.

Exogenous plasmid capture samples

Two samples of poultry litter were collected from a broiler chicken house on Wenger's Mill Road, in Linville VA, in January 2018, one labeled 'Old' and one 'New'. The 'Old' litter was collected from the corners of the chicken house, where litter is changed less frequently and the 'New' litter was taken from the middle of the chicken house where the litter is changed more frequently. This litter was stored at room temperature (RT) in sealed, plastic, gallon sized bags.

Exogenous plasmid capture

E.coli LA61^{RifR} was cultivated shaking at 220 RPM. One mL of culture was centrifuged at 5,800 *x g* for 10 minutes and the supernatant was removed. One mL of 1X PBS was added and the tube vortexed until the cell pellet was resuspended. The culture was again centrifuged at 5,800 *x g* for 10 minutes before again resuspending the pellet in 1mL of 1X PBS.

In order to release potential plasmid donor cells from the litter, 10g of poultry litter was mixed with 90mL of 0.1% sodium pyrophosphate in a sterile 250mL Erlenmeyer flask. The flask was sealed with parafilm and vortexed for 30 seconds until the mixture was homogenous, and the suspension settled for 5 minutes at RT. After settling, 1.5mL of liquid was removed and centrifuged for 10 minutes at 5,800 *x g*. The pellet was aspirated and dissolved in 1mL 1X PBS (pH 7). The mixture was centrifuged a second time for 10 minutes and resuspended in PBS.

Control plates were prepared by separately spread plating the capture and donor cells on individual TSA+tet (25μg/mL) and TSA+rif (100μg/mL) plates.

In order to conjugate the capture and donor cells, 500µL of re-suspended capture cells were added to 500µL of re-suspended donor cells in a 1.5mL microcentrifuge tube, and allowed to incubate static at RT for 30 seconds. For liquid conjugation, 400µL of the capture-donor cell mixture were added to 600µL of TSB in a 1.5 ml tube and incubated while shaking at 220 RPM. Filter conjugation was prepared by pipetting 200µL of the capture-donor cell mixture onto a 45µm membrane disc filter placed on a TSA plate and incubated static.

After incubation, the filter was aseptically removed and washed in 10mL of PBS to remove transconjugants. One hundred microliters µL of the liquid conjugated and filter conjugated cells were spread plated onto separate TSA+tet (25µg/mL)+rif (100µg/mL) plates and incubated for 5 days static. Plating on media with added tetracycline and rifampicin selected for plasmids carrying tetracycline resistance conjugated into the rifampicin-resistant LA61^{RifR} capture strain.

Twelve transconjugants were randomly selected (named LA61::pEH1-12), 6 from filter conjugation and 6 from liquid conjugation, and isolation by streaking onto eosin methylene blue (EMB) agar. This was done to confirm that all transconjugants were *E. coli* LA61. *E. coli* LA61^{RifR} was used as a control.

Mini plasmid prep and visualization of pEH1-12

All pDNA was extracted using a plasmid DNA extraction protocol developed previously in the Herrick lab (19). One and a half milliliters of turbid LA61^{RifR}::pEH11 culture was transferred to a sterile microcentrifuge tube and centrifuged at 10,000 *x g* for 5 minutes at RT. The supernatant was poured off and the pellet resuspended in 100µL of resuspension buffer (50mM dextrose, 10mM EDTA 10mM Tris-Cl, pH 8). Subsequently, 100µL of 0.2M NaOH/1%

SDS was added, inverted 5X, and allowed to sit at RT for 5 minutes. One hundred fifty microliters each of 7.5M ammonium acetate and chloroform were added, inverted 5X to mix, and chilled on ice for a minimum of 10 minutes. The tube was again inverted 5X before RT centrifugation for 10 minutes at 10,000 *x g*. After centrifugation, the aqueous phase was transferred to a new microcentrifuge tube containing 200µL 30% PEG/1.5M NaCl, inverted 3-4X, and chilled on ice for another 10-15 minutes. The tube was again centrifuged for 10 minutes, RT, at 10,000 *x g* and supernatant removed by aspiration with care to not disturb the pellet. One mL of freshly-prepared - 20°C 70% EtOH was added and the centrifuged again at RT for 5 minutes at 10,000 *x g*. The supernatant was removed and the pellet air dried inside a fume hood for 10 minutes. The pellet was then resuspended in 50µL of sterile distilled and deionized H_2O stored at -20°C.

Results of the mini plasmid prep were visualized on a 1% agarose gel. Ten microliters of λ/HinDIII ladder was loaded into the first and last lanes, respectively, of the gel. To load the plasmid DNA (pEH1-12) and the negative control ($LA61^{RifR}$), 13 beads of 4µl loading dye were laid out on a piece of parafilm and 16µl of pDNA were added to the corresponding loading dye bead. The 20µl of each pDNA+loading dye bead was transferred to the designated well. The voltage source was set to 70V for 60 minutes and the gel was stained in 3X Gel Red (Sigma-Aldrich, St. Louis) for 30 minutes before imaging.

Assessing antibiotic susceptibility using a modified Stokes test

Twelve transconjugants, six from filter conjugation and six from liquid conjugation, were randomly selected from the exogenous plasmid capture plates. After isolation and purification, the twelve selected transconjugants were tested for susceptibility to ten antibiotics (in addition to

tetracycline) using a modified Stokes test (13). The antibiotics tested were tobramycin (10 µg), streptomycin (10 µg), gentamicin (10 µg), aztreonam (30 µg), kanamycin (30 µg), imipenem (10 µg), piperacillin (100 µg), piperacillin/tazobactam (100 µg/10 µg), ceftazidime (30 µg), and ciprofloxacin (5 µg). The Stokes test is similar to a Kirby Bauer test, where antibiotic discs are placed on a lawn of growth and zones of inhibition measured. However, a Stokes test directly compares the transconjugant's and host strain's antibiotic susceptibility (Figure 3).

The turbidity of each transconjugant, LA61^{RifR}::pEH1-12, and the capture strain, LA61^{RifR}, was standardized using a 0.5 McFarland standard. For each transconjugant and the capture strain, single colonies were added to a tube filled with PBS until the turbidity of each culture matched the turbidity of the 0.5 McFarland standard. The capture strain lawn was swabbed, with a sterile swab, onto the first and third sections of a Meuller Hinton agar plate and the transconjugant was swabbed onto the second section. On the lines dividing each section, each antibiotic disc was placed (Figure 3). Post incubation, the zones of inhibition were measured in millimeters with an electronic caliper and recorded.

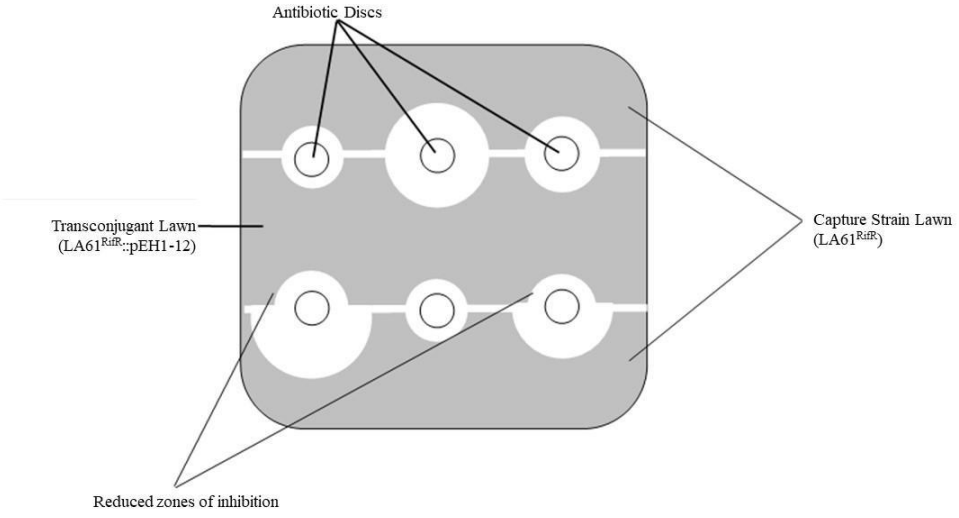

Figure 3. Modified Stokes test for testing antibiotic susceptibility of transconjugant in comparison to capture strain growth. Figure modified from Herrick et al., 2014.

Whole genome extractions

An overnight culture of LA61^{RifR}::pEH11 was cultivated at 37°C and 220 RPM. Whole genome extraction was carried out using the QIAGEN DNeasy® Blood and Tissue Kit (Hilden, Germany) and following the manufacturer's protocol 'Pretreatment of Gram Negative Bacteria and Purification of Total DNA from Animal Tissues' (20). One and a half mL of each culture were centrifuged for 10 min at 5000 *x g*, and resuspended in 180µl Buffer ATL. Twenty µl of proteinase K was then added. The tubes were briefly vortexed and then incubated at 56℃ in a Bioshake (Q Instruments, Jena, Germany) for 25 minutes at 300RPM until the tubes were clear and viscous, indicating lysis. The tubes were again vortexed for 15 seconds before 200µl of Buffer AL were added and mixed by vortexing. Two hundred microliters of ice-cold 100% ethanol were added and the tubes were vortexed to mix. Mixtures were pipetted into the DNeasy Mini spin columns, placed into a two ml collection tube, and centrifuged three times (6,000 *x g*, 6,200 *x g*, and 6,200 *x g*, one minute each) so that all supernatant was collected as flow through. The centrifuge speed was increased to 6,200 *x g* for two subsequent centrifugation steps to increase flow through. After all supernatant had been collected, the collection tube with the flow through was discarded, the spin column placed into a second collection tube, 500µl Buffer AW1 added to the spin column, and subsequently centrifuged for one minute at 6,200 *x g*. After centrifugation, the spin column was placed into a third collection tube, 500µl of Buffer AW2 was added to dry the DNeasy membrane, and centrifuged for 3 minutes at 20,000 *x g*.

After the DNeasy membrane was dry, the DNA was eluted into a 1.5 ml microcentrifuge tube by pipetting 100µl of H_2O directly onto the DNeasy membrane, incubating for one minute

at RT, and centrifuging for one minute at 6,200 *x g*. This was repeated with an additional 100µl of H_2O in order to increase DNA yield.

DNA concentration and purity were determined using the Qubit 2.0 fluorometer (Invitrogen, Carlsbad) and H1 Synergy Multi-Mode Reader (BioTek, Winooski), respectively. The concentration of LA61^{RifR}::pEH11 was determined using the Qubit® dsDNA BR Assay Kit. Qubit® working solution was prepared in a sterile microcentrifuge tube by diluting the Qubit® dsDNA BR reagent 1:200 into the Qubit® dsDNA BR Buffer. Two hundred microliters of working solution was prepared for the LA61^{RifR}::pEH11 DNA extraction and an additional 400µl to account for two standards. One hundred-ninety microliters of Qubit working solution was added to 10µl of Qubit® standard #1 in a thin-walled, clear, half ml, Qubit® assay tube, and Qubit® standard #2 in a second tube. Two microliters of LA61^{RifR}::pEH11 dsDNA were added to 198µl of Qubit® working solution in the third tube. All assay tubes, with a final volume of 200µl, were vortexed for two to three seconds -- with care to not create bubbles -- and incubated at RT for two minutes. The Qubit® 2.0 Fluorometer was set to the "DNA and dsDNA Broad Range" settings. Both standards were read and the concentration of LA61^{RifR}::pEH11 was then recorded.

DNA purity of the LA61^{RifR}::EH11 extraction was determined using the BioTek Reader. The reader was blanked using H_2O and two µl of the DNA extraction was added to two wells of the plate. The software used was Gen5 2.09 and double stranded DNA was selected as the sample type. An A260/280 ratio of ~1.8 was considered pure.

Illumina MiniSeq & Oxford Nanopore MinION library prep & sequencing

The whole genome of strain *E. coli* LA61^{RifR} was sequenced by the Virginia Department of Consolidated Laboratory Services (DCLS) using an Illumina MiSeq short-read sequencer. The strain was shipped to the DCLS in a 1.2 mL Fisherbrand screw-top cryogenic vial. TSA stabs were prepared and inoculated with a colony of LA61^{RifR}. The stab was incubated and stored, at 4°C, until shipping. In order to ship the inoculated TSA stab, the cryogenic tube was wrapped in a paper towel and placed inside a primary biohazard shipping container. The primary container was placed in a secondary container filled with absorbent material, which was placed in an outer container with the necessary permits, papers, and directions for shipping and handling biohazardous materials.

The whole genome of LA61^{RifR}::pEH11 (54.5 ng/µl, 260/280: 1.964) was also sequenced (in-house) using the Oxford Nanopore Technologies^{TM} MinION long-read sequencer in June 2019. The FlowCell was quality checked for number of active pores by connecting the MinION to the Oxford Nanopore Technologies^{TM} MinKNOW software and connecting to the MinIT remote computer, adding the new FlowCell, selecting the FLO-MIN106 FlowCell, and pressing *Check Flow Cells* and *Start Test*. A minimum of 800 active pores was needed for a successful run.

In order to prepare the DNA library for MinION sequencing, the *O*xford Nanopore™ Rapid Barcoding Sequencing (SQK-RBK004) and Flow Cell Priming (EXP-FLP002) kits were used. The Fragmentation Mix RB01-12 and Rapid Adapter (RAP) were briefly centrifuged and mixed by pipetting. The Sequencing Buffer (SQB), Loading Beads (LB), Flush Buffer (FLB), and Flush Tether (FLT) were thawed at RT, mixed by pipetting, and stored on ice until use.

The DNA extraction was prepared in nuclease-free water by transferring ~400 ng genomic DNA into a DNA LoBind tube (7.5µl of LA61^{RifR}::pEH11extracted DNA), adjusting the volume to 7.5µl with water, mixing the tube by gentle flicking (avoiding unwanted shearing), and briefly spinning down then mixing well by pipetting. In another 0.2µl thin-walled PCR tube the 7.5µl of the 400ng template DNA was mixed with 2.5µl Fragmentation Mix (LA61^{RifR}::pEH11- RB11), mixed by flicking the tube, spinning down, and incubating at 30℃ for 1 minute, 80℃ for 1 minute, and incubated on ice for 1 minute. The 30℃ and 80℃ steps of incubation were completed using a thermocycler. This was repeated for each extraction being sequenced on the flow cell. There were 11 additional DNA extractions, from the Herrick laboratory, on the flow cell with LA61^{RifR}::pEH11. All DNA extractions were pooled together in the desired ratio to the final volume (10µl/extraction X 12 extractions = 120 µl).

In order to achieve improved yield (max. 30 Gpb, nanoporetech.com), the pooled extractions were cleaned and concentrated using AMPure XP beads (stored at 4℃). The beads were resuspended by vortexing right before adding an equal amount of beads to the volume of pooled and barcoded extraction mixture. This mixture was flicked with a finger to gently mix and incubated for 5 minutes at RT. While incubating, the tubes were flicked every 1 minute. The extraction mixture was spun down, placed on a magnet (Beckman Coulter Life Sciences, Agencourt SPRIStand), and after the magnetic beads had formed a pellet on the magnet, the

supernatant pipetted off. The beads were washed in 200µl freshly prepared 70% ethanol (prepared in nuclease-free water) without disturbing the pellet. The ethanol was removed using a pipette and the wash repeated before removing the ethanol a second time. The tubes were spun down and placed on the magnetic rack and the pellet resuspended in 10µl of 10mM Tris-HCl (pH 7.5) with 50mM NaCl before being incubated for two minutes at RT. The beads were kept, still and pelleted, on the magnet until the eluate became clear and colorless. This took approximately 30 seconds. The eluate was removed by pipetting into a clean 1.5ml Eppendorf LoBind tube. To the concentrated library, 1µl of RAP was added and mixed by flicking. The reaction was spun down and incubated for 5 minutes at RT.

The next step was to prime and load the flow cell. Separate tubes of SQB, LB, FLT, and FB were again mixed by vortexing and returned to the ice. The MinION Mk 1B lid was opened and the flow cell was slid in under the clip. In order to open the priming port, the priming port cover was slid clockwise. To remove the bubble under the priming port, a P1000 pipette was set to 200µl, the tip inserted into the port, and the wheel of the pipette turned until a small volume of buffer was seen entering the pipette tip. It was extremely important to keep the membrane covered in the buffer by not removing more than the air bubble and a few microliters of the buffer. In order to prepare the flow cell priming mix, 30µl of thawed and mixed FLT were added directly to the FB tube and vortexed to mix. Eight hundred microliters of the priming mix was loaded into the flow cell through the priming port, avoiding the introduction of any air bubbles. The flow cell was left at RT for 5 minutes while the tubes of SQB and LB were each mixed well by pipetting. In a new tube, 34µl of SQB, 25.5µl of LB, 4.5µl Nuclease-free water, and 11µl of the DNA library were added. It is important to note that the LB tube was mixed immediately before adding since the beads settle very quickly.

In order to prime the SpotON sample port, the cover was lifted and 200µl of the priming mix was added to the priming port. The prepared library was mixed by pipetting immediately before 75µl of the library was loaded into the flow cell via the SpotON sample port. This was done drop by drop, ensuring that each drop flowed into the port before the addition of the next. The SpotON sample port cover was replaced and the priming port closed before the MinION lid was replaced.

On the MinKNOW graphical user interface, a new experiment was set up and named, the SQK-RAD004 control experiment was selected, *Basecalling* was turned on and *Fast basecalling* was selected, and all run and output options remained on their default settings. Then the experiment was started. The MinION was connected to the ONT MinIT, a computing unit equipped with a 256-core GPU, so that MinKNOW could locally basecall the real-time data using Guppy (21), before the experiment was complete and FAST5 files output for assembly and analysis.

Assembly of DNA raw reads

Short-read assembly of LA61^{RifR}

Raw sequencing read files of LA61^{RifR} were placed by the DCLS in the DCLS-JMU data library on GalaxyTrakr (galaxytrakr.org; 22), as well as on the Illumina BaseSpace website. Raw reads were assembled using short read tools on GalaxyTrakr. The uploaded reads were quality checked and trimmed using FastQC (23) and Trimmomatic (24). Trimmomatic was started using three operations, in the following: AVGQUAL with a 27bp minimum average quality per read, SLIDINGWINDOW with a 4bp window and 20 average quality requirement, and MINLEN with

a 20bp minimum for reads kept. Fast QC was, again, utilized to verify the improvement in read quality.

SPAdes (25) is a tool used to combine the individual short reads into a draft assembly. The assembler works by identifying overlapping sequences on the ends of reads and joining corresponding sequences into as few contigs as possible. The output files from the Trimmomatic tool, R1 (paired) and R2 (paired), were input as separate forward and reverse fastq.gz files. The following parameters were used in SPAdes: "*Single-cell?"*, *"Run only assembly? (without error correction)"*, *"Careful correction?"*, and *"Automatically choose K-mer values?"* were not selected; K-mer values were 21,33,55,77, 99, and 127; *"Coverage cutoff"* was turned off; the input libraries were not IonTorrent reads; the *"Library type"* was *"Paired-end / Single reads"* with a forward / reverse direction *"-> <- (fr)"*; the file formats were separate input files; *"Yes"* was selected for *"Output assembly graph (contigs)?"* and "*No"* was selected for *"Output assembly graph with scaffolds?"*. After all output files had been generated, the assembly quality was analyzed using Quast (26) tool.

The final assembly was kept only if the QUAST report met the following parameters. The assembly had to have a mean Q (Phred Score measures quality of individually base called called nucleotides) greater than 30, a mean depth of coverage greater than 30x, an N_{50} greater than 200,000, less than 200 contigs, and a sequence length comparable to reference *E.coli* genomes, 4.6-5.6 Mbp. The assembly was visualized using Bandage (27).

Assembling pEH11

Fastq files of $LA61^{RifR}$::pEH11 and $LA61^{RifR}$ were input into Unicycler on Galaxy (usegalaxy.org; 23) to generate a hybrid assembly. The long read fastq files of $LA61^{RifR}$::pEH11

generated by the ONT MinION run on the ONT MinIT, as well as the Illumina short read fastq files generated by the Virginia DCLS, were input into the program. In order for Unicycler to accept the files, the ".fastq" file extension was replaced with ".fastqsanger". However, it is important to note that the files remained fastq files and, despite the name change, were not converted to fastqsanger files. The output was a fasta assembly file and a graph file that was input into Bandage for visualization. The generated fasta file was input into QUAST in order to analyze the quality of the Unicycler assembly.

The LA61^{RifR}::pEH11 FASTQ long-reads were also assembled using Canu (28) on the Galaxy Europe platform (https://usegalaxy.eu/; 22). FAST5 files were input into Canu and run with the mode set to nanopore raw, no restrictions on Canu, a 5 million bp estimated genome size, no maximum raw overlap mismatch or corrected overlap mismatch, 1000 bp minimum read length, 500 bp minimum overlap, and a 40X target coverage for corrected reads. There were no additional options selected. The contigs.fasta file was loaded into Geneious Prime (Geneious Prime 2020.0.3. (https://www.geneious.com)), and the size, GC content, and predicted circularization noted. The Contig #2 FASTA file was input into NCBI BLASTn (29) and the results used to identify Contig #2 as the contig containing the plasmid, pEH11, and to determine pEH11's incompatibility group. Other output contigs from Canu were also input into NCBI BLASTn to identify additional contigs making up the pEH11.

Annotation of pEH11

After using NCBI Blast to confirm Canu's output contig "*Contig #2"*As pEH11, the top three BLASTn matches were aligned with pEH11 using Mauve to determine and visualize the

varying regions between the genomes. The sequence of contig #2, now referred to as pEH11, was input into ABRicate (https://github.com/tseemann/abricate) in GalaxyTrakr using the databases *Resfinder* (30), *Argannot* (31), *Card* (32), *NCBI betalactamase* (33), *Plasmidfinder* (34), and *Vfdb* (35). Each database was searched without *Suppress header* and a minimum DNA identity of 75. Additionally, the pEH11 sequence was input into *Prokka* (36) on Galaxy Europe with the following parameters: a single dataset; no *Locus tag prefix*; The *Locus tag counter increment* was 1; The *GFF version* was 3; *Force GenBank/ENA/DDJB compliance* was turned off; *Gene* features were not added for each *CDS* feature; The *Minimum contig size* was 200; The *Sequencing centre ID, Genus name, Species name, Strain name, Plasmid name or identifier* were left without parameters; The *Kingdom* selected was *bacteria*; The *Genetic code* was 11; *Use genus-specific database* was off; Nothing was selected for the *Additional FASTA file of proteins to first annotate from; Improve predictions for highly fragmented genomes, Fast mode, Don't run rRNA search with Barrnap,* and *Don't run tRNA search with Aragorn* were not selected; The *Similarity e-value cut-off* was set to 0.000001; All *Additional outputs* were selected. The pEH11 sequences was through *Integron Finder* (37) on Galaxy Pasteur (https://galaxy.pasteur.fr/; 38) with the following parameters: The *Replicon files* was a *single dataset*; *Thorough local detection, Use linear replicon, Annotate cassettes using HMM profiles, Use different HMM banks, Just look for attC sites,* and *Use your own covariance matrix* were turned off; The *Threshold for clustering (in base)* was 4000; The *attC e-value* was 1; The *maximum* & *minimum value*(s) *for attC size* were 200 and 40, respectively. The last tool used to annotate pEH11 was *Rapid Annotation using Subsystem Technology* or RAST (https://rast.nmpdr.org/rast.cgi; 39). The FASTA sequence file of pEH11 was input into RAST's online annotation tool which generated a categorical distribution of features within the plasmid genome as well as a list of

predicted features. Using these online annotation tools as reference, pEH11 was manually annotated in Geneious Prime by inputting the genes with predicted functions and a methionine starting amino acid. BLASTn and HHPred were also used to evaluate uncertain genes.

Results:

Exogenous plasmid capture and phenotypic analysis

Two samples were collected from a chicken house and plasmids conferring resistance to tetracycline were selected for in an exogenous plasmid capture. After incubation for 5 days, the new litter produced no transconjugants and the old litter transconjugant plates were covered in individual colonies (Figure 4). Twelve of the colonies from the transconjugant plates (six each from liquid and filter conjugation) were randomly selected for further analysis.

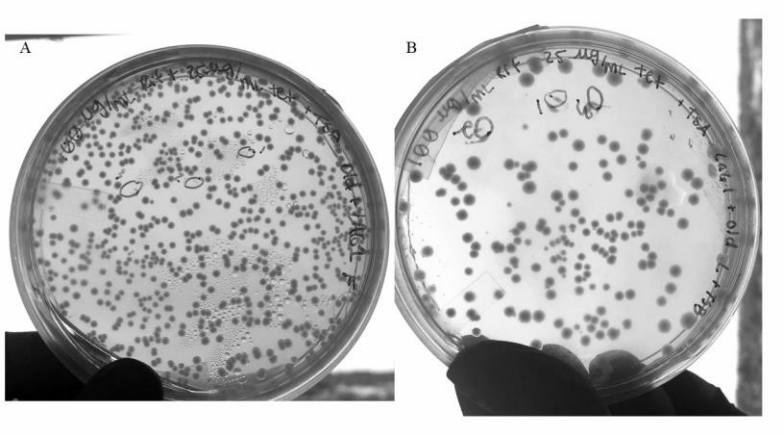

Figure 4. Examples of transconjugants on TSA amended with tetracycline (25 μg/mL) and rifampicin (100 μg/mL). (A) transconjugants captured via filter conjugation (B) transconjugants captured via liquid conjugation.

After incubation, the LA61^{RifR} on the rifampicin-containing positive control plate had a lawn of growth and the washed litter (old and new) on the tetracycline-containing positive control plates had individual colonies. Nothing grew on the LA61^{RifR} negative control plate. However, there were a few colonies on the washed litter (old and new) rifampicin negative control plates. This unexpected result gave reason to plate the twelve selected colonies on EMB and confirm the presence of plasmid HGT with a mini plasmid prep. All 12 re-plated colonies produced the typical purple and green sheen morphology of *E. coli* on EMB agar. The mini plasmid prep results showed that all selected transconjugants contained a plasmid (Figure 5).

There was no evidence of plasmid DNA in the negative control lane (containing LA61^{RifR} only). The exogenous plasmid capture control plates, in conjunction with the transconjugant colony morphology on EMB agar and the presence of plasmids after the plasmid mini prep, suggested that all twelve colonies were LA61^{RifR} transconjugants.

While all the plasmids were larger than 23kb, the plasmid in transconjugant EH11 was larger than the other captured plasmids, all of which appear to be approximately the same size (Figure 5).

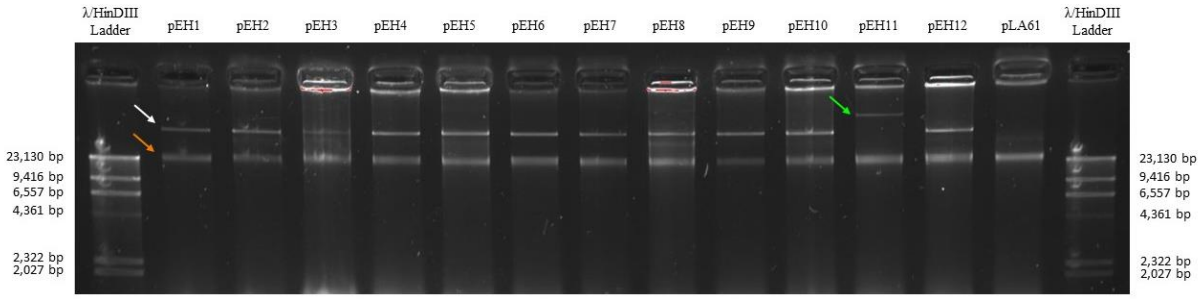

Figure 5. Gel electrophoresis of the mini plasmid prep of twelve transconjugants. The white arrow indicates the plasmid DNA, the orange arrow indicates the band of primarily sheared chromosomal DNA, and the green arrow indicates the large plasmid in transconjugant EH11.

Antibiotic susceptibility

A modified Stokes test was used to determine the antibiotic susceptibility of the 12 transconjugants relative to capture strain LA61^{RifR}. Eleven of the 12 transconjugants conferred resistance to more than one antibiotic (excusing tetracycline), some of which were late generation, clinically-significant antibiotics (Table 1).

Table 1. Results of modified Stokes test of antibiotic susceptibility. Tetracycline resistance (TET^{R}) was a control, as it was selected for during the exogenous plasmid capture. Transconjugants conjugated on a filter plate are noted with an "F" and those conjugated in TSB liquid media are noted with an "L".

Transconjugant	Conjugation method	Antibiotic Resistance phenotype[a]
LA61^{RifR}::pEH1	L	TET^{R}, PIP^{R}, TZP^{R}
LA61^{RifR}::pEH2	L	TET^{R}, PIP^{R}, CIP^{R}
LA61^{RifR}::pEH3	F	TET^{R}, PIP^{R}, TZP^{R}
LA61^{RifR}::pEH4	F	TET^{R}, ATM^{R}, PIP^{R}, TZP^{R}
LA61^{RifR}::pEH5	F	TET^{R}, PIP^{R}, TZP^{R}
LA61^{RifR}::pEH6	F	TET^{R}, PIP^{R}, TZP^{R}
LA61^{RifR}::pEH7	F	TET^{R}, PIP^{R}, TZP^{R}
LA61^{RifR}::pEH8	F	TET^{R}, PIP^{R}, TZP^{R}
LA61^{RifR}::pEH9	L	TET^{R}, PIP^{R}
LA61^{RifR}::pEH10	L	TET^{R}, PIP^{R}, TZP^{R}
LA61^{RifR}::pEH11	L	TET^{R}, S^{R}, ATM^{R}
LA61^{RifR}::pEH12	L	TET^{R}, PIP^{R}, TZP^{R}, CAZ^{R}

[a]S = streptomycin, ATM = aztreonam, PIP = piperacillin, TZP^{R} = piperacillin/tazobactam, CAZ^{R} = ceftazidime, CIP^{R} = ciproflaxocin.

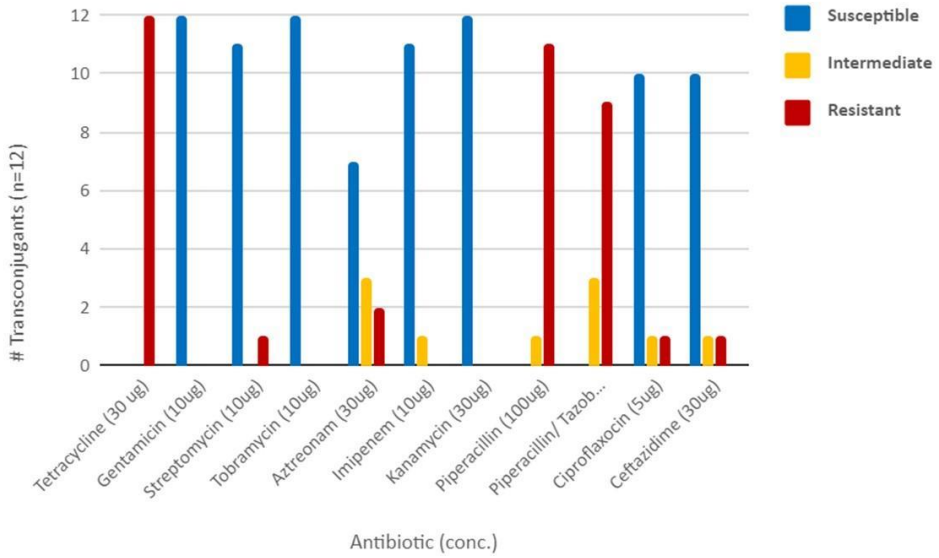

Figure 6. Phenotypic resistance observed for each of 11 antibiotics, tested on the twelve transconjugants EH1-12.

Phenotypic "resistance" was defined as having a difference of 3 mm between the clearing zone radii of transconjugant and LA61^{RifR}. Intermediate resistance was defined as a radius difference of between 2 and 3 mm between the transconjugant and LA61^{RifR}'s inhibitory zone. LA61^{RifR}::pEH3,8,12 were intermediately resistant to aztreonam (30µg), LA61^{RifR}::pEH1 was

intermediately resistant to imipenem (10µg), LA61^{RifR}::pEH11 was intermediately resistant to piperacillin (100µg) and ciproflaxocin (5µg), LA61^{RifR}::pEH3,9,11 were intermediately resistant to piperacillin/tazobactam (100µg + 10µg), and LA61^{RifR}::pEH8 was intermediately resistant to ceftazidime (30µg). None of the transconjugants tested were resistant to gentamicin (10µg), tobramycin (10µg), or kanamycin (30µg).

Because of its unique size and antibiotic resistance phenotype, whole genome sequencing was carried out on transconjugant LA61^{RifR}::pEH11.

LA61^{RifR} genome assembly

In order to analyze transconjugant LA61^{RifR}::pEH11, it was necessary to accurately assemble the capture strain, LA61^{RifR}. The FastQC report of the raw LA61^{RIfR} short reads demonstrated that, out of a total 1,179,567 reads, no reads were flagged for poor quality and all ranged between 35-251bp in length. After trimming, there were a total of 1,076,091 reads with improved quality (Figure 7).

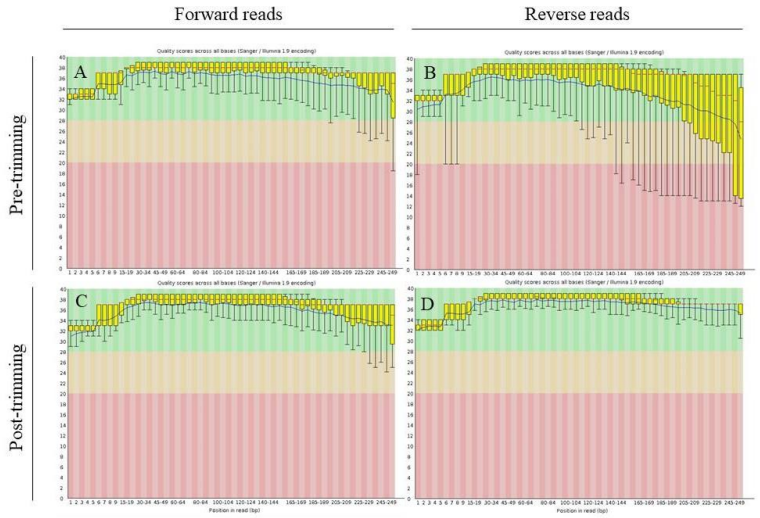

Figure 7. Improvement in LA61^{RifR} short-read quality output from the Illumina MiSeq. Panels A and C show the forward reads pre- and post-Trimmomatic. Panels B and D

show the reverse strands pre- and post- Trimmomatic. The reduction of reads from the red and orange area into solely the green area, due to higher phred scores, indicate better read quality.

After trimming using Trimmomatic, and assembling the short reads using SPAdes, 97 output contigs resulted. These were then visualized in Bandage and a graph generated (Figure 8A). According to Quast, the largest contig was 633,582 bp in length and the assembly had a total length of 5,128,965 bp and a GC% of 50.51%, which is consistent with the length and GC content of an *E. coli* chromosome. The depth of coverage graph shows that the longest contig was 1,500,000 bp in length and had a coverage depth of ca. 35.5 (Figure 8D). The average depth of coverage of the assembly was 36.2X and the N_{50} number was 225,150. BLAST results of the first, unassembled contig (Figure 8C) *E. coli* isolate SC457 chromosome.

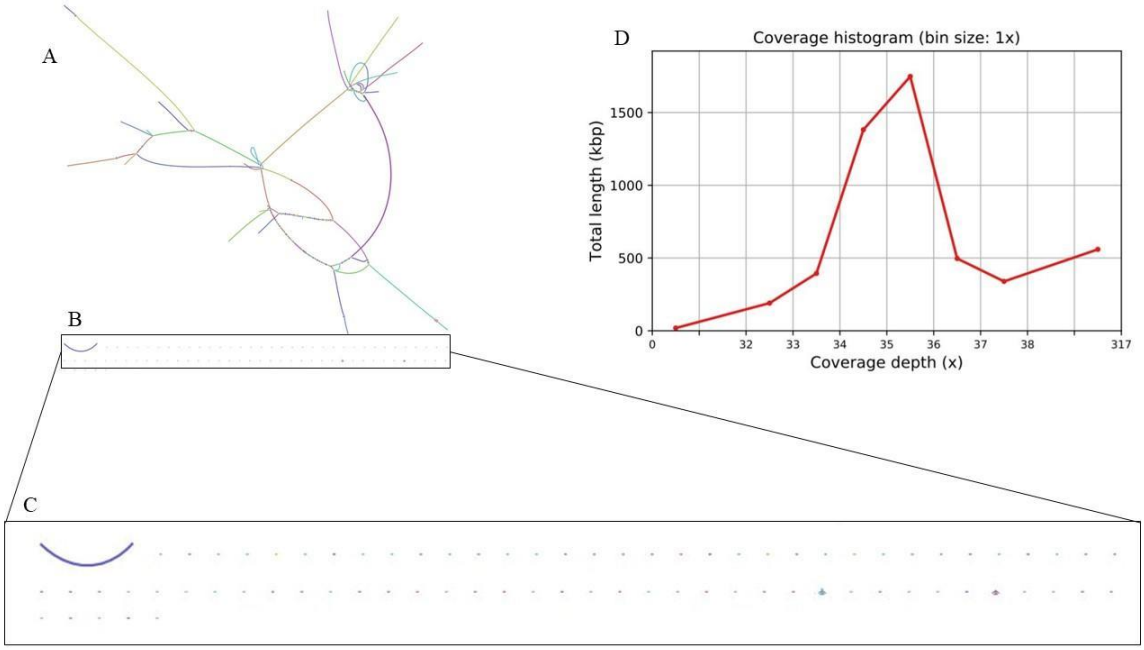

Figure 8. Bandage graph of 97 assembled contigs from LA61^{RifR} short-reads. Assembly generated with SPAdes. (A) Map of assembled and connecting contigs. (B) Assembled, unconnected contigs. (C) Zoomed view of assembled, unconnected contigs. (D) Coverage histogram of LA61^{RifR} short-read SPAdes assembly.

pEH11 assembly and annotation

LA61^{RifR}::pEH11 long-reads and LA61^{RifR} short-reads were hybrid-assembled using Unicycler resulting in 9 contigs. The largest contig was 3,699,052 bp and the assembly length was 5,163,374 bp, according to QUAST. A Bandage visualization of the assembly graph can be seen in Figure 9. When each contig was input into NCBI BLASTn, none of the contigs aligned to a plasmid or to a feature that would indicate a plasmid.

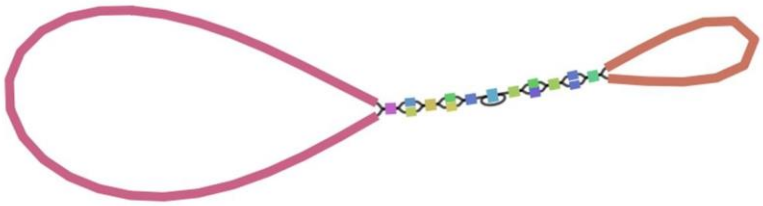

Figure 9. Bandage graph of LA61^{RifR}::pEH11 long-reads and LA61^{RifR} short-reads.

Assembly of only the LA61^{RifR}::pEH11 long-reads using Canu generated only 5 contigs, fewer than the 9 contigs generated by the Unicycler hybrid assembly. Contig #1 was the largest and the size of the capture strain chromosome while the other conigs were much smaller in length and had varying GC contents (Table 2).

Table 2. Data, generated by Geneious Prime, on LA61^{RifR}::pEH11 contigs

Contig	Reads	Length (bp)	GC%	Predicted structure[a]
Contig #1	18,715	5,183,052	50.6	circular
Contig #2	471	177,343	49.7	circular
Contig #3	16	3,839	52.1	linear
Contig #4	51	4,412	54.9	linear
Contig #5	30	3,995	56.0	linear

[a] Structure predicted by Geneious Prime

The results after each contig were input into NCBI BLASTn demonstrated that *Contig #1* was the *E. coli* LA61^{RifR} chromosome, *Contig #2* was pEH11, and *Contig #3*, *Contig #4*, and *Contig #5* matched to fragments of plasmids isolated from a variety of gram negative species such as *Escherichia*, *Salmonella*, *Enterobacter*, *Klebsiella*, and *Proteus*. A Mauve alignment provided visual comparison of the high similarity between *Contig #2* and three *Salmonella enterica* plasmids (Figure 10). This visual result was quantitated by the query cover and percent identity NCBI BLASTn values (Table 3).

Table 3. NCBI BLASTn results of *Contig #2*

Description	Max Score	Total Score	Query Cover	E value	Percent Identity	Accession
Salmonella enterica subsp. enterica serovar Kentucky plasmid pSSAP03302A, complete sequence	1.720E+05	3.645E+05	100%	0.0	99.24%	CP002089.1
Salmonella enterica subsp. enterica str. CVM29188 plasmid pCVM29188_146, complete sequence	1.489E+06	3.645E+05	100%	0.0	99.21%	CP001122.1
Salmonella enterica subsp. enterica serovar Kentucky plasmid pCS0010A, complete sequence	1.489E+06	3.643E+05	100%	0.0	99.21%	CP002090.1

Figure 10. Mauve alignment of pEH11 to its three closest NCBI BLASTn matches, from *Salmonella enterica* subsp. enterica serovar Kentucky plasmid complete sequences pCS0010A, pSSAP03302A, and pCVM29188_146.

Integron Finder did not identify any integrons from the fasta assembly file of Contig #2. However, when this file was input into ABRicate several antibiotic resistance genes were identified (Table 4). Each of these genes was confirmed using the whole genome annotators *Prokka* and RAST. A map of all annotated genes is displayed in Figure 11. These annotations suggested regions of pEH11 with specific functions (Figure 11B).

Table 4. Antibiotic resistance genes identified on pEH11.

Gene	Start	Stop	% Coverage	% Identity	Database(s)	Accession #
Tet(B)[a]	72682	73882	99.59	99.58	resfinder, argannot, card	AF326777
StrA[b]	78563	79363	99.63	99.63	resfinder, argannot, card	AF321551
StrB[b]	79363	80195	99.52	99.16	resfinder, argannot, card	M96392
Tet(D)	74670	75086	99.76	99.76	card	BX664015

[a]Argannot database identified this gene as Tet(A) (Accession: NC_020418).
[b]StrA and StrB were also identified as genes contributing to general aminoglycoside resistance.

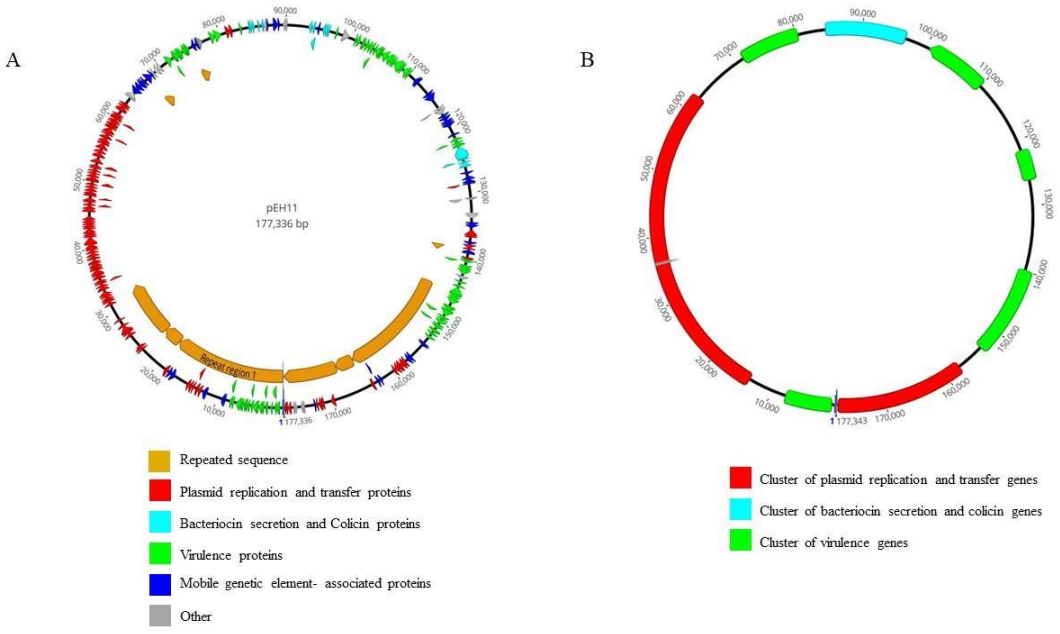

Figure 11. Annotated pEH11 map as visualization of motifs and positions of similar gene functions on accessory genome. A) Map of genes annotated for function. B) Regions of plasmid genome associated specific functions. Plasmid replication and transfer genes are not considered accessory genes, as they are vital to plasmid function.

Prokka and RAST identified a total of 47 transposase genes. Nine repeat regions were identified in RAST. Two main regions of the pEH11 genome contain plasmid replication, stability, and transfer genes. These replication and transfer proteins (*Rep* and *Tra*) alongside NCBI BLASTn results, identified pEH11 as a plasmid in the IncF incompatibility group. RAST also identified *doc* toxin genes. In addition, pEH11 encoded many colicin and unspecified bacteriocin genes, which were identified by both Prokka and RAST. RAST also identified a bacteriophage tail assembly protein which suggests the presence of an incomplete prophage.

There were several distinct regions of the pEH11 chromosome, each with a singular predicted function (Figure 11B). There were two regions that contain genes necessary for plasmid replication and conjugation. There were several regions that contain virulence proteins and ARGs. Three such regions that facilitate iron uptake, labeled in Table 5 as "aerobactin system", are each ~7-8 Kbp in length. pEH11 also contains a region that produces Colicin and other bacteriocin-related proteins.

Table 5. Annotated genes on plasmid EH11.

Gene	Predicted function	Start[a]	Stop	Strand
iucA	Aerobactin system	617	724	F
iucA	Aerobactin system	758	1105	F
iucA	Aerobactin system	1102	1248	F
iucA	Aerobactin system	1215	1664	F
iucA	Aerobactin system	1630	1842	F
iucB	Aerobactin system	2325	2465	F
iucB	Aerobactin system	2462	2758	F
iucB	Aerobactin system	2733	2927	F
iucC	Aerobactin system	2927	3160	F
iucC	Aerobactin system	3160	3702	F
iucC	Aerobactin system	3764	3877	F

Table 5. continued

iucC	Aerobactin system	3864	4535	F
iucD	Aerobactin system	4637	4816	F
iucD	Aerobactin system	4777	4971	F
iucD	Aerobactin system	5036	5290	F
iucD	Aerobactin system	5408	5611	F
iucD	Aerobactin system	5605	5910	F
iutA	Aerobactin system	5992	6216	F
iutA	Aerobactin system	6260	6466	F
iutA	Aerobactin system	6463	7095	F
iutA	Aerobactin system	7451	7792	F
iutA	Aerobactin system	7789	7938	F
iutA	Aerobactin system	7916	8248	F
iutA	Transposase IS1	8700	8999	F
insB	Transposase IS1	9001	9240	F
insA	Transposase IS1	11603	11902	F
insB	Transposase IS1	11922	12278	F
parA	Plasmid partitioning	12587	12994	F
parA	Plasmid partitioning	13045	13212	F
parA	Plasmid partitioning	13212	13454	F
parA	Plasmid partitioning	13427	13753	F
parB	Plasmid partitioning	13753	14046	F
parB	Plasmid partitioning	14079	14507	F
parB	Plasmid partitioning	14607	14720	F
tnp	Transposase IS110	17360	16971	R
tnp	Transposase IS110	17554	17369	R
tnp	Transposase IS110	17993	17643	R
klcA	Plasmid anti-restriction	18358	18483	F
klcA	Plasmid anti-restriction	18480	18782	F
ssb	DNA replication, recombination, & repair	23198	23464	F
ssb	DNA replication, recombination, & repair	23570	23971	F
parB	Plasmid partitioning	25865	26263	F

Table 5. continued

psiB	SOS response	26315	26590	F
psiB	SOS response	26575	26748	F
psiA	SOS response	26745	27530	F
flmC	Toxin-antitoxin system	27680	27895	F
ssb	DNA replication, recombination, & repair	28632	28757	F
traM	Plasmid transfer	31333	31665	F
traJ	Plasmid transfer	31780	31995	F
traJ	Plasmid transfer	31992	32213	F
traJ	Plasmid transfer	32222	32455	F
traY	Plasmid transfer	32559	32960	F
traA	Pilin- Plasmid transfer	32993	33364	F
traE	Plasmid transfer	33801	34199	F
traK	Plasmid transfer	34301	34540	F
traK	Plasmid transfer	34666	34800	F
traK	Plasmid transfer	34767	34970	F
traB	Plasmid transfer	34970	35329	F
traB	Plasmid transfer	35513	35635	F
traB	Plasmid transfer	35629	36042	F
traB	Plasmid transfer	36045	36185	F
traP	Plasmid transfer	36191	36814	F
traP	Plasmid transfer	36796	36963	F
trbD	Plasmid transfer	36956	37147	F
trbG	Plasmid transfer	37178	37408	F
traV	Plasmid transfer	37405	37734	F
traV	Plasmid transfer	37731	37925	F
traR	Plasmid transfer	38050	38271	F
traC	Plasmid transfer	38431	38601	F
traC	Plasmid transfer	38598	38867	F
traC	Plasmid transfer	38973	39299	F
traC	Plasmid transfer	39423	39773	F
traC	Plasmid transfer	39841	40098	F
traC	Plasmid transfer	40127	40768	F

Table 5. continued

traW	Plasmid transfer	40765	41826	F
traW	Plasmid transfer	41823	42044	F
traU	Plasmid transfer	42041	42421	F
traU	Plasmid transfer	42418	43029	F
trbC	Plasmid transfer	43035	43259	F
trbC	Plasmid transfer	43321	43617	F
traN	Plasmid transfer	43669	43959	F
traN	Plasmid transfer	43956	44333	F
traN	Plasmid transfer	44330	44842	F
trbE	Plasmid transfer	45491	45694	F
traF	Plasmid transfer	45713	45979	F
traF	Plasmid transfer	45976	46374	F
traF	Plasmid transfer	46343	46456	F
trbA	Plasmid transfer	46494	46820	F
traQ	Plasmid transfer	46959	47240	F
trbB	Plasmid transfer	47227	47484	F
trbB	Plasmid transfer	47441	47593	F
trbF	Plasmid transfer	47927	48061	F
traH	Plasmid transfer	48048	48263	F
traH	Plasmid transfer	48269	48484	F
traH	Plasmid transfer	48543	49085	F
traH	Plasmid transfer	49093	49419	F
traG	Plasmid transfer	49412	49549	F
traG	Plasmid transfer	49546	49722	F
traG	Plasmid transfer	49680	49973	F
traG	Plasmid transfer	50025	50162	F
traG	Plasmid transfer	50147	50275	F
traG	Plasmid transfer	50295	50672	F
traG	Plasmid transfer	50669	51136	F
traG	Plasmid transfer	51133	51330	F
traG	Plasmid transfer	51288	51557	F
traG	Plasmid transfer	51533	51898	F
traG	Plasmid transfer	51981	52214	F

Table 5. continued

traS	Plasmid transfer	52229	52405	F
traS	Plasmid transfer	52362	52730	F
traT	Conjugation regulation	52779	53513	F
traD	Plasmid transfer	53741	53881	F
traD	Plasmid transfer	53902	54390	F
traD	Plasmid transfer	54462	54806	F
traD	Plasmid transfer	55041	55736	F
traD	Plasmid transfer	55739	55939	F
traI	Plasmid transfer	55939	56157	F
traI	Plasmid transfer	56141	56689	F
traI	Plasmid transfer	56701	56895	F
traI	Plasmid transfer	56892	57134	F
traI	Plasmid transfer	57089	57295	F
traI	Plasmid transfer	57277	57648	F
traI	Plasmid transfer	57692	58336	F
traI	Plasmid transfer	58318	58593	F
traI	Plasmid transfer	58609	58797	F
traI	Plasmid transfer	58797	58940	F
traI	Plasmid transfer	58940	59302	F
traI	Plasmid transfer	59299	59517	F
traI	Plasmid transfer	59504	59767	F
traI	Plasmid transfer	59730	59873	F
traI	Plasmid transfer	59891	60064	F
traI	Plasmid transfer	60061	60192	F
traI	Plasmid transfer	60192	60545	F
traI	Plasmid transfer	60628	60834	F
traX	Plasmid transfer	61261	61533	F
traX	Plasmid transfer	61599	61997	F
finO	Conjugation regulation	61994	62362	F
finO	Conjugation regulation	62398	62550	F
hha	Haemolysin expression modulating protein	63636	63845	F
yihA	GTP-binding	63882	64472	F

Table 5. continued

tnp	Tn3 family transposase	64532	65281	R
tnp	Tn3 family transposase	65278	65904	R
tnp	Tn3 family transposase	65975	66346	R
tnp	Tn3 family transposase	66430	66660	R
tnp	Tn3 family transposase	66657	67010	R
tnp	transposase	67402	68202	F
tnp	transposase	68265	68492	F
lysR	Transcriptional regulator	68615	68740	R
gltS	Sodium/glutamate symporter	68970	69509	R
gltS	Sodium/glutamate symporter	69503	69967	R
tetR	Tetracycline resistance	71317	72015	F
tetR	Tetracycline resistance	71980	72195	R
tetR	Tetracycline resistance	72436	72600	R
tetB	Tetracycline resistance	72664	73182	F
tetB	Tetracycline resistance	73206	73346	F
tetB	Tetracycline resistance	73367	73483	F
tetB	Tetracycline resistance	73491	73901	F
tetR	Tetracycline resistance	73998	74375	R
tetR	Tetracycline resistance	74389	74658	R
tetD	Tetracycline resistance	74671	74973	F
tnp	Tn5 family transposase	75499	75978	R
tnp	Tn3 family transposase	76401	76526	R
hin	DNA convertase	76651	77265	F
strA	Aminoglycoside resistance	78563	79363	F
strB	Aminoglycoside resistance	79363	80195	F
repA	Replication protein	81122	81643	F
repB	Replication protein	81636	81779	F
repA	Replication protein	81776	81940	F
doc	Toxin-antitoxin system	83174	83308	F

Table 5. continued

Colicin-la[b]	Bacteriocin ion-channel formation	84445	84690	F
Colicin-la[b]	Bacteriocin ion-channel formation	84687	84989	F
Colicin-la[b]	Bacteriocin ion-channel formation	85049	85222	F
Colicin[b]	Bacteriocin	85871	86008	F
imm	Colicin immunity	86339	86614	R
-	Phage tail assembly protein	86777	87031	R
tnpB	Transposase IS66 family	87144	87314	F
tnp	Transposase IS66 family	88122	88262	F
tnp	Transposase IS66 family	88256	88894	F
tnp	Transposase IS66 family	88891	89106	F
ydeA	Unknown function	89517	89882	R
ydeA	Unknown function	89873	90106	R
macB	Macrolide export-Efflux system	93399	93797	R
macB	Macrolide export-Efflux system	93910	94218	R
macB	Macrolide export-Efflux system	94194	94649	R
insL	Transposase for insertion element 186	94752	94991	R
cvaA	Colicin V secretion	95600	95821	R
cvaA	Colicin V secretion	95839	96138	R
cvaA	Colicin V secretion	96387	96542	R
doc	Toxin-antitoxin system	97515	97697	F
aroH	Aromatic amino acid biosynthesis	98656	99546	F
iroN	Aerobactin system	100637	100750	F
iroN	Aerobactin system	101020	101298	F
iroN	Aerobactin system	101630	101899	F
iroN	Aerobactin system	101914	102141	F
iroN	Aerobactin system	102323	102697	F

Table 5. continued

iroE	Aerobactin system	102842	102973	R
iroE	Aerobactin system	102973	103398	R
iroE	Aerobactin system	103355	103693	R
iroD	Aerobactin system	103778	103915	R
iroD	Aerobactin system	103990	104535	R
iroD	Aerobactin system	104539	104754	R
iroD	Aerobactin system	104751	104999	R
iroC	Aerobactin system	105103	105792	R
iroC	Aerobactin system	105822	106514	R
iroC	Aerobactin system	106543	106902	R
iroC	Aerobactin system	106889	107156	R
iroC	Aerobactin system	107144	108385	R
iroC	Aerobactin system	108382	108504	R
iroC	Aerobactin system	108521	108751	R
iroB	Aerobactin system	108929	109621	R
iroB	Aerobactin system	109651	109998	R
tnp	Transposase	111226	111519	R
tnp	Transposase	111516	111785	R
insC	Transposase for element IS2	114000	114407	F
insD	Transposase for element IS3	114431	115270	F
-	Metal chaperone-Zn homeostasis	116607	116774	F
-	Metal chaperone-Zn homeostasis	116726	116917	F
-	Metal chaperone-Zn homeostasis	116967	117671	F
insD	Transposase for insertion element IS2	117888	118055	F
tnp	Transposase	118304	118741	F
tnp	Transposase	118834	119067	F
tnp	Transposase	119143	119253	F
tnp	Transposase	119301	119465	F

Table 5. continued

tnp	Transposase	119486	119773	F
tnp	Transposase IS200 family	120952	121179	R
-	Efflux transport system	121647	122000	R
-	Efflux transport system	121981	122271	R
-	Efflux transport system	122381	122533	R
-	Efflux transport system	122679	122810	R
-	Efflux transport system	122810	123010	R
macB	Macrolide export-Efflux system	123089	124846	R
macA	Macrolide export-Efflux system	124948	125103	R
macA	Macrolide export-Efflux system	125064	125324	R
macA	Macrolide export-Efflux system	125324	125563	R
macA	Macrolide export-Efflux system	125614	126153	R
insO	Transposase for insertion element 911	127022	127312	F
tnp	Transposase IS3 family	127583	128047	R
insK	Transposase	128084	128470	R
tnp	Transposase IS3 family	128451	128741	R
traI	Plasmid transfer	128696	128968	F
ompT	protease	130839	131048	F
ompT	protease	131014	131340	F
mig-14	Transcription activator	133043	133168	F
mig-14	Transcription activator	133297	133755	F
rec	Resolvase	134521	134877	F
rec	Resolvase	134859	135251	F
repFIB	Replication protein	135520	136329	R
repFIB	Replication protein	136363	136668	R
tnp	Transposase	137270	137674	F
umuC	SOS response	137774	138178	F
pndC	Plasmid stability protein	138245	138421	R

Table 5. continued

insB	Transposase- IS1 family	138577	138798	R
insA	Transposase- IS1 family	138818	139117	R
ssb	SOS response	139721	139879	F
sitB	Aerobactin system	139892	140098	F
sitB	Aerobactin system	140095	140427	F
sitC	Aerobactin system	140396	140659	F
sitC	Aerobactin system	140831	141172	F
sitD	Aerobactin system	141245	141562	F
sitD	Aerobactin system	141544	142095	F
eno	Enolase	142561	142995	F
crcB	Aerobactin system	143332	143352[illegible]	F
nhaA	H+- antiporter	144085	144201	R
nhaA	H+- antiporter	144278	145276	R
iucA	Aerobactin system	145412	145813	F
iucA	Aerobactin system	145810	146139	F
iucA	Aerobactin system	146176	146883	F
iucB	Aerobactin system	147311	147607	F
iucB	Aerobactin system	147604	147879	F
iucB	Aerobactin system	147876	148070	F
iucC	Aerobactin system	148070	148852	F
iucC	Aerobactin system	148785	149141	F
iucC	Aerobactin system	149096	149410	F
iucC	Aerobactin system	149398	149688	F
iucD	Aerobactin system	149790	150305	F
iucD	Aerobactin system	150277	150426	F
iucD	Aerobactin system	150563	150718	F
iucD	Aerobactin system	150718	151068	F
iutA	Aerobactin system	151417	151623	F
iutA	Aerobactin system	151620	151853	F
iutA	Aerobactin system	152023	152250	F
iutA	Aerobactin system	152581	153009	F
insA	IS1 transposase	153858	154157	F
insB	IS1 transposase	154159	154398	F

Table 5. continued

insA	IS1 transposase	156765	157064	F
insB	IS1 transposase	157084	157416	F
parA	Plasmid partitioning	157748	158155	F
parA	Plasmid partitioning	158165	158737	F
parA	Plasmid partitioning	158745	158918	F
parB	Plasmid partitioning	158918	159187	F
parB	Plasmid partitioning	159243	159671	F
parB	Plasmid partitioning	159771	159884	F
tnp	Transposase	162218	162400	R
tnp	Transposase	162534	162815	R
tnp	Transposase	162809	163099	R
klcA	Antirestriction protein	163524	163649	F
KklcA	Antirestriction protein	163646	163948	F
ssb	DNA replication, recombination, & repair	168298	168723	F
ssb	DNA replication, recombination, & repair	168739	169056	F
parB	Plasmid partitioning	169777	170349	F
psiB	Plasmid SOS inhibition	171478	171723	F
psiB	Plasmid SOS inhibition	171740	171913	F
psiA	Plasmid SOS inhibition	171910	172053	F
psiA	Plasmid SOS inhibition	172340	172627	F
tnp	Transposase	172701	172823	R
traM	Plasmid transfer	176409	176738	F
traJ	Plasmid transfer	176920	177066	F
traJ	Plasmid transfer	177063	177191	F

[a] distance from position 1 of repeat region 1.
[b] protein product

Discussion

Exogenous plasmid capture

The large number of transconjugants conferring tetracycline resistance on an *E. coli* host strain used for their capture indicated the presence of self-transmissible antibiotic resistance plasmids in poultry litter. One caveat to selecting plasmids using tetracycline is that only plasmids conferring resistance to tetracycline could be detected, however tetracycline has been routinely used in the poultry industry both as a growth promoter and for prophylactic and therapeutic use (40) and these results suggest that transmissible Tet^R plasmids may be common in poultry litter and, therefore, in and on the poultry themselves.

All transconjugants were resistant to tetracycline, as expected, since tetracycline was used to select the transconjugants. However, the resistance phenotyping results for the other antibiotics were somewhat unexpected. Past experiments in the Herrick laboratory have shown that plasmids carrying tetracycline resistance genes usually also carry streptomycin and gentamicin resistances (13). However, of the twelve transconjugants analyzed in this study, only pEH11 was resistant to streptomycin and all transconjugants were susceptible to gentamicin. More importantly, resistance to a number of (human) clinically-relevant antibiotics was observed. Phenotypic resistance to aztreonam, ceftazidime, and ciproflaxocin are not typically observed on plasmids transmitted in agriculture reservoirs (13). Aztreonam is a monobactam antibiotic that is used to treat persistent respiratory infections that are resistant to first-line antibiotics (41). Ceftazidime is a third generation, broad-spectrum cephalosporin antibiotic typically reserved for use in ICUs with patients that have otherwise antibiotic resistant infections (42-43), and ciproflaxocin is a second-generation fluoroquinolone antibiotic. Each of these antibiotics is typically used in human clinical settings.

Although tetracycline is an antibiotic that is used as an agricultural growth promoter and prophylactic agent, the use of tetracycline in agricultural environments may add selective pressure that encourages the transmission of mobile genetic elements harboring multiple antibiotic resistant genes. Poultry livestock production may provide a reservoir for clinically relevant antibiotic resistance, even for antibiotics not directly used on poultry.

LA61^{RifR} genome assembly

QUAST data conveyed that the assembly of LA61^{RifR} was good. The quality of the assembly was determined by there being less than 600 contigs, a number recommended by the CDC assembly thresholds for *Escherichia* (44). In addition, the assembly in its entirety was 5,128,965 bp, the correct size for an *E. coli* genome. The depth of coverage of the LA61^{RifR} sequence, 35.5X, exceeded the 30X depth of coverage associated with good sequence quality. The N_{50}, or the median length of contigs, was greater than 200,000, which is also considered to be acceptable.

Ideally, all contigs would be assembled into one single contig that would circularize, which is rarely if ever achieved using only short-reads. In order to decrease the number of contigs and improve the assembly quality, LA61^{RifR} should be long-read sequenced on the ONT MinION and a hybrid assembly should be generated using Unicycler. The combination of Illumina and Oxford Nanopore sequencing improves the sequence quality as well as assembly quality.

pEH11 assembly and annotation

Comparison of the hybrid short-read LA61^{RifR} and long-read LA61^{RifR}::pEH11 Unicycler assembly to the long-read only LA61^{RifR}::pEH11 Canu assembly shows that using the short-read LA61^{RifR} chromosome lacking the plasmid (pEH11) to generate a hybrid assembly did not improve the assembly quality. This is seen by the greater number of nodes, 17, and contigs, 9, in the Unicycler assembly. In addition, NCBI BLASTn did not identify any hybrid contigs that matched a plasmid. This was unexpected due to the fact that, generally, using both short and long-read data improves the genome assembly (16).

In addition to a low number of contigs, the Canu assembly of pEH11 had a N_{50} of 5,183,052bp, a complete length of 5,372,641 bp, and a GC% of 50.61%. The length and GC content were consistent with an *E.coli* genome.

The Mauve alignment (Figure 10) of pEH11 (Contig #2) and its three top NCBI BLASTn results (Table 4) show that there is >99% similarity between pEH11 and pSSAP03302A, pCVM29188_146, and pCS0010A. However, pEH11 is a longer plasmid by ~30,000 bp. It is possible that this region was inaccurately assembled into the plasmid due to common repeat regions. However, by looking at the positions of this inserted region on the Mauve graph and searching for homologous sequences using NCBI BLASTn, the same three *Salmonella enterica* serovar Kentucky plasmids were the top results. This could be caused by the repeat regions within that extra region of the genome, 166,410-21,328. Therefore, it is also possible that this region was correctly assembled and that it is part of a genetic element that was repeated and integrated into a new part of the plasmid genome, as indicated in Figure 11.

Annotation of the pEH11 genome provided insight into the plasmid's potential to add to the virulence of a host bacterium. Output from RAST, an auto-annotation tool, categorized the predicted genes to display key functions of the plasmid (Figure 12). It's important to note the

high prevalence of iron acquisition and metabolism protein-coding genes as well as (especially) membrane transport protein genes. Closer investigation of the plasmid revealed the two aerobactin systems, consistent with iron acquisition. The reason the membrane transport protein category was so large may be because RAST includes efflux systems and plasmid replication and transfer proteins within this category and these genes collectively occupied a large proportion of the plasmid genome.

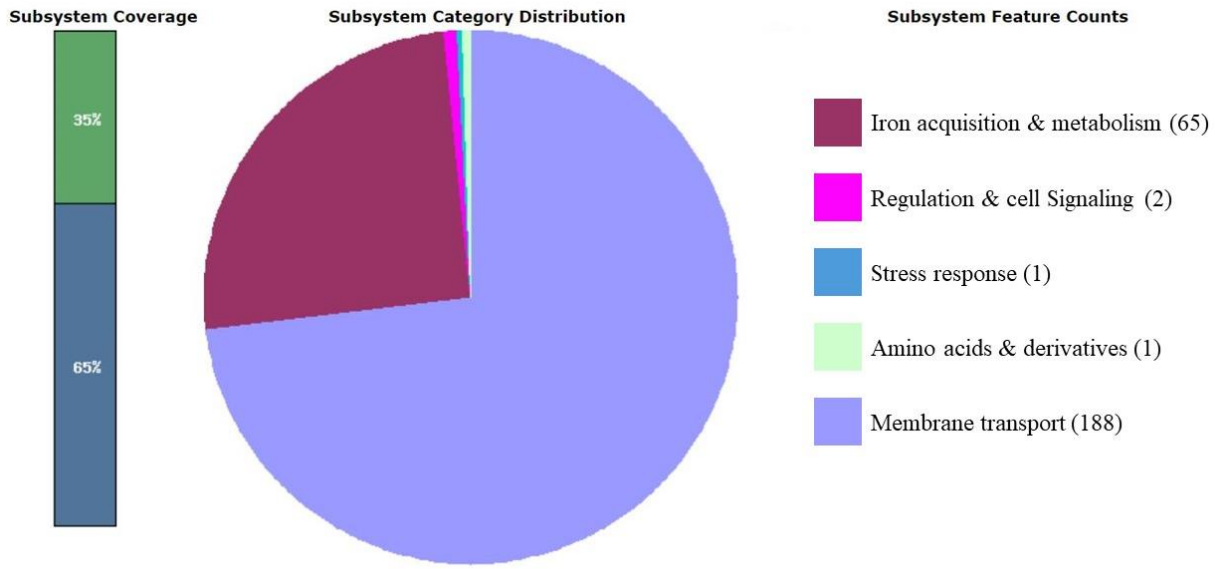

Figure 12. Categorical distribution of predicted genes on plasmid pEH11 generated by RAST and modified. Subsystem coverage identifies the percentage of features that are within a subsystem. A subsystem is the set of functional roles that contribute to a biological process (45).

On pEH11 there are several regions that are responsible for plasmid replication and transfer. Genes specific for replication and transfer are the *tra* genes as well as *repFIB*, *repA*, *repB*, and *psiB*. The pilin gene, *traA* encodes a pilus, which plays a critical role in plasmid conjugation and provides further evidence that pEH11 is transmissible. The plasmid partitioning genes *parA*, *parB*, *psiA*, *psiB*, *ssb*, and *umuC* enable the stable transfer of plasmid DNA. These genes include genes responsible for SOS inhibition, which prohibits transfer of damaged DNA.

The conjugation regulator genes *ylpA* and *finO* regulate the transfer of a replicated plasmid into a new host cell. Regulation of plasmid conjugation is influenced by plasmid replication genes like *repFIB*, which indicates the IncF incompatibility group (46). This is additionally supported by the IncF incompatibility grouping of all three of the top BLASTn matches to pEH11, pCS0010A, pSSAP03302A, and pCVM29188_146.

The gene annotation results produced by *Prokka* and RAST provided insight into the potential virulence that pEH11 contributes to its host cell. Three regions contain genes encoding aerobactin systems. These plasmid-associated genes are commonly found in uropathogenic *E. coli* that live in iron-depleted environments (47). Genes *iuc*, *iut*, and *iro* are associated with iron-uptake chelate and transport. The length of each of these regions is comparable to the length of the iron-regulated aerobactin operon found on the ca. 8Kbp plasmid ColV-K30p (47), suggesting that each of these aerobactin system regions are complete operons.

Additionally, there are a number of genes on pEH11 that can contribute to its host cell's antibiotic resistance. One region of the plasmid contains genes *tetR*, *tetB*, and *tetD*, which are responsible for LA61^{RifR}::pEH11's tetracycline resistant phenotype. The presence of these genes was not surprising because tetracycline was used to select ARG-carrying plasmids during the exogenous plasmid capture, from which LA61^{RifR}::pEH11was a transconjugant. Two additional ARGs were discovered that explain LA61^{RifR}::pEH11's phenotypic streptomycin resistance. Streptomycin is an aminoglycoside antibiotic, which acts by binding to the 30S or 50S ribosomal subunits and prohibiting bacterial protein synthesis (48). Two genes, *strA* and *strB*, encode an aminoglycoside acetyltransferase which enzymatically modifies the aminoglycoside antibiotic (49), rendering the drug ineffective. The presence of these two genes explain the phenotypic resistance to streptomycin observed in LA61^{RifR}::pEH11. However, these genes did not provide

resistance to additional aminoglycoside antibiotics that were tested during the Stokes tests like gentamicin and tobramycin. LA61^{RifR}::pEH11 was phenotypically resistant to aztreonam, however, there were no corresponding monobactam resistance genes identified on the plasmid chromosome. One possible explanation is that the monobactam resistance gene responsible is a new or uncharacterized gene. Another possible explanation is that the gene is sufficiently divergent from previously identified aztreonam resistance genes and was not recognized.

In addition, two regions of pEH11 contain *macA* and *macB* genes that encode a macrolide efflux system. Macrolides such as erythromycin are a class of antibiotics that prohibit protein synthesis by binding and inhibiting the bacterial 50S ribosomal subunit (50). Efflux pumps enable bacteria to export toxic substances out of the cell, thus disabling drugs and other toxins. LA61pEH11^{RifR}::pEH11 was not tested for resistance to macrolides. It would be interesting to determine the susceptibility of LA61^{RifR}::pEH11's macrolide drugs. Regardless of the resistance phenotype, the presence of these genes on plasmid pEH11 demonstrate the presence of ARGs able to be transferred between bacteria in a specific reservoir.

Plasmid pEH11 also contains genes associated with the production of bacteriocins. Bacteriocins are toxic proteins that a bacterium can secrete in order to outcompete similar strains in competition for space and nutrients. There is a region on pEH11 that is associated with colicin, which is a specific bacteriocin produced by coliform bacteria like *E. coli*. This region included protein products *colicin-la* and *colicin* as well as genes *cvaA*, and *imm*. This region is located in a cluster between positions 84,445 and 96,139. Each of these genes encodes a different part of the system needed to export a bacteriocin. The *Colcin-la* protein forms an ion channel that is necessary for bacteriocin export, *colicin* encodes the bacteriocin itself, *cvaA* contributes to colicin V secretion, and *imm* is a gene associated with colicin immunity. The gene *imm* does

provide immunity for the cell secreting the bacteriocin, preventing colicin from having a suicidal effect on the host cell.

Under typical low-stress conditions, it costs a host cell energy to replicate and maintain a plasmid (51). When there is no selective-pressure (like an antibiotic) in its environment, a host bacterium has no reason to maintain a plasmid, especially a large one like pEH11. In order to maintain themselves, plasmids often encode a toxin/antitoxin (TA) system. The toxin will accumulate in a cell and lead to cell death without the presence of an antitoxin to suppress it. The gene *flmC* is associated with a Type I TA system on IncF plasmids (52). The essential genes to the system are *flmA* and *flmB*, which were not identified on IncF plasmid pEH11. However, there were many unknown coding (CDS) regions of pEH11, which could encode *flmA* and *flmB*, thus completing the TA system. The presence of the *flmC* gene also points toward the presence of TA systems on related plasmids, which may have interacted and exchanged genetic material with pEH11 ancestors. This is an example of gene mobility on and between plasmids in a host cell. One other gene annotated on pEH11, *doc*, points towards the presence of a TA system. *Doc* toxins work by stabilizing mRNA which leads to translation arrest, however the exact mechanism is unknown (53). Like the *flm* TA system, it is unknown whether all components of the *doc* TA system are encoded on pEH11, perhaps due to many unidentified CDS regions. Many antitoxins are also noncoding RNAs (54) which were not analyzed on pEH11and would not have been annotated by Prokka, RAST, or NCBI BLASTn.

Interestingly, there was a bacteriophage tail assembly protein encoded on plasmid pEH11. The presence of this gene points to the presence of a prophage in host cells of pEH11 ancestors. This gene could have been mobilized to plasmid by a transposon. This is supported by the nearby presence of several transposases on the pEH11 genome.

In addition to pEH11 being a MGE in itself, capable of transferring virulence and genetic material between cells, there is genetic evidence of MGEs within the pEH11 genome. In total there were 47 transposase genes on pEH11. The effort to identify a more accurate tool for predicting transposons was unsuccessful. However, the presence of transposase genes for insertion sequences and transposons provides evidence for these elements and thus for the movement of genetic material between pEH11 and other plasmids or host cell genomes. There was no evidence of integrons on pEH11.

Conclusion

This study demonstrates that the use of tetracycline, an antibiotic commonly used in the Shenandoah Valley agriculture industry, provides selective pressure for many additional ARGs and virulence factors carried on transmissible plasmids. The use of tetracycline may therefore contribute to the role poultry litter and other large-scale poultry-farming practices play as reservoirs for antibiotic resistance and virulence genes transferrable by horizontal gene transfer. These genes, exhibited by their presence on plasmid pEH11, may increase the host's ability to survive in stressful environments that contain antibiotics and/or high levels of competition from other bacteria. Selection for plasmids encoding tetracycline resulted in the co-selection of multiple antibiotic resistance genes, some of which encoded resistance to antibiotics that are clinically-relevant in human healthcare.

Future directions

The ability to assemble pEH11 from whole genome sequencing data removes the additional time and resources needed to extract and purify plasmid DNA. However, it was still necessary to isolate and purify LA61^{RifR}::pEH11 from the transconjugant plate. In further studies, it would be beneficial to develop a pipeline for identifying MGEs and virulence genes, harbored on plasmids, that are transferred in reservoirs like the Shenandoah poultry industry. The method of assembling multiple plasmids from whole genome data sequenced using one barcoding tag, is referred to as a "metaplasmidome".

As proof of principle, an artificial metaplasmidome should first be completed with known plasmids. Whole genome extraction and sequencing of transconjugants containing known and previously-sequenced plasmids were completed as part of this study. Assembly of these plasmids, electroporated into commercial *E. coli*, EC100, should be attempted. If assembly is unsuccessful, a pipeline of bioinformatics tools to analyze the virulence genes and MGEs, transferable within one reservoir, will be attempted and compared to known virulence genes from previous individual plasmid extractions and assemblies (17, 19).

If successful, this proof of principle will be applied to a real metaplasmidome. The real metaplasmidome will be the single analysis of all transconjugants on one plate from an exogenous plasmid capture. The plate would be washed, the whole transconjugant genome sequenced as one extraction, and virulence genes present on plasmids analyzed bioinformatically.

If the idea of a metaplasmidome remains unsuccessful, analysis of other plasmids captured in the previous exogenous plasmid capture (pEH1-12, excluding pEH11), could be assembled and analyzed from whole genome or plasmid genome sequencing data.

References

1. Nordmann P, Naas T, Fortineau N, Poirel L. 2007. Superbugs in the coming new decade; multidrug resistance and prospects for treatment of *Staphylococcus aureus*, Enterococcus spp. and *Pseudomonas aeruginosa* in 2010. Curr Opin Microbiol 10:436-40.

2. McEwen SA, Fedorka-Cray PJ. 2002. Antimicrobial use and resistance in animals. Clin Infect Dis 34:S93–S106.

3. Hollis A, Ahmed Z. 2013. Preserving antibiotics, rationally. N Engl J Med 369:2474–2476.

4. Kapsak C. 2018. Isolation, sequencing, and characterization of four transmissible antibiotic resistance plasmids captured from bacteria in stream sediments. Masters Thesis James Madison University. https://commons.lib.jmu.edu/master201019/577/

5. Tortora GJ, Funke BR, Case CL. 2016. Microbial Genetics, p. 227-263. *In* Tortora, 12 (ed.), Microbiology: An Introduction. Pearson Education, London.

6. Deng Y, Bao X, Ji L, Chen L, Liu J, Miao J, Chen D, Bian H, Li Y, Yu G. 2015. Resistance integrons: class 1, 2 and 3 integrons. Ann Clin Microbiol Antimicrob..

7. Bouvier M, Demarre G, Mazel D. 2005. Integron cassette insertion: A recombination process involving a folded single strand substrate. EMBO J 24:4356–4367.

8. Johnson CM, Grossman AD. 2015. Integrative and conjugative elements (ICEs): what they do and how they work. Annu Rev Genet 49:577–601.

9. Salyers AA, Shoemaker NB, Stevens AM, Li L-Y. 1995. Conjugative transposons: an unusual and diverse set of integrated gene transfer elements. microbiological reviews 59: 579–590.

10. Bertelli C, Tilley KE, Brinkman FSL. 2019. Microbial genomic island discovery, visualization and analysis. Brief Bioinform 20:1685–1698.

11. Chopra I, Roberts M. 2001. Tetracycline antibiotics: mode of Action, applications, molecular biology, and epidemiology of bacterial resistance. Microbiol Mol Biol Rev 65:232–260.

12. Speer BS, Shoemaker NB, Salyers AA. 1992. Bacterial resistance to tetracycline: mechanisms, transfer, and clinical significance. Clinical Microbiology Review 5(4): 387–399.

13. Herrick JB, Haynes R, Heringa S, Brooks JM, Sobota LT. 2014. Coselection for resistance to multiple late-generation human therapeutic antibiotics encoded on tetracycline resistance plasmids captured from uncultivated stream and soil bacteria. J Appl Microbiol 117:380–389.

14. Mikheyev AS, Tin MMY. 2014. A first look at the Oxford Nanopore MinION sequencer. Mol Ecol Resour 14:1097–1102.

15. Smits THM. 2019. The importance of genome sequence quality to microbial comparative genomics. BMC Genomics 20:662.

16. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Computational Biology 13: e1005595.

17. Gehr E. 2013. The potential for replication and transmission of antibiotic resistance plasmids in an E. coli population in agriculturally impacted stream sediment. Masters Thesis, James Madison University..

18. Szybalski W, Bryson V. 1952. Genetic studies on microbial cross resistance to toxic agents. I. Cross resistance of *Escherichia coli* to fifteen antibiotics. J Bacteriol 64:489–499.

19. Libuit K. 2016. Next-generation sequencing of a multi-drug resistance plasmid captured from stream sediment. Masters Thesis, James Madison University.

20. 2006. DNeasy ® Blood & Tissue Handbook. http://www.bea.ki.se/documents/EN-DNeasy%20handbook.pdf

21. Wick RR, Judd LM, Holt KE. 2019. Performance of neural network basecalling tools for Oxford Nanopore sequencing. Genome Biol 20:129-139.

22. Afgan E, Baker D, Batut B, Van Den Beek M, Bouvier D, Ech M, Chilton J, Clements D, Coraor N, Grüning BA, Guerler A, Hillman-Jackson J, Hiltemann S, Jalili V, Rasche H, Soranzo N, Goecks J, Taylor J, Nekrutenko A, Blankenberg D. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res 46:W537–W544.

23. Babraham Bioinformatics - FastQC A quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

24. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

25. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin A V., Sirotkin A V., Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477.

26. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075.

27. Wick R, Schultz M, Zobel J, Holt K. 2015. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics 31:3350–3352.

28. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive κ-mer weighting and repeat separation. Genome Res 27:722–736.

29. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410.

30. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. Journal of Antimicrobial Chemotherapy 67:2640–2644.

31. Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, Rolain JM. 2014. ARG-annot, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob Agents Chemother 58:212–220.

32. Jia B, Raphenya A, Alcock B, Waglechner N, Guo P, Tsang K, Lago B, Dave B, Pereira S, Sharma A, Doshi S, Courtot M, Lo R, Williams L, Frye J, Elsayegh T, Sardar D, Westman E, Pawlowski A, Johnson T, Brinkman F, Wright G, McArthur A. 2017. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic database. resistance. Nucleic Acids Res 45:566–573.

33. Naas T, Oueslati S, Bonnin RA, Dabos ML, Zavala A, Dortet L, Retailleau P, Iorga BI. 2017. Beta-lactamase database (BLDB)–structure and function. J Enzyme Inhib Med Chem 32:917–919.

34. Carattoli A, Zankari E, Garciá-Fernández A, Larsen MV, Lund O, Villa L, Aarestrup FM, Hasman H. 2014. *In silico* detection and typing of plasmids using Plasmidfinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother 58:3895–3903.

35. Chen L, Zheng D, Liu B, Yang J, Jin Q. 2016. VFDB 2016: Hierarchical and refined dataset for big data analysis--10 years on - PubMed. Nucleic Acids Res 44:694–697.

36. Seeman T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.

37. Cury J, Jove T, Touchon M, Neron B, Rocha EP. 2016. Identification and analysis of integrons and cassette arrays in bacterial genomes. Nucleic Acids Research 44:4539–4550.

38. Mareuil F, Doppelt-Azeroual O and Ménager H. 2017. A public Galaxy platform at Pasteur used as an execution engine for web services. *F1000Research* **6**:1030 (poster).

39. Overbeek R, Olson R, Pusch GD, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the rapid annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res 42:206–214.

40. Roberts MC, Schwarz S. 2016. Tetracycline and phenicol resistance genes and mechanisms: importance for agriculture, the environment, and humans. J Environ Qual 45:576–592.

41. Andrews R, Fasoli R, Scoggins WG, Algozzine GJ, Spann RW, Sundaresh KV, Mathers JA Jr., Babb R 3rd, Kuppinger M, Cooper B. 1994. Combined aztreonam and gentamicin

therapy for pseudomonal lower respiratory tract infections. Clinical Therapeutics 16:236-252.

42. Carmeli Y, Armstrong J, Laud PJ, Newell P, Stone G, Wardman A, Gasink LB. 2016. Ceftazidime-avibactam or best available therapy in patients with ceftazidime-resistant Enterobacteriaceae and *Pseudomonas aeruginosa* complicated urinary tract infections or complicated intra-abdominal infections (REPRISE): a randomised, pathogen-directed, phase 3 study. Lancet Infect Dis 16:661–673.

43. Richards DM, Brogden RN. 1985. Ceftazidime. A review of its antibacterial activity, pharmacokinetic properties and therapeutic use. Drugs 29:105–161.

44. Schroeder M. 2019. Assessment and interpretation of sequence quality in bionumerics. Presentation.

45. Overbeek R, Begley T, Butler RM, Choudhuri J V., Chuang H-Y, Cohoon M, Crecy-Lagard V de, Diaz N, Disz T, Edwards R, Fonstein M, Frank ED, Gerdes S, Glass EM, Goesmann A, Hanson A, Iwata-Reuyl D, Jensen R, Jamshidi N, Krause L, Kubal M, Larsen N, Linke B, McHardy AC, Meyer F, Neuweger H, Olsen G, Olson R, Osterman A, Portnoy V, Pusch GD, Rodionov DA, Ruckert C, Steiner J, Stevens R, Thiele I, Vassieva O, Ye Y, Zagnitko O, Vonstein V. 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res 33:5691–5702.

46. Thomas CM. 2014. Plasmid incompatibility, p. 1–3. *In* Molecular Life Sciences. Springer New York, New York, NY.

47. De Lorenzo V, Bindereif A, Paw BH, Neilands JB. 1986. Aerobactin biosynthesis and transport genes of plasmid colV-K30 in *Escherichia coli* K-12. J Bacteriol 165:570–578.

48. Krause KM, Serio AW, Kane TR, Connolly LE. 2016. Aminoglycosides: an overview. Cold Spring Harb Perspect Med 6.

49. Vetting MW, Chi HP, Hegde SS, Jacoby GA, Hooper DC, Blanchard JS. 2008. Mechanistic and structural analysis of aminoglycoside N-acetyltransferase AAC(6′)-Ib and its bifunctional, fluoroquinolone-active AAC(6′)-Ib-cr variant. Biochemistry 47:9825–9835.

50. Dinos GP. 2017. The macrolide antibiotic renaissance. British Journal of Pharmacology. 17: 2967–2983.

51. Ponciano JM, De Gelder L, Top EM, Joyce P. 2007. The population biology of bacterial plasmids: A hidden Markov model approach. Genetics 176:957–968.

52. Loh SM, Cram DS, Skurray RA. 1988. Nucleotide sequence and transcriptional analysis of a third function (Flm) involved in F-plasmid maintenance. Gene 66:259–268.

53. Cruz JW, Rothenbacher FP, Maehigashi T, Lane WS, Dunham CM, Woychik NA. 2014. Doc toxin is a kinase that inactivates elongation factor Tu. J Biol Chem 289:7788–7798.

54. Unterholzner SJ, Poppenberger B, Rozhon W. 2013. Toxin–antitoxin systems. Mob Genet Elements 3:e26219.