The Australian Group on Antimicrobial Resistance

http://antimicrobial-resistance.com

2020 Australian Group on Antimicrobial Resistance (AGAR) Surveillance Programmes – Additional Projects, GNSOP.

TRIM: D18-25786

04 June 2021 to 13 May 2022

Final Report May 2022

GNSOP, Project 2: Understanding AMR plasmids in Gram-negative bacteria in Australia.

S. PARTRIDGE[1,2,3], A. FAJARDO LUBIAN[1,2], J. DRAPER[4], E. MARTINEZ[4], J. IREDELL[1,2,3], on behalf of AGAR GNSOP

[1]Sydney Medical School, The University of Sydney, Sydney, Australia.
[2]The Westmead Institute for Medical Research (WIMR), Westmead, Australia.
[3]Centre for Infectious Diseases and Microbiology, Westmead Hospital, Western Sydney Local Health District (WSLHD), Westmead, Australia.
[4]Centre for Infectious Diseases and Microbiology, Laboratory Services, NSW Health Pathology, Westmead, Australia.

Identifying the source and spread of infection is essential in understanding the transmission of antimicrobial resistance (AMR) in Gram-negative bacteria. AMR in these organisms is largely due to acquisition and spread of mobile genes that each cause resistance to a subset of antibiotics. AMR genes, especially those that cause failure of the main hospital antibiotics in the most common pathogens (e.g., *E. coli* and *Klebsiella*), are largely spread on genetic vehicles (plasmids) that travel between bacteria, including different species.

Incorrect assignment of a gene transmission event to a plasmid and/or failure to recognise the relatedness of plasmids is a major barrier to infection management and control. Short-read WGS is sufficient to identify AMR genes but often cannot accurately assign an AMR gene to particular plasmid: resistance plasmids typically contain many long repeats, particularly in regions containing AMR genes, and the same components are often found in different plasmids in the same cell. This means that assembly of complete plasmid sequences directly from short read data is almost impossible.

Aim. to generate complete accurate assemblies of the most important plasmids in AMR Gram-negative pathogens in Australian hospitals.

Materials and Methods.

1. Isolates used in this study. 48 Gram-negative isolates with relevant phenotype (resistance to third generation cephalosporins and or carbanepem resistance) and specific plasmid content were selected, based on our short read WGS analysis of 2020 GNSOP collection to perform long read sequencing (PacBio Sequel II HiFi).

2. *Long read WGS (PacBio Sequel II HiFi).* Total DNA was extracted using Quick-DNA HMW MagBead kit (Zymo Research) following manufacturer's instructions, with some modifications. Long read WGS is being performed by AGRF Brisbane. Data will be analysed using Use SMRT Link v11.0 Microbial Genome Analysis application (HiFi software v11.0).

3. Manual analysis of short read data. Ambiguities, potential multiple gene copies/variants and gene location (plasmid or chromosomal) were checked manually (Geneious software). Bandage was used for visualization of the connections between assembled contigs [1]. For this work we have designed a workflow to define antibiotic resistance gene location. Briefly, these steps were followed for each individual resistance gene and isolate: **1)** identify and annotate the contig where the resistance gene is located by Geneious and GAMR (https://galileoamr.arcbio.com/mara), **2)** identify location of important genetic features (such as insertion sequences: IS*26*, IS*Ecp1*), **3)** map reads to resistance genes and insertion sequences to decipher the presence of multiple copies and/or different variants, **4)** look for flanking regions of insertion sequences and resistance genes and blast against publicly available genomes (NCBI) to identify antibiotic resistance gene context and location.

Results.

1.Analysis of short read WGS. For this study we used short-read sequencing data corresponding to routine WGS from GNSOP 2020 [2] and extra project "GNSOP 1, Understanding how important AMR genes are becoming fixed in the chromosome". Isolates (n=700) were examined followed the steps described in section 3 of Material and Methods. **Table 1** summarizes the list of isolates selected for long read sequencing, based on their AMR gene profile and plasmid content. As seen in **Table 1**, Some isolates carry more than one copy of the same AMR gene or different AMR genes. Isolates that **[1]** the location of the AMR gene could not be identified ("Unknown", in Table 1) or **[2]** the transposon unit inserted on the chromosome or plasmid sequences could not be completed by short read sequencing and manual analysis were selected for PacBio long read sequencing.

Table 1. Isolated selected for long read PacBio Sequel II HiFi

Isolate	Specie	ST*	AMR gene	Plasmids	Comments / Location (chromosomal, plasmid, unknown)**
0736	*E. coli*	ST38	$bla_{\text{CTX-M-14a}}$	IncF	seven copies of AMR gene, chromosomal
0005	*E. coli*	H41	$bla_{\text{CTX-M-27}}$	IncF	unknown
0035	*K. pneumoniae*	ST20	$bla_{\text{IMP-4}}$	IncHI2 and IncM	Plasmid
0073	*E. coli*	ST648	$bla_{\text{CTX-M-14a}}$, $bla_{\text{DHA-1}}$	IncF	Chromosomal ($bla_{\text{CTX-M-14a}}$), unknown ($bla_{\text{DHA-1}}$)
0075	*E. coli*	H30Rx	$bla_{\text{CTX-M-15}}$	IncF, IncY	two copies of AMR gene, unknown
0115	*E. coli*	ST12	$bla_{\text{CTX-M-14a}}$	none identified	at least two copies of AMR gene, chromsomal possibly
0144	*K. pneumoniae*	ST15	$bla_{\text{NDM-1}}$, $bla_{\text{CTX-M-15}}$	IncNDM-MAR, IncA/C	plasmid ($bla_{\text{NDM-1}}$), chromosomal ($bla_{\text{CTX-M-15}}$)
0175	*E. coli*	ST12	$bla_{\text{CTX-M-15}}$	IncF	Chromosomal
0182	*K. pneumoniae*	ST436	$bla_{\text{OXA-48}}$	IncF	unknown
0195	*E. coli*	ST648	$bla_{\text{OXA-48}}$	IncF	unknown
0209	*E. coli*	H30Rx	$bla_{\text{CTX-M-15}}$	IncF, IncX1	probably more than one copy of AMR gene, unknown
0232	*E. coli*	ST648	$bla_{\text{CTX-M-15}}$	IncF	chromosomal
0233	*E. coli*	H30	$bla_{\text{CTX-M-14a}}$	IncF	at least two copies of AMR gene, one copy in chromsome possibly
0255	*E. coli*	H30Rx	$bla_{\text{CTX-M-15}}$	IncF, IncI1	at least two copies of AMR gene, unknown
0268	*E. coli*	ST405	$bla_{\text{NDM-5}}$, $bla_{\text{CMY-2}}$	IncF	unknown
0269	*E. coli*	ST12	$bla_{\text{CTX-M-14a}}$	IncF, IncB/O	chromosomal
0297	*E. coli*	H41	$bla_{\text{CTX-M-27}}$	IncF, IncB/O	unknown
0331	*E. cloacae complex*	ST125	$bla_{\text{IMP-4}}$	IncF	plasmid
0341	*E. coli*	ST38	$bla_{\text{CTX-M-27}}$	IncF	Plasmid
0362	*E. coli*	ST38	$bla_{\text{CTX-M-143}}$	IncF	Plasmid
0374	*E. coli*	H30Rx	$bla_{\text{CTX-M-15}}$	IncF	unknown
0402	*E. coli*	ST38	$bla_{\text{NDM-5}}$, $bla_{\text{CTX-M-14}}$	None identified	unknown ($bla_{\text{NDM-5}}$), chromosomal ($bla_{\text{CTX-M-14}}$)
0410	*E. coli*	H30Rx	$bla_{\text{CTX-M-15}}$, $bla_{\text{CTX-M-127}}$	IncF	unknwon
0418	*E. coli*	H30Rx	$bla_{\text{CTX-M-15}}$	IncF	unknown
0422	*E. coli*	ST205	$bla_{\text{OXA-181}}$	IncF	unknown
0445	*E. coli*	ST38	$bla_{\text{CMY-2}}$	IncF	unknown
0477	*E. coli*	H30Rx	$bla_{\text{CTX-M-15}}$	IncF	unknown
0487	*E. coli*	ST38	$bla_{\text{CTX-M-14a}}$	none identified	one copy chromosomal and copy unknown
0503	*E. coli*	ST648	$bla_{\text{CTX-M-15}}$, $bla_{\text{DHA-1}}$	IncF	Chromosomal ($bla_{\text{CTX-M-15}}$), unknown ($bla_{\text{DHA-1}}$)
0542	*E. coli*	ST1193	$bla_{\text{CTX-M-15}}$	IncF	chromosomal
0569	*E. coli*	ST1193	$bla_{\text{CTX-M-27}}$	IncF	unknown
1272	*S. marcescens*	-	$bla_{\text{IMP-4}}$	IncF, IncA/C	unknown
0706	*K. pneumoniae*	ST15	$bla_{\text{CTX-M-15}}$	IncF	one copy chromosomal and copy unknown
0708	*E. coli*	H30Rx	$bla_{\text{CTX-M-14a}}$	none identified	chromosomal
0729	*E. coli*	H41	$bla_{\text{CTX-M-15}}$	IncF	unknown
0859	*E. coli*	H41	$bla_{\text{CTX-M-27}}$	IncF	unknown
0866	*E. cloacae complex*	ST365	$bla_{\text{NDM-1}}$, $bla_{\text{CTX-M-9}}$	IncF, IcnHI2	unknwon

0898	*E. coli*	H41	*bla*$_{CTX-M-27}$	IncF	unknown
0903	*E. coli*	H41	*bla*$_{CTX-M-27}$	IncF	unknown
0935	*E. coli*	H41	*bla*$_{CTX-M-27}$	IncF	unknown
0996	*E. coli*	ST12	*bla*$_{CTX-M-15}$	IncF	chromosomal
1017	*E. coli*	H30	*bla*$_{CTX-M-14a}$	IncF	four copies of AMR gene, chromosomal
1037	*E. coli*	ST38	*bla*$_{CTX-M-27}$	IncF	unknown
1108	*E. coli*	ST648	*bla*$_{CTX-M-14}$	IncF	chromosomal
1129	*E. coli*	H30Rx	*bla*$_{CTX-M-15}$	none identified	unknown
1159	*E. coli*	H30Rx	*bla*$_{CTX-M-15}$	IncF	three copies of AMR gene, unknown
1162	*E. coli*	ST38	*bla*$_{CTX-M-14b}$	IncF	two copies of AMR gene, chromosomal

*H41, H30 and H30Rx are different FimH types inside ST131 clonal group.

**Some isolates carry more than one copy of the same AMR gene. "Unknown" means we could not identify if AMR was located on chromosome or plasmid. "Chromosomal" and "Plasmid" mean that although we can locate AMR gene on the chromosome or the plasmid, we cannot identify the complete transposon unit or the plasmid sequence cannot be completed by short read sequencing.

2. Extraction of High Molecular Weight DNA from bacterial samples. For this step of the study we needed to do few rounds of optimization and troubleshooting to get high DNA quality and yield. Although we have used the recommended commercial kit (Quick-DNA HMW MagBead kit, Zymo Research), each type of sample usually needs adjustments to the original protocol. For bacterial isolates, period of incubation with proteinase K, number of beads used and elution volumes had to be modified in order to get high molecular weight DNA (DIN > 8) and high yield (> 1 µg).

3. Long read sequencing PacBio Sequel II HiFi. Currently ongoing.

Future directions. [1] Finish PacBio sequencing and analysis. [2] Submit publication (estimated time: first half 2023).

References

[1] Wick RR, Schultz MB, Zobel J, Holt KE. Bioinformatics 2015, 3350-3352.

[2] Bell J, Fajardo-Lubian A, Partridge S, et al. Commun Dis Intell. Australian Group on Antimicrobial Resistance (AGAR) Australian Gram-negative Sepsis Outcome Programme (GnSOP) Annual Report 2020. (2022) 46.

Additional information

1. Project completion.

All bacterial isolates selected for this study have been processed and DNA extraction is completed. Long read PacBio sequencing is currently ongoing.

There have been two main issues that we had to face during the development of this study: **[1]** COVID closed the NSW research laboratories for extended periods and shut down work completely for few months in the second part of 2021. The Pathogen Genomics Service at Westmead Hospital was not able to begin the sequencing of the isolates until the end of January, due to COVID genome sequencing demands, which delayed the start of the analysis of our data. **[2]** Once we had all the short read data analysed and samples selected, DNA extraction for long read sequencing has proven challenging. From the start we sought help from Australian Genome Research Facility (AGRF) and Westmead Research Hub but even with their assistance and expertise high quality DNA sample preparation needed several steps for optimization.

2. Important findings from our research.

The **primary aims** of this study were

To examine plasmid vehicles and strains carrying relevant AMR genes in isolates from the Australian Group on Antibiotic Resistance GNSOP 2020 and generate a trusted "library" of accurate references for the most important AMR plasmids.

progress - research laboratories were closed for most of the second part of 2021 due to the COVID pandemic situation in NSW, which impacted the work for this project. Also, the need for important optimization of DNA extraction delayed the start of long read sequencing.

outcome - This aim has not been fully completed but it is currently in progress. Estimated end of analysis: 3-6 months.

Secondary aim was

(I) to develop rapid testing for key plasmids or plasmid/AMR gene combinations without the need for WGS.

outcome – This aim has also been delayed by COVID genome sequencing demands on equipment and staff at the Pathogen Genomics Service (Westmead Hospital) and by manual optimization of DNA extraction for long read sequencing, but it is now proceeding.

As a result of this work,

- *We have stablished a systematic workflow for analysis of WGS data to identify location of AMR genes that can be used for other bacterial species.*
- *We are generating a database of plasmids to use as references for efficient tracking of most prevalent AMR genes.*
- *We will design rapid tests (PCR based) to quickly reveal outbreak clusters of dangerous AMR transmission in hospitals without having to wait for WGS.*

3. Published papers, reports, media or promotional material relating to the project.

Paper in preparation: long read sequencing ongoing (estimated time for submission: first half 2023).

4. Project expenditure

Activity Item	Notes/Basis of estimate	Budget total	Actual expenditure
Sequencing costs (n=48 samples)	[1] Microbial multiplexing, low-plexing & hi-plexing, SMRTbell Template Prep- Multiplexed Microbial, 16-48 plex [2] PacBio Sequel II Sequencing PacBio Sequel II SMRT Cell 8M	$38,500	$38,500
DNA extraction	[1] media, [2] DNA extraction kits, [3] WIMR Research Hub support	$12,500	$12,500
	TOTAL	**$50,000**	**$50,000**

<u>Support for all other consumables provided from Iredell research grant.</u>