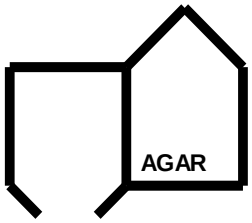

The Australian Group on Antimicrobial Resistance

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 1: Understanding how important AMR genes are becoming fixed in the chromosome.

A. FAJARDO LUBIAN[1,2], S. PARTRIDGE[1,2,3], 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.

Antimicrobial resistance (AMR) genes in Gram-negative organisms commonly disseminate on plasmids, but critical AMR genes are increasingly reported on bacterial chromosomes [1,2,3,4]. If this leads to a permanent change in multiple species, restricting antibiotic use (the primary policy setting globally) will be powerless to reverse the threat of AMR.

In this study, *Escherichia coli* from bacteriaemia episodes (Gram-negative Sepsis Outcome Program, Australian Group on Antimicrobial Resistance, AGAR GNSOP, 2020) carrying extended-spectrum β-lactamase (ESBL) genes were assessed. Our data demonstrate that the most common AMR genes ($bla_{\text{CTX-M}}$ genes) conferring resistance to third generation cephalosporins (ceftazidime and ceftriaxone) are commonly located on the chromosome in major globally disseminated lineages.

Aim. To define the extent of chromosomal AMR gene carriage in Gram-negative bacteria in Australia.

Materials and Methods.

1. Isolates used in this study. 445 *E. coli* isolates with relevant phenotype (resistance to third generation cephalosporins) were selected, based on our PCR screening analysis of 2020 GNSOP collection (**Table 1**), for whole genome sequencing (WGS) to identify the extent to which antimicrobial resistant genes are found in the chromosomes of Gram-negative bacteria under surveillance.

Table 1. *E. coli* isolates selected for WGS.

	Genes identified by PCR* (n=445)			
Organism	**CTX-M-group 1**	**CTX-M-group 9**	**CMY-like**	**DHA**
E. coli	195	221	27	29

*Some isolates have more than one β-lactam gene identified by PCR.

2. *Whole genome sequencing (WGS)*. Total DNA was extracted using DNeasy blood and tissue kit (Qiagen) following manufacturer's instructions. WGS was performed by the Microbial Genomics Reference Laboratory (CIDMLS, ICPMR, Westmead Hospital) using the Illumina NextSeq™ 500 platform. Data were analysed using a modification of the Nullarbor bioinformatic pipeline [5], followed by a custom AMR-specific pipeline. Also, when needed, long-read sequencing (PacBio Sequel II HiFi) at AGRF Brisbane was performed (see details in GNSOP Project 2).

3. Manual analysis of chromosomal integration of AMR genes. 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 [6]. 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. Prevalence of ESBL genes in major E. coli STs in Australia. For this study we performed short-read sequencing in 445 *E. coli* isolates from GNSOP 2020. As part of routine GNSOP programme we already had sequencing data corresponding to 201 *E. coli* isolates from 2019 and 260 *E. coli* isolates from 2020 [7,8]. **Table 2** summarizes the list of most common ESBL genes found in the main *E. coli* STs in this study. Out of 445 *E. coli* isolates sequenced for this project, analysis is completed in 380 isolates (**Table 2**). Unsurprisingly, *E. coli* ST131, a globally disseminated pathogen, is the dominant clonal group also

in Australia (at least 51% of *E. coli* isolates). Data from routine WGS in 2019 and 2020 also showed the same trend (41% and 48% of *E. coli* isolates belonged to ST131, respectively).

Table 2. ESBL genes carried by major *E. coli* STs in Australia.

***E. coli* (major, n>10 isolates, Sequence Type, ST)**	**Genes identified by WGS* (n=380), major types**				
	***bla*$_{CTX-M-15}$**	***bla*$_{CTX-M-14}$**	***bla*$_{CTX-M-27}$**	***bla*$_{CMY-2}$**	***bla*$_{DHA-1}$**
ST38 (n=17)	2	8	4	2	
ST69 (n=20)	10	3	1		5
ST131 (n=197)	75	7	105		
ST1193 (n=29)	7	1	16	2	

*Some isolates have more than one β-lactam gene identified by WGS.

2. Prevalence of chromosomal bla$_{CTX-M}$ in E. coli in Australia. Our ongoing analysis of routine *E. coli* short-read WGS data for 2019 (n=201), 2020 isolates (n=260) and current project (n=445) reveal that the most prevalent AMR genes (*bla*$_{CTX-M}$ genes, **Table 2**) conferring resistance to widely used extended-spectrum β-lactam antibiotics (ceftazidime and ceftriaxone) are mostly found on the chromosome in major pandemic lineages of *E. coli* (**Table 3**). The frequency of chromosomal location of these genes is clearly dependent of ST and type of gene analysed.

Table 3. Prevalence of chromosomal ESBL genes in most common *E. coli* sequence types (STs) in Australia.

E. coli	**ESBL gene**	**Chromosomal location (%)**
ST12	*bla*$_{CTX-M-14}$ / *bla*$_{CTX-M-15}$	100%
ST38	*bla*$_{CTX-M-14}$	100%
ST131-FimH30	*bla*$_{CTX-M-14}$	87%
ST131-FimH30x	*bla*$_{CTX-M-15}$	*At least 23%
ST131-FimH41	*bla*$_{CTX-M-27}$ / *bla*$_{CTX-M-15}$	*At least 80%
ST648	*bla*$_{CTX-M-15}$	100%
ST1193	*bla*$_{CTX-M-15}$	100%

*For some isolates, location of ESBL genes could not be determined. Summary of ongoing analysis of isolates from routine WGS data from 2019 (n=201), 2020 (n=260) and current project 2020 (n=445).

3. Chromosomal location of bla$_{CTX-M}$ genes. By manual checking of sequencing data (section 3, Materials and Methods) we have been able to identify most of the chromosomal contexts where *bla*$_{CTX-M}$ has integrated. As shown in **Table 4**, the chromosomal locations vary depending on the specific bacterial lineage-ESBL gene pair examined. Most of the isolates carry one copy of the AMR gene but some isolates carry more than one *bla*$_{CTX-M}$ variant and/or more than one copy of a specific *bla*$_{CTX-M}$ gene in their chromosome and/or plasmids. Although few chromosomal contexts have been already identified as spots for *bla*$_{CTX-M}$ integration [1], during this study we have define mostly new chromosomal insertions. The fact that for some STs these chromosomal locations and transposon units are conserved and predominant (such as *mgtC/sapB* and *dacD* in ST38, SHPF domain protein and *gspD* in ST131) will allow us to develop in silico PCRs to automate the analysis of short-read data. When multiple copies of *bla*$_{CTX-M}$ genes were found or the exact chromosomal location of the *bla*$_{CTX-M}$ gene was not identifiable, isolates where selected for PacBio long-read sequencing (see **Table 1** from project GNSOP 2, "Understanding AMR plasmids in Gram-negative bacteria in Australia").

Table 4. Chromosomal location of ESBL genes.

Bacterial gene interrupted	***E. coli***	**ESBL gene***
bglA (glucosidase), *dtd* (tRNA deacylase), fimbrial type 1 protein	ST12	$bla_{\text{CTX-M-15, -14}}$
rbsR (transcriptional regulator), *yicI* (hydrolase), *glgX* (hydrolase), *mgtC/sapB* (magnesium transporter), *spoT* (synthase/hydrolase), *btuB* (outer membrane transporter), *dacD* (carboxypeptidase).	ST38	$bla_{\text{CTX-M-14}}$
SHPF domain protein and unknown locations	ST131-FimH30	$bla_{\text{CTX-M-14}}$
Nucleoside transporter, *LacI (*transcriptional regulator)	ST131-FimH41	$bla_{\text{CTX-M-15}}$
gspD (Type II secretion system protein), prophage region	ST131-FimH41	$bla_{\text{CTX-M-27}}$
mdtN (efflux pump), *recQ* (recombinase), *nanM* (carbon metabolism)	ST648	$bla_{\text{CTX-M-15}}$

*Some isolates carry multiple copies of $bla_{\text{CTX-M}}$. Summary of ongoing analysis of isolates from routine WGS data from 2019 (n=201), 2020 (n=260) and current project 2020 (n=445).

Conclusion.

Our results indicate that some global epidemic bacterial strains might be permanently incorporating AMR genes into their core genome. The chromosomal fixation of AMR genes will have enormous implications for antibiotic use policies as the view that less antibiotic use will allow AMR to vanish might not be valid in the near future.

Future directions. [1] Finish PacBio sequencing and analysis. [2] Submit publication (estimated time: submission by end 2022).

References

[1] Yoon EJ, Choi YJ, Kim D, Won D et al., mSpectrum (2022) Apr. 25
[2] Yoon EJ, Gwon B, Liu C, Kim D, Wong D, et al. mSystems (2020) 29;5(5).
[3] Mathers AJ, Stoesser N, Chai W, Carrol J, Barry K, et al. Antimicrob Agents Chemother (2017) 23;61(3):e01823-16.
[4] Turton JF, Doumith M, Hopkins KL, Perry C, Meunier D, Woodford N. J Med Microbiol (2016) 65(6):538-546.
[5] Lee RS, Seemann T, Heffernan H, Kwong JC, Goncalves da Silva A, et al. J Antimicrob Chemother. (2018) 73: 353–364.
[6] Wick RR, Schultz MB, Zobel J, Holt KE. Bioinformatics 2015, 3350-3352.

[7] 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 2019. (2020) 44.

[8] 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 originally selected for this study have been processed and WGS is completed. Currently we have finished analysis for most of the isolates (380 out of 445 isolates).

There has been one main issue that we had to face during the development of this study: 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 Genomic 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, what delayed the start of the analysis of our data.

The approach developed during this study will be implemented in the analysis of isolates for subsequent GNSOP programs. We are also planning to revise periodically publicly available genomes and include local isolates referred to the state reference laboratory (Pathogen Genomic Service, Westmead Hospital) in our analysis as an ongoing process. AMR and plasmid epidemiology are permanently changing so continuous monitoring of these processes is needed.

2. Important findings from our research.

The **primary aim** of this study was

To determine the extent to which antimicrobial resistant genes are located on the chromosomes of bacteria under surveillance.

progress - research laboratories were closed for major second part of 2021 due to COVID pandemic situation in NSW what has impacted the work for this project.

outcome - This aim was nevertheless ***achieved*** as we have been able to analyse over 80% of all isolates sequenced. The results obtained from this study will help to streamline this labour-intensive analysis for successive surveillance projects.

Secondary aims were

(I) to automate the manual checking of gene location by developing in silico PCRs.

(II) to perform long read sequencing of additional key, representative subsets of isolates (see **Table 1** from project GNSOP 2, "Understanding AMR plasmids in Gram-negative bacteria in Australia").

outcome – These two aims have been delayed by COVID genome sequencing demands on equipment and staff at The Pathogen Genomic Service (Westmead Hospital), but they are now proceeding.

As a result of this work,

- *We have determined major E. coli sequence types for most prevalent ESBL genes found in Australia.*
- *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 designing tools to automate the identification of chromosomal integration of most prevalent ESBL genes from short-read sequencing data.*
- *We are generating a database of transposon units to track the chromosomal integration of $bla_{CTX\text{-}M}$ genes. For this, we are using local isolates and publicly available genomes.*

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

Paper in preparation: final analysis ongoing (estimated time for submission: end 2022).

Presentation in Conference: submitted for oral presentation at ASM (Sydney, Australia, July 2022) and confirmed oral presentation at IC2AR (Caparica, Portugal, September 2022).

4. Variation to initial proposal.

E. coli isolates only, instead of *E. coli* and *K. pneumoniae* isolates, were sequenced for this project. *E. coli* is by far the most predominant pathogen isolated from bacteriemia episodes in Australia (over 55% of the total). However, *K. pneumoniae* isolates (n ~100) with relevant phenotypes (resistance to third generation cephalosporins) are also being analysed from routine WGS data obtained from 2019 and 2020 to identify if similar rates of chromosomal integration of ESBL genes occur in this bacterial specie.

5. Project expenditure

Activity Item	Notes/Basis of estimate	Budget total	Actual expenditure
Sequencing costs (n=445 samples)	1) gDNA extraction ($1,700 per 250 samples) 2) library preparation and sequencing ($100 per sample) 3) bioinformatic software (Geneious license, $1,100 per year)	$50,000	$50,000
	TOTAL	**$50,000**	**$50,000**

Support for all other consumables provided from Iredell research grant.