Genetic Determinants Underlying the Progressive Phenotype of β-lactam/β-lactamase Inhibitor Resistance in *Escherichia coli*

Running Title: WGS assessment of progressive BL/BLI resistance

William C Shropshire[1*], Hatim Amiji[2*], Jordan Bremer[1], Selvalakshmi Selvaraj Anand[3], Benjamin Strope[3], Pranoti Sahasrabhojane[1], Marc Gohel[1], Samuel Aitken[4], Sarah Spitznogle[4], Xiaowei Zhan[5], Jiwoong Kim[5], David E Greenberg[6,7], Samuel A Shelburne[1,8#]

[1]Department of Infectious Diseases, Infection Control, and Employee Health, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

[2]Frank H. Netter MD School of Medicine, Quinnipiac University, Hamden, CT, USA

[3]Program in Diagnostic Genetics and Genomics, MD Anderson Cancer Center School of Health Professions, Houston, TX, USA

[4]Division of Pharmacy, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

[5]Quantitative Biomedical Research Center, Peter O'Donnell Jr. School of Public Health, University of Texas Southwestern Medical Center, Dallas, TX, 75390

[6]Department of Internal Medicine, UT Southwestern, Dallas, TX 75390, USA

[7]Department of Microbiology, UT Southwestern, Dallas, TX 75390, USA

[8]Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

*These authors contributed equally to the manuscript

#Address for correspondence: Samuel A Shelburne, MD, PhD, 1901 East Rd., South Campus Research Building, 4SCR5.1042, Houston, TX 77054; phone: (713) 792-3629, e-mail: sshelburne@mdanderson.org

Abstract Word Count: 250

Importance Word Count: 148

Main Text: 4911

Abstract

Currently, whole genome sequencing (WGS) data has not shown strong concordance with *E. coli* susceptibility profiles to the commonly used β-lactam/β-lactamase inhibitor (BL/BLI) combinations: ampicillin-sulbactam (SAM), amoxicillin-clavulanate (AMC), and piperacillin-tazobactam (TZP). Progressive resistance to these BL/BLIs in absence of cephalosporin resistance, also known as extended-spectrum resistance to BL/BLI (ESRI), has been suggested to primarily result from increased copy numbers of bla_{TEM} variants, which is not routinely assessed in WGS data. We sought to determine whether addition of gene amplification could improve genotype-phenotype associations through WGS analysis of 147 *E. coli* bacteremia isolates with increasing categories of BL/BLI non-susceptibility ranging from ampicillin-susceptible to fully resistant to all three BL/BLIs. Consistent with a key role of bla_{TEM} in ESRI, 112/134 strains (84%) with at least ampicillin non-susceptibility encoded bla_{TEM}. Evidence of bla_{TEM} amplification (i.e., bla_{TEM} gene copy number estimates > 2×) was present in 40/112 (36%) strains. There were positive correlations between bla_{TEM} copy numbers with minimum inhibitory concentrations (MICs) of AMC and TZP (P-value < 0.05), but not for SAM (P-value = 0.09). The diversity of β-lactam resistance mechanisms, including non-ceftriaxone hydrolyzing bla_{CTX-M} variants, bla_{OXA-1}, as well as *ampC* and bla_{TEM} strong promoter mutations, were greater in AMC and TZP non-susceptible strains but rarely observed within SAM and AMP non-susceptible isolates. Our study indicates a comprehensive analysis of WGS data, including β-lactamase encoding gene amplification, can help categorize *E. coli* with AMC or TZP non-susceptibility but that discerning the transition from SAM susceptible to non-susceptible using genetic data requires further refinement.

Importance

The increased feasibility of whole genome sequencing has generated significant interest in using such molecular diagnostic approaches to characterize difficult-to-treat, antimicrobial resistant (AMR) infections. Nevertheless, there are current limitations in the accurate prediction of AMR phenotypes based on existing AMR gene database approaches, which primarily correlate a phenotype with the presence/absence of a single AMR gene. Our study utilized a large cohort of cephalosporin-susceptible *E. coli* bacteremia samples to determine how increasing dosage of narrow-spectrum β-lactamase encoding genes in conjunction with other diverse BL/BLI genetic determinants contribute to progressively more severe BL/BLI phenotypes. We were able to characterize the complexity of the genetic mechanisms underlying progressive BL/BLI resistance including the critical role of β-lactamase encoding gene amplification. For the diverse array of AMR phenotypes with complex mechanisms involving multiple genomic factors, our study provides an example of how composite risk scores may improve understanding of AMR genotype/phenotype correlations.

Introduction

Rapid and accurate characterization of antimicrobial resistant (AMR) infections is necessary to combat their increasing threat to public health (NEED REFS: 28528569, 36636534 – these focus on getting effective antibiotics into patients quickly). There has been recent interest both from a clinical and research standpoint in using whole genome sequencing (WGS) to predict antimicrobial susceptibility patterns as molecular diagnostic approaches become more feasible (1, 2). When AMR is driven by a single gene, such as $bla_{\text{CTX-M}}$ for an extended-spectrum β-lactamase phenotype, AMR database query approaches generally detect a strong concordance between WGS predicted and observed phenotypes (3). However, discordant genotype-phenotype predictions can occur due to complicated resistance mechanisms that involve multiple contributing genetic factors (4). Such issues have been consistently observed when trying to use WGS to predict the antimicrobial susceptibility pattern of β-lactamase/β-lactamase inhibitor (BL/BLI) combinations (3, 5, 6).

Although recent years have seen the introduction into clinical practice of such novel BL/BLI combinations as ceftolozane/tazobactam and ceftazidime/avibactam, the vast majority of BL/BLIs currently used are ampicillin-sulbactam (SAM), amoxicillin-clavulanate (AMC), and piperacillin-tazobactam (TZP) (7-9). Given the clinical impact of *Escherichia coli* and its highly varied susceptibility profile to these three BL/BLI combinations, *E. coli* is among the most well studied organisms in terms of BL/BLI resistance (6, 10-19). Cephalosporinases and carbapenemases that hydrolyze broad spectrum β-lactams generally also inactivate SAM, AMC, and TZP (20). There has been increasing interest in characterizing cephalosporin-susceptible *E. coli* strains with varying susceptibility patterns amongst these three common BL/BLI combinations (11-14). Nevertheless, the characterization of these increasingly more severe BL/BLI resistant phenotypes using WGS to detect potential associations in well-defined cohorts are less understood.

To date, most studied *E. coli* ceftriaxone-susceptible, BL/BLI resistant strains harbor class A or class D beta-lactamases such as bla_{TEM-1B} or bla_{OXA-1} (11-13). It has been suggested that the BL/BLI resistance amongst such strains evolves in a gradient fashion in order from ampicillin-sulbactam non-susceptible (SAM-NS) to amoxicillin-clavulanate non-susceptible (AMC-NS) to piperacillin-tazobactam non-susceptible (TZP-NS) (11, 21). Based on studies in a limited number of strains under laboratory passage as well as in clinical isolates, bla_{TEM} amplification has been proposed as a major mechanism of what has been called extended-spectrum resistance to BL/BLIs or ESRI (11-13, 16, 19). However, there are many other genetic determinants that can contribute to reduced susceptibility to BL/BLIs such as augmented expression of bla_{TEM} due to promoter mutations (22, 23), de-repression of the chromosomal *ampC* gene (24, 25), decreased outer membrane permeability (18, 26), inhibitor resistant TEM variants (27), as well as CTX-M enzymes with decreased cephalosporin affinity but increased TZP hydrolysis activity (28). Furthermore, it is not clear how these BL/BLI genetic determinants contribute to BL/BLI resistance across the full ESRI spectrum.

There has been in-depth examination of *E. coli* BL/BLI resistance mechanisms with regards to each respective BL/BLI combination. For example, Noguchi et al. used a phenotypic approach and found that hyperproduction of TEM-1 and altered cell membrane permeability accounted for a large proportion of SAM resistance (18). Conversely, WGS based analyses of AMC and TZP resistance showed a complex variety of mechanisms including bla_{TEM-1} overexpression as well as increased copy numbers of various *bla* encoding enzymes (6, 14). A critical missing piece in understanding the role of genetic analyses in BL/BLI resistance is the lack of WGS investigation of a large number of cephalosporin-susceptible *E. coli* isolates with varying BL/BLI resistance profiles. Thus, we performed WGS of 147 ceftriaxone-susceptible *E. coli* which had BL/BLI resistance phenotypes ranging from ampicillin (AMP) susceptible to TZP non-susceptible. We

sought to determine whether genetic analysis of known BL/BLI resistance mechanisms could be used to classify amongst the various phenotypic ESRI categories. Our data suggest that progressive ESRI is driven in part by *bla*$_{TEM}$ amplification but that many complex mechanisms detectable by WGS analysis also contribute to AMC and TZP non-susceptibility. The lack of ability to separate SAM susceptible from SAM non-susceptible using WGS data, including analysis of *bla*$_{TEM}$ copy number, indicates that improvements of these genotype/phenotype correlations may require alternative approaches to current AMR genetic database query methods.

Results

Distribution of ESRI groups amongst CRO-S *Escherichia coli* BSI isolates

We identified 1,026 *E. coli* bloodstream isolates from May 1st, 2015, to April 30th, 2020. Of these, 389 (38%) had a ceftriaxone (CRO) MIC ≥ 4 mg/L. Consistent with a high rate of BL/BLI resistance among CRO resistant isolates, 91% and 76% were non-susceptible to SAM and AMC respectively. We selected the remaining 637 (62%) *E. coli* isolates in our sampling frame with a CRO MIC < 4 mg/L to study cephalosporin susceptible, BL/BLI associated resistance mechanisms (Fig. S1A). We grouped isolates by their susceptibility patterns according to CLSI guidelines as follows: β-lactam pan-susceptible (PAN-S; Group 1), ampicillin non-susceptible (AMP-NS; Group 2), ampicillin-sulbactam non-susceptible (SAM-NS; Group 3), amoxicillin-clavulanate non-susceptible (AMC-NS; Group 4), and piperacillin-tazobactam non-susceptible (TZP-NS; Group 5) as shown in **Fig. 1**.

The plural majority (230/637; 36%) of CRO susceptible *E. coli* BSI isolates were pan-susceptible to the studied β-lactams (Group 1). The next most common isolates were SAM-NS (Group 3; 27%) and AMC-NS (Group 4; 20%) whereas the lowest frequencies were found within AMP-NS (Group 2; 9%) and TZP-NS (Group 5; 8%). We next looked at the minimum inhibitory concentration (MIC) distributions for each respective BL/BLI across each of the BL/BLI groups (Fig. S1B). For AMP and TZP antimicrobial susceptibility testing (AST), there was a bifurcation of susceptible and resistant isolates with very few isolates near the respective MIC breakpoints (Fig. S1B). Conversely, for SAM and AMC AST, there were large numbers of isolates which had intermediate resistant phenotypes, which likely reflects a more diffuse spectrum of resistance for SAM and AMC vs. AMP and TZP (Fig. S1B).

From the CRO susceptible *E. coli* BSI cohort, we selected 147 isolates for whole genome sequencing across the spectrum of BL/BLI susceptibilities with the group distribution of

sequenced strains relative to the total cohort shown in **Fig. 1**. We under- and over-sampled Group 1 and Group 5 respectively whereas Groups 2, 3, and 4 had similar proportions of isolates sequenced within each group respective to the total cohort (148/637; 23%) as shown in Figure 1.

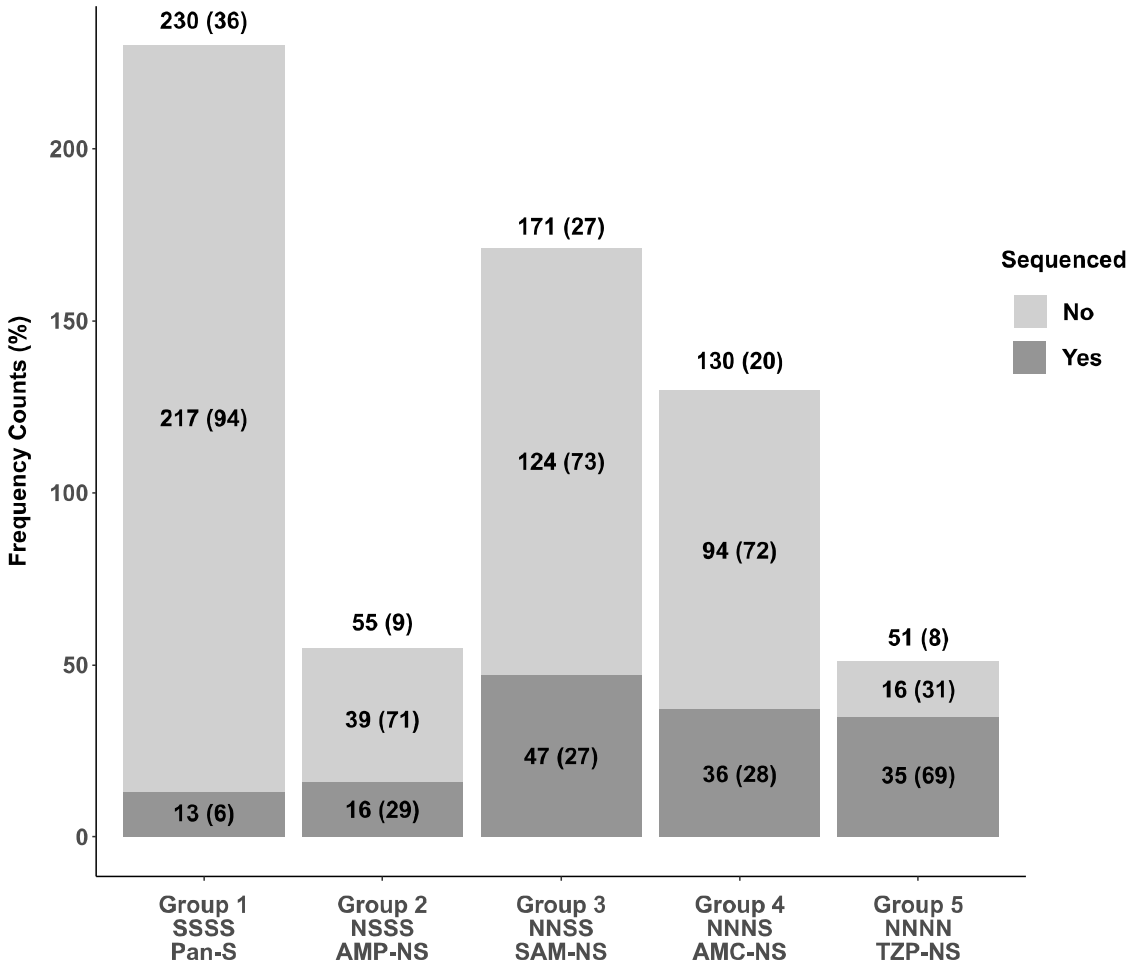

FIG 1 Ceftriaxone susceptible (*i.e.*, CRO MIC < 4 mg/L) *Escherichia coli* bloodstream isolates (n=637) stratified by increasing BL/BLI resistant phenotype (Group 1 – 5). Unique BSI samples are grouped into progressively more resistant BL/BLI phenotypes from pan-susceptible (Group 1) to resistant to all four drug combinations in our study (Group 5). Frequency counts above each group give group totals with percentage of total population in parenthesis. We further split groups into sequenced (dark gray) vs non-sequenced (light gray) with frequency counts and percentages of each respective group labelled. N = Non-susceptible; S = Susceptible; Pan-S = pan-susceptible; AMP-NS = ampicillin non-susceptible; SAM-NS = ampicillin-sulbactam non-susceptible; AMC = amoxicillin-clavulanate non-susceptible; TZP-NS = piperacillin-tazobactam non-susceptible.

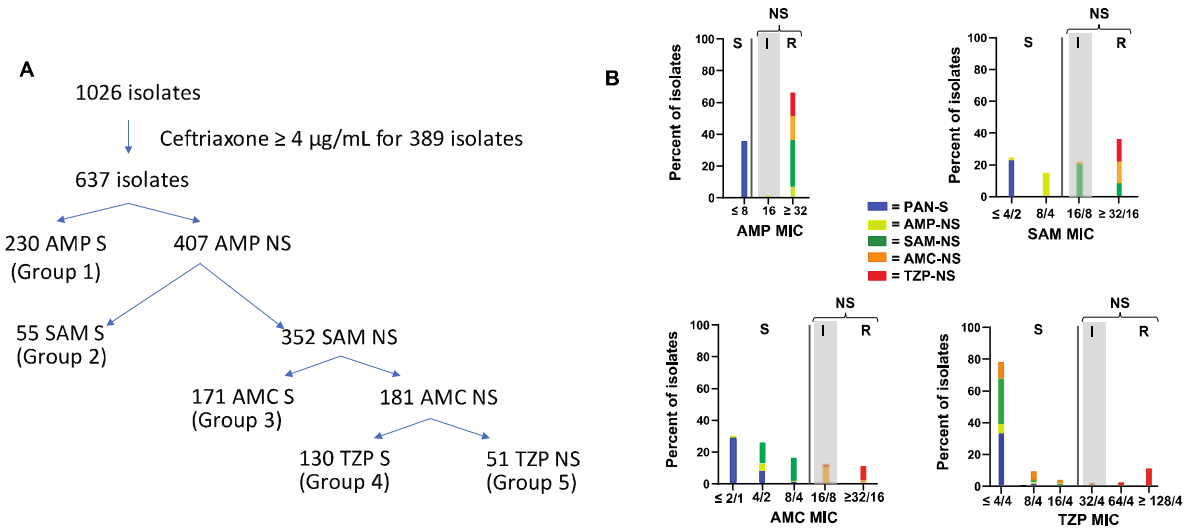

FIG S1. Study Population and BL/BLI MIC Distribution Across Each Non-Susceptible Group. (A) Overview of total *E. coli* bacteremia isolates detected from May 1st, 2015, to April 30th, 2020, with their respective ESRI phenotype patterns. S = susceptible, NS = non-susceptible. (B) Categorical distribution of minimum inhibitory concentration (MIC; µg/mL) values across each respective BL/BLI non-susceptibility group designation for AMP, SAM, AMC, and TZP respectively.

Exogenous β-lactamase gene detection contribution to BL/BLI genetic determinants

The majority of isolates which were at least AMP-NS had ≥ one exogenous β-lactamase encoding gene detected (129/134; 96%). The bla_{TEM} variants were the most commonly identified exogenous β-lactamase encoding genes being present in 84% (112/134) of AMP-NS or greater strains with the non-inhibitor resistant TEM (N-IRT) gene, bla_{TEM-1B}, accounting for the vast majority (n = 110) of N-IRT detected. The bla_{OXA-1} gene was the second most commonly identified exogenous β-lactamase encoding gene being present in 19/134 strains (14%). Previously characterized inhibitor resistant TEM (IRT) variants were quite rare, with bla_{TEM-31} and bla_{TEM-40} each found once. Other rare β-lactamase encoding genes included bla_{SHV-1} (n=1), bla_{CARB-2} (n=2), bla_{HER-3} (n=1), bla_{LAP-2} (n=1) and two bla_{CTX-M} variants further described below.

For both AMP-NS and SAM-NS strains, nearly all strains carried bla_{TEM-1} alone with a single SAM-NS strain containing bla_{CARB-2} whereas bla_{LAP} and bla_{CARB-2} were present in two AMC-NS strains. Conversely, strains in the TZP-NS had the most diverse array of β-lactamase encoding genes including two strains with IRT β-lactamase variants (n = 2) and CTX-M enzymes (n = 2) respectively. One CTX-M enzyme was a derivative of CTX-M-15 (*i.e.,* CTX-M-189) with an S133G polymorphism whereas the other was a derivative of CTX-M-27 (*i.e.,* CTX-M-255) and contained a G239S polymorphism. Both of these polymorphisms have been shown in an experimental model to engender resistance to β-lactamase inhibitors but also to diminish the ability to hydrolyze cephalosporins (28). There was a statistically significant difference in exogenous β-lactamase gene content between the groups (χ^2 P-value < 0.05). Specifically, AMP-NS strains were significantly less likely to contain any exogenous β-lactamase whereas TZP-NS strains were more likely to contain bla_{OXA-1} (Fisher's exact test P-value < 0.001 for each). When only Groups 2 through 4 (*i.e.,* AMP-NS, SAM-NS, AMC-NS) were analyzed, no statistically significant differences in β-lactamase encoding gene content were observed. Thus,

we conclude that the fully susceptible and fully resistant groups can be distinguished from the other groups using exogenous β-lactamase content alone, but the middle three groups cannot.

Contribution of additional mechanisms to BL/BLI non-susceptibility

Given that exogenous β-lactamase encoding genes only separated the fully susceptible and fully resistant strains from the remainder of the cohort, we next focused on other factors previously associated with BL/BLI resistance (6). First, *E. coli* strains encode an AmpC β-lactamase which is typically transcriptionally silenced but can become active in the presence of promoter mutations (24). We identified three strains, all in AMC-NS isolates, which contained *ampC* promoter variations previously associated with AMC resistance (24). Similarly, we assessed for variation in the bla_{TEM} promoter region and found nine instances of strong bla_{TEM} promoter variants (*i.e.,* eight *Pa/Pb* and one *P5*) (22, 23). For eight of the nine cases, strains were in the TZP-NS group with the remaining strain having a borderline TZP MIC of 16 mg/L indicating that bla_{TEM} promoter mutations were associated with TZP non-susceptibility. Finally, we analyzed the OmpC and OmpF content of our cohort inasmuch as variation in outer membrane protein profile has been associated with BL/BLI resistance (18). Only four strains had predicted inactivating *ompC* (n = 1) or *ompF* (n = 3) mutations and the strains were present in varied groups (one in AMP-NS, one in SAM-NS, and two in TZP-NS). Thus, 16 strains (11%) had genetic changes predicted to increase *ampC* or bla_{TEM} expression or eliminate *ompC/ompF* production which have been previously associated with BL/BLI resistance.

Association of bla_{TEM} amplification and BL/BLI resistance

Increased TEM-1 activity has previously been shown to be an integral aspect of progressive BL/BLI resistance (14, 16) with mechanisms involving increased transcription due to promoter variation and/or increase in bla_{TEM-1} copy number (16, 22, 23). The median copy number

estimate of *bla*$_{TEM}$ was 1.57× with a maximum of 45× observed **(Fig. 2A)**. Taking a cut-off of 2.0-fold DNA coverage as evidence of amplification (29), 40/115 *bla*$_{TEM}$ NIR containing strains (35%) had *bla*$_{TEM}$ amplification. To understand the relationship more clearly between *bla*$_{TEM}$ amplifications and BL/BLI resistance we focused on the 91 strains which only contained N-IRT genes as a β-lactam resistance mechanism (*i.e.,* no other exogenous β-lactamase, no *ampC*/*bla*$_{TEM}$ promoter variation, and no *ompC/ompF* disruptions). We found statistically significant increases in *bla*$_{TEM}$ amplification levels between SAM-NS and TZP-NS isolates **(Fig. 2B)**. However, even for groups where there was a statistically significant difference, there remained overlap such that *bla*$_{TEM-1}$ amplifications did not clearly distinguish between the various resistance profiles **(Fig. 2B)**. Consistent with the inability of *bla*$_{TEM-1}$ amplifications to distinguish between strains from AMP-NS and SAM-NS, there was no significant correlation between SAM MIC and *bla*$_{TEM-1}$ amplification (*P*-value = 0.09, **Fig. 2C**). However, AMC and TZP MICs were significantly associated with *bla*$_{TEM-1}$ amplification (*P*-value < 0.01, **Fig. 2D, 2E**). When *bla*$_{TEM}$ amplification > 2× was considered as a potential β-lactam mechanism, there was a statistically significant difference in β-lactamase content between SAM-NS and AMC-NS (Fisher's exact test *P*-value < 0.001) whereas AMP-NS and SAM-NS continued to have similar mechanisms (Fisher's exact test *P*-value = 0.17). Thus, these data suggest that *bla*$_{TEM-1}$ copy numbers help distinguish AMC resistant from AMC susceptible strains but do not assist with assessing SAM susceptibility.

As previously noted, *bla*$_{OXA-1}$ containing strains were only present in AMC-NS and TZP-NS isolates with the vast majority of *bla*$_{OXA-1}$ containing strains belonging to TZP-NS (16/19). The three *bla*$_{OXA-1}$ AMC-NS strains did not contain another BL/BLI resistance mechanism whereas six TZP-NS strains contained both *bla*$_{OXA-1}$ and *bla*$_{TEM-1}$, and one TZP-NS strain had a TEM-promoter variation, *bla*$_{OXA-1}$, and *bla*$_{TEM-1}$. Similar to *bla*$_{TEM-1}$, an increase in *bla*$_{OXA-1}$ CNV

estimates was commonly observed in our cohort with 11/19 strains (57%) having > 2.0× normalized coverage depths. When strains containing only *bla*$_{OXA-1}$ as a β-lactam resistant mechanism were considered, there was an increase in CNV positively correlated with TZP MIC albeit this did not reach statistical significance **(Fig. 2F)**. Consistent with these data, all strains with a *bla*$_{OXA-1}$ copy number ≥ 2× were in the fully resistant category.

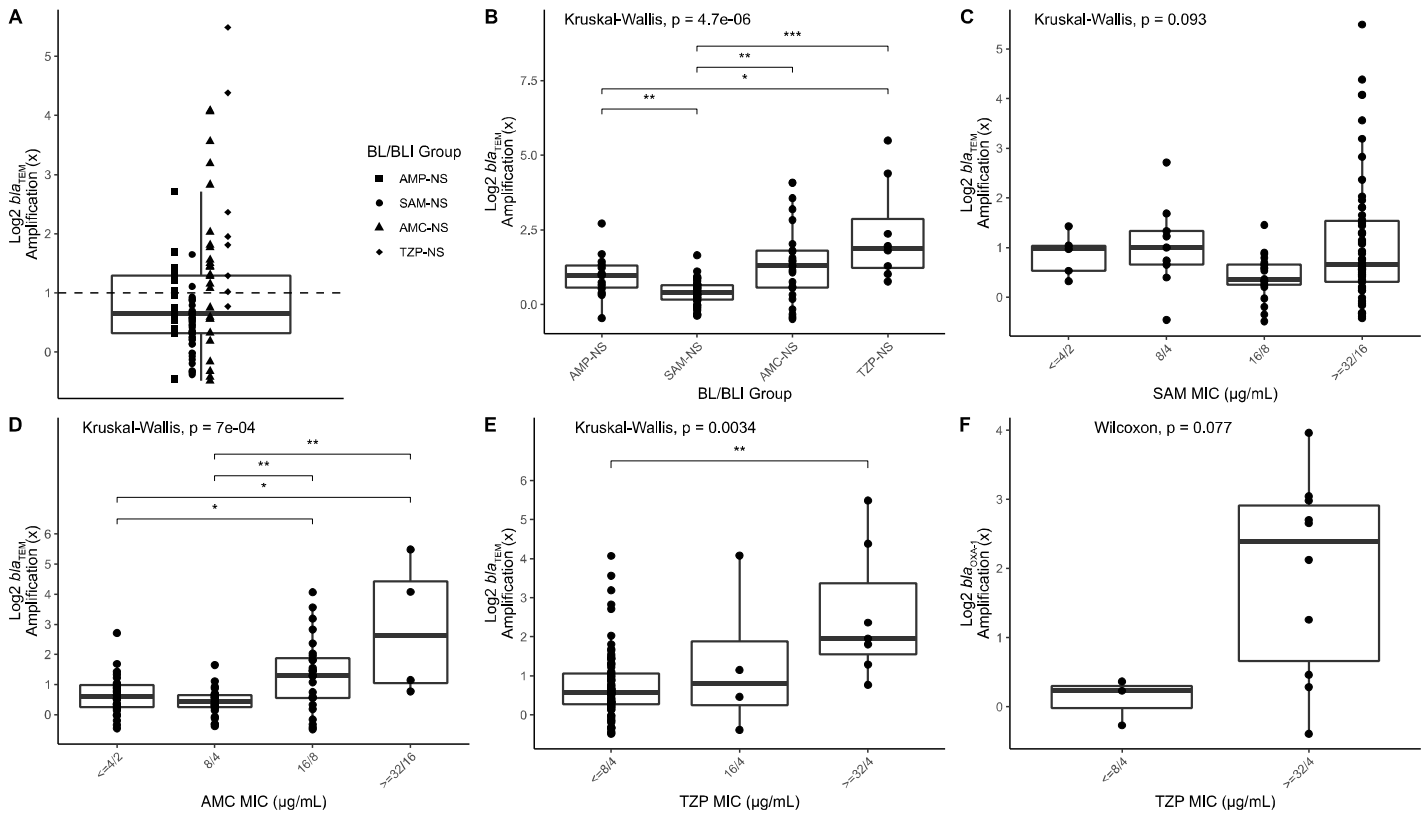

FIG 2. Log2 transformed *bla* copy number variant (CNV) estimates for BL/BLI isolates with single exogenous β-lactamase gene excluding non-exogenous BL/BLI genetic determinant factors. (A) CNV estimates for bla_{TEM-1} for 91 *E. coli* stratified by BL/BLI group indicated by shape. Horizontal dashed line indicates cut-off for 2x CNV; (B) bla_{TEM-1} CNV across each of the BL/BLI groups; (C) bla_{TEM-1} CNV by SAM MIC; (D) bla_{TEM-1} CNV by AMC MIC; (E) bla_{TEM-1} CNV by TZP MIC; (F) bla_{OXA-1} CNV by TZP MIC. Global *P*-values for Kruskal-Wallis or Wilcoxon Rank-sum Tests reported above each respective panel. For global *P*-value < 0.05, pairwise Wilcoxon Rank-sum test adjusted *P*-values are reported as follows: * < 0.05; ** <0.01; *** <0.001; **** <0.0001.

Correlations of BL/BLI genetic determinants with BL/BLI groups across population structure

A core gene alignment inferred maximum-likelihood phylogeny is presented in **Fig. 3A**. There were eight core gene inferred clusters identified through the hierBAPS algorithm (30) that are highlighted in the phylogeny **(Fig. 3A)**. Each of the cluster levels were associated with previously established phylogroups (31) or more closely related sequence types (STs) with highly related STs labelled on branch tips grouping together **(Fig. 3A)**. The majority of isolates in this CRO-S cohort belonged to the B2 clade (60%; 88/147) although a total of 37 distinct STs were present. Strains of the pandemic ST131 clade (33%; 49/147) or the emergent ST1193 (16%; 23/147) were the most frequently observed STs with both belonging to the B2 phylogroup. There was only one other ST with more than 10 observations, which was ST69 (7%; 11/147). Other STs comprised of five or more strains included ST73 (n = 8), ST648 (n = 6) and ST10 (n = 5). We grouped STs with less than 5 strains together as “Rare STs” and such strains accounted for 31% of the cohort (n = 45). When comparing all STs comprised of ≥ 5 isolates, there was a statistically significant relationship between ST and AMR grouping (χ^2 simulated *P*-value < 0.001) **(Fig. 3B)**. Specifically, ST648 strains were more likely to be TZP-NS, and ST1193 strains were more likely to be SAM-NS (Pairwise Fisher’s Test adj. *P*-value < 0.05). Conversely, ST131 strains were not statistically significantly associated with any particular susceptibility profile. While not statistically significant, similar trends can be seen on **Fig. 3A** where ST73/ST12 isolates (hierBAPS level 7; light blue) have 85% isolates with AMC-NS or TZP-NS phenotypes whereas ST69 have more susceptible patterns with 91% of isolates having SAM-NS or a more susceptible phenotype.

We further characterized each of the BL/BLI groups by the number of BL/BLI genetic determinants, which could range from 0 to 10 (see Methods for details) **(Fig. 3C)**. Each of these

BL/BLI genetic determinants were based on presence/absence of genetic features that have been previously identified as contributing to BL/BLI resistance (11, 18, 22, 24, 28, 32). The number of unique combinations of BL/BLI genetic determinants across each of the BL/BLI phenotype groups are shown in Fig S2. All PAN-S isolates (n=13) had no identifiable BL/BLI determinants whereas almost all isolates (130/134; 97%) with AMP-NS or higher had at least one genetic determinant identified (Fig S2). Interestingly, all isolates in our cohort with 3+ BL/BLI determinants (n=10) were TZP-NS **(Fig. 3C)**. Importantly, there was a positive monotonic relationship (ρ = 0.62, *P*-value = 2.2e-16) between BL/BLI group status and BL/BLI genetic determinant score thus underscoring how increasing numbers of BL/BLI genetic determinants impact BL/BLI non-susceptibility. Nevertheless, there was similar distribution of BL/BLI genetic determinant scores across Group 2 and Group 3 isolates (Table S1), further highlighting the difficulty in discriminating genetic differences across these isolates.

Finally, we next sought to combine genetic determinants associated with BL/BLI resistance along with our phylogenomic structure to assess the odds of being in a greater or lesser BL/BLI non-susceptibility group using ordinal logistic regression methods. **Table 1** provides an overview of BL/BLI and population level covariate relationships with the odds of being in each BL/BLI phenotypic group. There were no statistically significant associations observed between increasing BL/BLI group and ST or phylogroup; however, there was a weak association between Cluster 2, predominantly consisting of B1/C isolates that had 0.3 times (95% CI: 0.1 – 0.9) lesser odds of belonging to a higher (i.e., less susceptible) BL/BLI group compared to Cluster 1 isolates (ST131). When looking at relationships of increasing BL/BLI non-susceptibility with the composite BL/BLI genetic determinant score, for every one unit increase in BL/BLI genetic determinant, the odds of being in a higher BL/BLI group was 10.9 times (95% CI: 6.2 - 20.2) greater than a lower BL/BLI group **(Table 1)**. Importantly, there were no associations between N-IRT gene carriage and BL/BLI group (OR = 1.1; 95% CI: 0.5 – 2.4) whereas bla_{TEM}

amplification, *bla*$_{OXA-1}$ presence, and other non-TEM *bla* amplifications all appeared to contribute to increasing odds of BL/BLI non-susceptibility. Thus, these data suggest that the presence and amplification of narrow-spectrum β-lactamases are the primary drivers of the progressive BL/BLI phenotype.

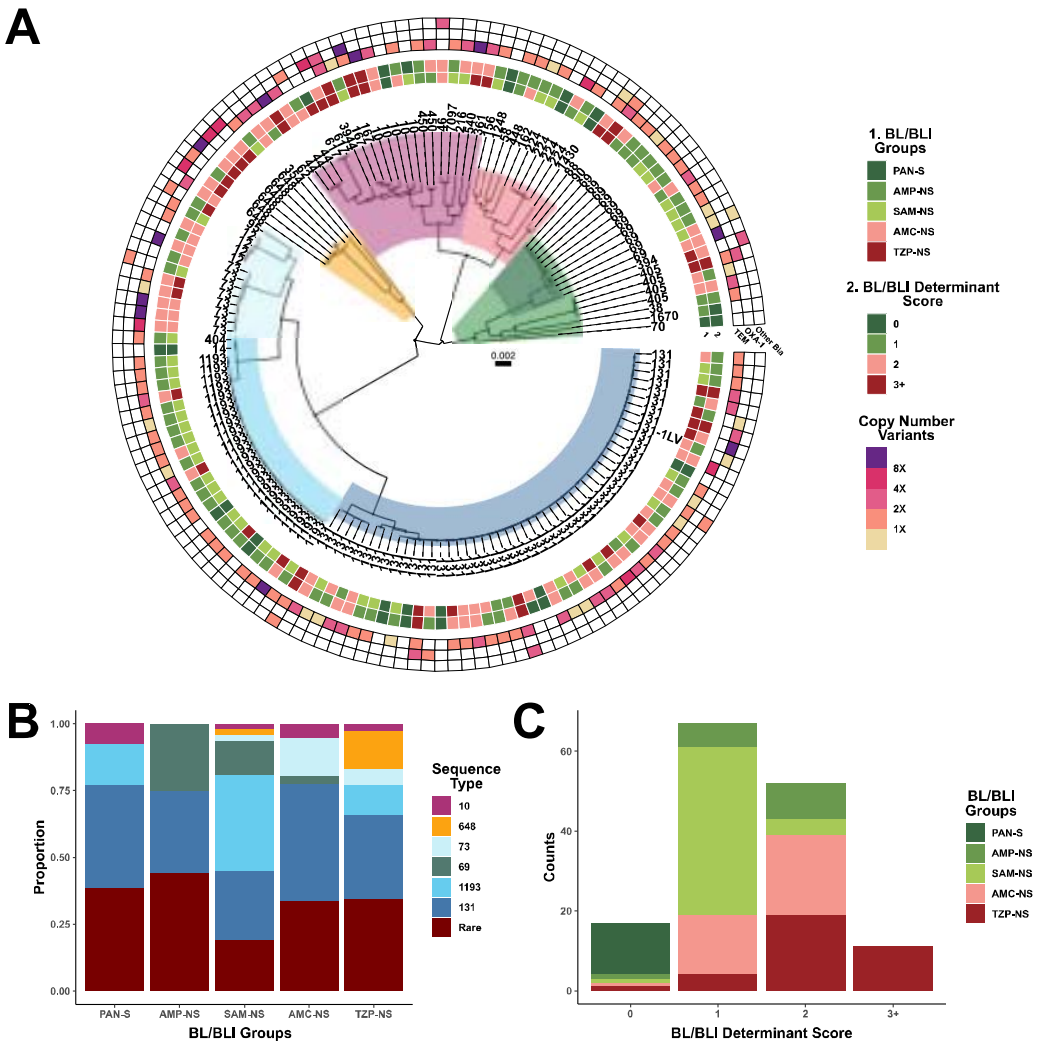

FIG 3. BL/BLI genetic determinant distribution across ESRI population. (A) Core gene alignment inferred maximum likelihood phylogeny of 147 CRO-S *E. coli* Bloodstream Isolates. The tree is mid-point rooted with the timescale indicating mean nucleotide substitutions per site. Clades are shaded by core hierarchical population structure that strongly corresponds with phylogroup. Branch tips are labelled by respective sequence type. Inner rings 1 and 2 correspond to BL/BLI groups and BL/BLI determinant scores respectively as labelled in legend. Outer ring corresponds to bla_{TEM}, bla_{OXA-1}, and other *bla* gene copy number variants with copy number estimates color coded as indicated in legend. (B) Stacked bar-chart of the total proportion of STs across each of the BL/BLI Groups. (C) Stacked bar-chart of frequency counts of BL/BLI groups across each of the BL/BLI determinant scores.

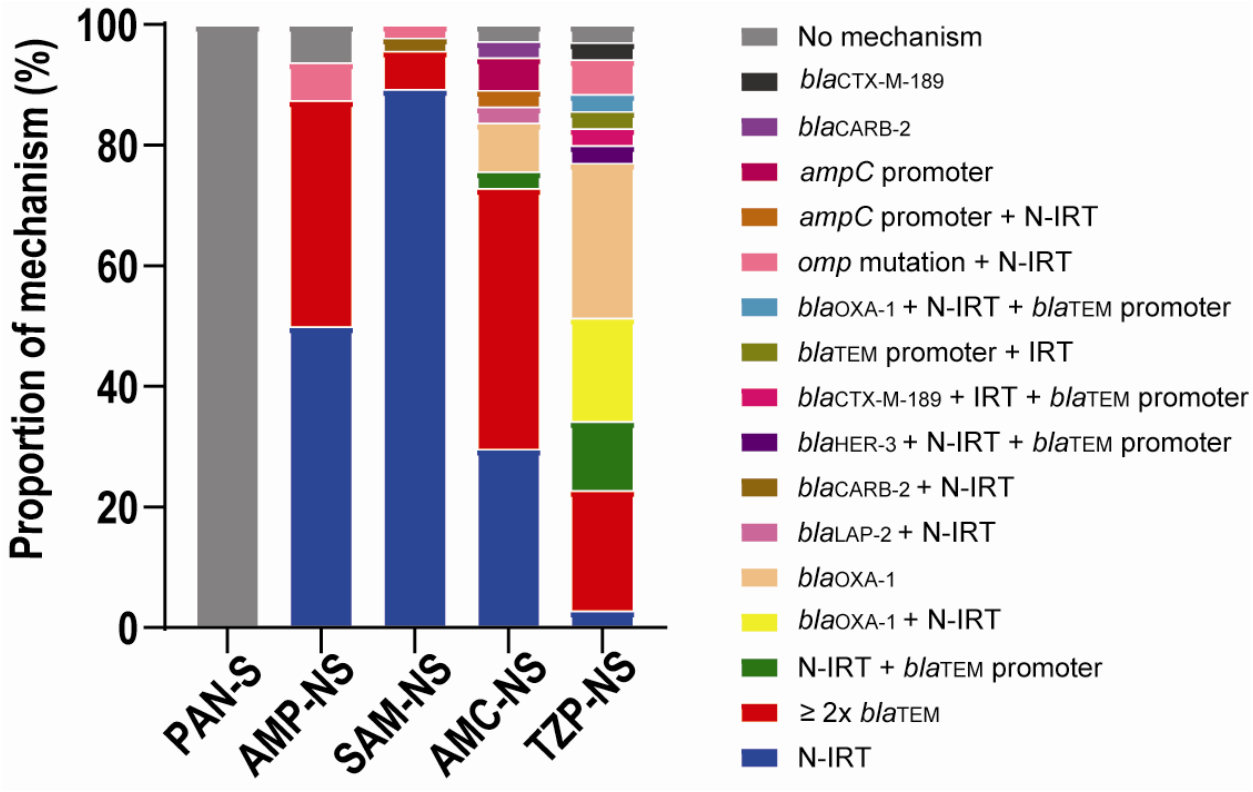

FIG S2. Proportion of BL/BLI genetic mechanisms detected across *E. coli* isolates by BL/BLI phenotype. Each of the respective combinations of genetic mechanisms are listed across each of the BL/BLI phenotypes and labelled accordingly in the legend.

Table 1. Ordinal Logistic Regression to measure covariate associations with BL/BLI Groups			
Feature	**Mean (SD)/n (%)**	**OR[a]**	**95% CI[b]**
BL/BLI Genetic Determinant Score[c]	1.4 (0.9)	10.9	6.2 – 20.2
N-IRT Presence	110 (74.8)	1.1	0.5 – 2.4
$bla_{TEM} \geq 2\times$	40 (27.2)	2.8	1.4 – 5.4
bla_{OXA-1}	19 (12.9)	33.6	9.0 – 125.1
Other *bla* $\geq 2\times$	14 (9.5)	67.8	8.5 – 543.5
Hierarchical Cluster[d]	-	-	-
Cluster 2 (B1/C)	12 (8.2)	0.3	0.1 – 0.9
Cluster 3 (D – ST405-like)	7 (4.8)	0.8	0.2 – 3.8
Cluster 4 (D – ST69-like)	12 (8.2)	0.3	0.1 – 1.0
Cluster 5 (B2 – ST1193-like)	25 (17.0)	0.5	0.2 – 1.1
Cluster 6 (A)	20 (13.6)	1.9	0.7 – 5.0
Cluster 7 (B2 – ST73-like)	13 (8.8)	2.1	0.8 – 5.9
Cluster 8 (F – ST648-like)	8 (5.4)	4.5	1.0 – 20.9

[a]Odds ratio
[b]95 percent confidence interval
[c]BL/BLI genetic determinant score composite variable of ten binary coded variables (See methods)
[d]Each group is compared to Cluster 1 which is B2 phylogroup ST131 isolates; parenthesis includes phylogroup and most common ST group where applicable

317

Discussion

The increasing affordability and availability of WGS data has markedly increased understanding of bacterial AMR to the point where it has been postulated that such analysis might augment or even replace routine phenotypic susceptibility testing (33). For fast growing bacteria, such as Enterobacterales, WGS has shown good concordance with AMR phenotype for many "bug-drug" combinations, but in general WGS has not accurately predicated Enterobacterales susceptibility to commonly used BL/BLIs such as SAM, AMC, and TZP (3, 5, 34). Herein, we sought to determine whether a comprehensive WGS analysis, including assessment of *bla* copy number, could separate cephalosporin-susceptible *E. coli* based on BL/BLI resistance patterns. Our data indicate that an increase in copy numbers of narrow spectrum β-lactamases does occur with progressive BL/BLI resistance but accounts for only part of this process in conjunction with an accumulation of a variety of other mechanisms (11).

A key strength of our study was using WGS data to analyze a phenotypically diverse cohort of *E. coli* bloodstream isolates from across the spectrum of BL/BLI resistance in contrast to previous studies which investigated a single BL/BLI combination (6, 14, 18). Moreover, we focused on cephalosporin susceptible strains and integrated analyses of *bla* gene copy numbers to elucidate the recently proposed concept of extended-spectrum resistance to BL/BLI (ESRI) (11). This approach allowed for augmented understanding of how each β-lactam resistance mechanism contributes to the spectrum of BL/BLI resistance in clinical isolates. For example, mutations in the *ampC* promoter were associated with resistance up through AMC but not TZP whereas mutations in the bla_{TEM} promoter were observed only in fully resistant strains except for a single isolate with a borderline TZP MIC of 16 mg/L. Similarly, although bla_{OXA-1} has been associated with TZP resistance (14), we observed that strains in which bla_{OXA-1} was the only β-lactamase and did not have increase copy number remained TZP susceptible although

resistant to AMC and SAM suggesting a necessary gene dosage and/or OXA-1 activity effect for progressive resistance. Given that *bla*$_{OXA-1}$ amplification is clearly associated with TZP resistance (14, 15), these and other data raise concerns regarding the clinical efficacy of TZP therapy of *bla*$_{OXA-1}$ containing strains even if they test initially susceptible given the capacity of such strains to readily increase *bla*$_{OXA-1}$ copy number (35).

A major surprising finding of our study was that whereas there was a progressive increase in *bla*$_{TEM}$ copy numbers moving from SAM-NS to AMC-NS to TZP-NS strains, there was conversely a small decrease in *bla*$_{TEM}$ copy numbers between AMP-NS (Group 2) and SAM-NS (Group 3) isolates. Such data were even more surprising given that *bla*$_{TEM-1}$ was nearly universal and the only β-lactam resistance mechanism for both groups, leading us to hypothesize that increased TEM-1 production due to gene amplification would be the major mechanism distinguishing these two groups (11). Differentiation between AMP-NS and SAM-NS *E. coli* has not been systemically studied using WGS data such as has been done for AMC-NS and TZP-NS isolates (6, 11, 13, 14, 17), and thus we are limited in our ability to benchmark our data against others. Noguchi et al. studied SAM resistant isolates from Japan using phenotypic assays and found some relationship between increased TEM-1 activity, decreased membrane permeability, and SAM resistance (18). Consistent with our data, they found only a small percentage of strains had inactivating mutations in *ompC* or *ompF*. Thus, together with other data regarding *E. coli* responses to antimicrobials, we hypothesize that *bla*$_{TEM-1}$ harboring *E. coli* develops SAM-NS through an increase in TEM-1 activity that is not detectable via assessment of gene copy number when the organism is grown in the absence of antimicrobial pressure, along with decreased SAM entry through outer membrane changes which are not due to defined genetic mutations. As *E. coli* moves from SAM-NS to AMC-NS to TZP-NS, the resistance mechanisms then become both more diverse and genetically fixed in the population such that they become detectable using WGS based methodologies. Thus, barring identification

of new genetic mechanisms that can separate AMP-NS and SAM-NS isolates, it seems unlikely that WGS methods using bacteria grown in the absence of antimicrobial pressure will be able to accurately separate these two groups. Consistent with data from other groups (6, 14), WGS analyses are likely to be useful in detecting AMC and TZP resistance in *E. coli*, but need to consider a large group of potential mechanisms, including single nucleotide polymorphisms (SNPs), which are additive in nature and thus much more difficult to predict relative to ESBL phenotypes primarily driven by single genes. There is growing interest to create 'mechanism agnostic' machine learning models that can classify AST phenotypes based on detecting associations with SNPs and gene presence/absence after controlling for population structure (36-39).

There are several limitations to our study worth noting. First, we sought to provide a "real-world" assessment of using WGS to analyze BL/BLI resistance and thus did not recapitulate the phenotypic data nor did we assess BL/BLI phenotypes comparing different methodologies such as broth microdilution vs. agar dilution. Thus, given the difficulties in the accuracy of BL/BLI susceptibilities using automated methodologies (21), it is likely that there was some misclassification of organisms into the various groups, particularly for those whose MICs were near the susceptibility breakpoints. Second, we used the WGS data to assess for known β-lactam resistance mechanisms and thus cannot be certain that other mechanisms were not present. Even though we assessed nearly 150 strains, our sample sizes for some groups, such as group 2, were relatively small which could have limited our power to detect statistically significant differences in various mechanisms. Finally, we only sequenced strains from a single center meaning we do not know the generalizability of our findings although our data are in accord with other data from geographically distant locales (11-13, 16, 19).

In summary, we present herein a comprehensive WGS analysis of a large cohort of cephalosporin-susceptible *E. coli* strains from across the susceptibility spectrum of commonly used BL/BLIs. Our data both support and add to the complexity of the progressive ESRI spectrum proposed by Rodriguez-Villodres et al (11). Our data suggest that the addition of β-lactam copy number and promoter variation assessment does assist with delineation of AMC-NS and TZP-NS but not SAM-NS suggesting that the initial movement of *E. coli* along the ESRI spectrum may not be detectable using methods that query AMR databases.

Materials and Methods

Bacterial strains and AST characterization

Escherichia coli bloodstream isolates are collected weekly through an IRB approved protocol and stored at -80C in 40% glycerol stocks. Initial antimicrobial susceptibility testing (AST) is performed on a Vitek®2 (bioMérieux) automated platform through the University of Texas MD Anderson Cancer Center clinical microbiology laboratory. Additional AST was performed on selected candidate isolates to determine cohort inclusion using gradient ETest strips (Liofilchem) for SAM, AMC, and TZP respectively on MHB plates. Isolates with 'intermediate' susceptibilities to respective BL/BLI were grouped with isolates with interpretations of 'resistant' within the same BL/BLI group and thus defined as non-susceptible to that particular BL/BLI combination. Further AST information is included in Table S2.

Whole genome sequencing of bacterial isolates

Isolates with no indication of contamination and that fit into one of five groups based on AST confirmation were sub-cultured in LB and incubated for 3 hours at 37°C with 225 rpm agitation. A culture cell pellet was used to perform genomic DNA extraction using the QIAGEN DNeasy

Blood & Tissue Kit following manufacturer's instructions. The quality of gDNA was assessed through a measure of gDNA concentration with the Qubit 4 fluorometer (Invitrogen) and high molecular weight with the 4200 Tapestation (Agilent) system using the Genomic DNA ScreenTape assay. The respective QC gDNA was submitted to the MDACC Advanced Technology Genomics Core (ATGC) for library preparation and short-read, paired-end (150 × 2) whole genome sequencing using the Illumina NovaSeq6000 instrument. Samples were multiplexed, barcoded, and sequenced to achieve coverage depths > 100×.

Short-read sequencing data QC and copy number quantification

Short-read data was QC'd and processed through a bespoke pipeline (Shropshire W, spades_pipeline, GitHub: https://github.com/wshropshire/SPAdes_pipeline). Briefly, paired-end fastq reads have Illumina adaptors and low-quality bases trimmed using Trimmomatic-v0.39 with a seed-match (16 bp) that allows up to 2 mismatches, a sliding window of 4 bp, minimum average quality = 15, and minimum length of 36 bp. Trimmed fastq reads were then quality assessed using fastqc-v0.11.9. Sequencing QC information as well as BioSample IDs are provided in Table S2.

Trimmed reads were inputted into the COpy Number Variant QuantifICation Tool (CONVICT; w Shropshire; GitHub: https://github.com/wshropshire/convict). A core-gene control file, which is created through the panaroo pan_reference.fa file, is used to normalize coverage depths. CONVICT uses a coverage-depth normalization approach whereby the coverage-depth of target genes are divided by the coverage-depth of control genes to give an estimation of ploidy-agnostic copy number. In the first step of the CONVICT algorithm, reads are aligned with a target set of antimicrobial resistance genes determined by kmerresistance (40, 41) with the resfinder database (v4.0) (42, 43). Aligned reads are then sorted with Samtools into a

corresponding bam file. Coverage measurements are determined by pileup.sh from the BBMAP suite of tools in bins of 100 bases along the length of a target gene. Bins are removed or maintained in an iterative process by comparing coverage-depth of a given bin to the mean coverage-depth of all bins for the respective gene until a specific coefficient of variation (CV = 0.175) is met. Coverage-depth for control genes is calculated in the same way as for the target genes. Copy numbers are then estimated by dividing the coverage-depth of target genes by the coverage-depth of control genes. Complete convict estimated copy number results are provided in Table S2.

Short-read assembly, database query, and phylogenetics

Trimmed reads are then used as input for short-read genome assemblies using SPAdes-v.3.15.5 using the –isolate parameter (44). Short-read assemblies were used as input to AMRfinder plus (v3.11.11) using database version 2023-04-17.1 to confirm AMR variant alleles (45) detected with the short-read based convict tool. We used known strong bla_{TEM} promoter variants (22) and *ampC* promoter variants (24) that have been experimentally characterized as database input into a variant annotator tool (Selvalakshmi27, VariantAnnotator, https://github.com/Selvalakshmi27/VariantAnnotator). Each variant was confirmed using blast searches as well as visually inspected on SnapGene-v5.0.8. Genome assemblies were first annotated using Prokka-v1.14.6 (46) Annotated .gff files and were used to produce core gene alignment file using panaroo-v1.2.10 (47). We used the MAFT-v7.505 aligner (48), with the panaroo parameters strict mode and a core gene threshold of greater than 99% present within the population (n=147). The core gene alignment file then was used to generate a maximum-likelihood (ML) phylogeny using IQtree2- 2.2.0-beta. Parameterization included 1000 replicates for SH approximate likelihood ratio test, 1000 replicates for ultrafast bootstrap (49), and model optimization using ModelFinder (50). The optimized ML model was an unrestricted model with optimized base frequencies using a maximum-likelihood with a FreeRate of heterogeneity

chosen according to BIC. Core gene alignment file was used as input to investigate population structure using an implementation of the hierarchical population clustering tool, RhierBAPS (30) with a max depth of hierarchical search =2 and max populations = 30. The ggtree-v3.3.1 R package was used for tree visualization and mapping metadata to each individual tree branch-tip (51).

Statistics

Genetic determinants were coded into ten binary variables to create a cumulative BL/BLI genetic determinant score (Table S3). These variables included N-IRT presence, TEM gene amplification (≥2×), IRT gene, strong bla_{TEM} promoter variant, *ampC* promoter variant, bla_{OXA-1}, bla_{CTX-M}, Other *bla* gene, Other *bla* gene amplification (≥2×), and predicted Omp mutations based on previous research (11, 18, 22, 24, 28, 32). A Spearman's rho coefficient was calculated to test for relationship between BL/BLI groups and BL/BLI genetic determinant score. Global Wilcoxon Rank Sum or Kruskal Wallis Tests were performed to detect significant differences in log transformed copy numbers between groups. For global comparisons with significant *P*-values, pairwise comparisons using Wilcoxon rank sum exact test were performed with a false discovery rate *P*-value adjustment method. Ordinal logistic regression was performed using BL/BLI group as an ordinal outcome and various genetic features as independent covariates with the MASS R package (v.7.3-54) and the 'polr' function with a proportional odds logistic regression model. The brant test was performed to determine if parallel regression assumption held for each model using the 'brant' R package (v0.3-0). Statistical analysis was performed using R-v4.0.4.

Data Availability

Whole genome sequencing short-read data has been submitted to BioProject PRJNA924946 and PRJNA836696. R scripts for analyses can be made available through request of corresponding author.

Acknowledgements

The MDACC clinical microbiology lab does a fantastic job identifying and transferring these organisms of interest to us and as always is much appreciated. Core grant CA016672(ATGC) and NIH 1S10OD024977-01 grant provide funding for the Advanced Technology Genomics Core (ATGC) sequencing facility at MD Anderson Cancer Center. The National Institute of Allergy and Infectious Diseases (NIAID) T32 AI141349 Training Program in Antimicrobial Resistance supports the work of W.C.S. The NIAID R21AI151536 and P01AI152999 NIAID grants support work for this project for S.A.S. The authors acknowledge the support of the High-Performance Computing for research facility at the University of Texas MD Anderson Cancer Center for providing computational resources that have contributed to the research results reported in this paper.

References

1. Alleweldt F, Kara Ş, Best K, Aarestrup FM, Beer M, Bestebroer TM, Campos J, Casadei G, Chinen I, Van Domselaar G, Dominguez C, Everett HE, Fouchier RA, Grant K, Green J, Höper D, Johnston J, Koopmans MP, Oude Munnink BB, Myers R, Nadon C, Patel A, Pohlmann A, Pongolini S, Reimer A, Thiessen S, Wylezich C. 2021. Economic evaluation of whole genome sequencing for pathogen identification and surveillance – results of case studies in Europe and the Americas 2016 to 2019. Eurosurveillance 26.
2. Gordon LG, Elliott TM, Forde B, Mitchell B, Russo PL, Paterson DL, Harris PNA. 2021. Budget impact analysis of routinely using whole-genomic sequencing of six multidrug-resistant bacterial pathogens in Queensland, Australia. BMJ Open 11:e041968.
3. Shelburne SA, Kim J, Munita JM, Sahasrabhojane P, Shields RK, Press EG, Li X, Arias CA, Cantarel B, Jiang Y, Kim MS, Aitken SL, Greenberg DE. 2017. Whole-Genome Sequencing Accurately Identifies Resistance to Extended-Spectrum beta-Lactams for Major Gram-Negative Bacterial Pathogens. Clin Infect Dis 65:738-745.
4. Doyle RM, O'Sullivan DM, Aller SD, Bruchmann S, Clark T, Coello Pelegrin A, Cormican M, Diez Benavente E, Ellington MJ, McGrath E, Motro Y, Phuong Thuy Nguyen T, Phelan J, Shaw LP, Stabler RA, Van Belkum A, Van Dorp L, Woodford N, Moran-Gilad J, Huggett JF, Harris KA. 2020.

513 Discordant bioinformatic predictions of antimicrobial resistance from whole-genome sequencing
514 data of bacterial isolates: an inter-laboratory study. Microbial Genomics 6.
515 5. Hujer AM, Long SW, Olsen RJ, Taracila MA, Rojas LJ, Musser JM, Bonomo RA. 2020. Predicting β-
516 lactam resistance using whole genome sequencing in Klebsiella pneumoniae: the challenge of β-
517 lactamase inhibitors. Diagnostic Microbiology and Infectious Disease 98:115149.
518 6. Davies TJ, Stoesser N, Sheppard AE, Abuoun M, Fowler P, Swann J, Quan TP, Griffiths D, Vaughan
519 A, Morgan M, Phan HTT, Jeffery KJ, Andersson M, Ellington MJ, Ekelund O, Woodford N,
520 Mathers AJ, Bonomo RA, Crook DW, Peto TEA, Anjum MF, Walker AS. 2020. Reconciling the
521 Potentially Irreconcilable? Genotypic and Phenotypic Amoxicillin-Clavulanate Resistance in
522 Escherichia coli. Antimicrob Agents Chemother 64.
523 7. Magill SS, O'Leary E, Ray SM, Kainer MA, Evans C, Bamberg WM, Johnston H, Janelle SJ,
524 Oyewumi T, Lynfield R, Rainbow J, Warnke L, Nadle J, Thompson DL, Sharmin S, Pierce R, Zhang
525 AY, Ocampo V, Maloney M, Greissman S, Wilson LE, Dumyati G, Edwards JR, Emerging Infections
526 Program Hospital Prevalence Survey T. 2021. Antimicrobial Use in US Hospitals: Comparison of
527 Results From Emerging Infections Program Prevalence Surveys, 2015 and 2011. Clin Infect Dis
528 72:1784-1792.
529 8. Control CfD, Prevention, Control CfD, Prevention. 2021. Outpatient antibiotic prescriptions—
530 United States, 2021. CDC: Atlanta, GA, USA.
531 9. Thompson ND, Stone ND, Brown CJ, Penna AR, Eure TR, Bamberg WM, Barney GR, Barter D,
532 Clogher P, DeSilva MB. 2021. Antimicrobial use in a cohort of US nursing homes, 2017. Jama
533 325:1286-1295.
534 10. Rajer F, Allander L, Karlsson PA, Sandegren L. 2022. Evolutionary Trajectories toward High-Level
535 β-Lactam/β-Lactamase Inhibitor Resistance in the Presence of Multiple β-Lactamases.
536 Antimicrobial Agents and Chemotherapy:e00290-22.
537 11. Rodriguez-Villodres A, Gil-Marques ML, Alvarez-Marin R, Bonnin RA, Pachon-Ibanez ME, Aguilar-
538 Guisado M, Naas T, Aznar J, Pachon J, Lepe JA, Smani Y. 2019. Extended-spectrum resistance to
539 beta-lactams/beta-lactamase inhibitors (ESRI) evolved from low-level resistant Escherichia coli. J
540 Antimicrob Chemother 75:77-85.
541 12. Hansen KH, Andreasen MR, Pedersen MS, Westh H, Jelsbak L, Schonning K. 2019. Resistance to
542 piperacillin/tazobactam in Escherichia coli resulting from extensive IS26-associated gene
543 amplification of blaTEM-1. J Antimicrob Chemother 74:3179-3183.
544 13. Hubbard ATM, Mason J, Roberts P, Parry CM, Corless C, Van Aartsen J, Howard A, Bulgasim I,
545 Fraser AJ, Adams ER, Roberts AP, Edwards T. 2020. Piperacillin/tazobactam resistance in a
546 clinical isolate of Escherichia coli due to IS26-mediated amplification of blaTEM-1B. Nature
547 Communications 11.
548 14. Edwards T, Heinz E, Van Aartsen J, Howard A, Roberts P, Corless C, Fraser AJ, Williams CT,
549 Bulgasim I, Cuevas LE, Parry CM, Roberts AP, Adams ER, Mason J, Hubbard ATM. 2022.
550 Piperacillin/tazobactam-resistant, cephalosporin-susceptible Escherichia coli bloodstream
551 infections are driven by multiple acquisition of resistance across diverse sequence types.
552 Microbial Genomics 8.
553 15. Shropshire WC, Aitken SL, Pifer R, Kim J, Bhatti MM, Li X, Kalia A, Galloway-Peña J,
554 Sahasrabhojane P, Arias CA, Greenberg DE, Hanson BM, Shelburne SA. 2021. IS26-mediated
555 amplification of blaOXA-1 and blaCTX-M-15 with concurrent outer membrane porin disruption
556 associated with de novo carbapenem resistance in a recurrent bacteraemia cohort. Journal of
557 Antimicrobial Chemotherapy 76:385-395.
558 16. Schechter LM, Creely DP, Garner CD, Shortridge D, Nguyen H, Chen L, Hanson BM, Sodergren E,
559 Weinstock GM, Dunne WM. 2018. Extensive gene amplification as a mechanism for piperacillin-
560 tazobactam resistance in Escherichia coli. MBio 9:e00583-18.

561 17. Livermore DM, Day M, Cleary P, Hopkins KL, Toleman MA, Wareham DW, Wiuff C, Doumith M,
562 Woodford N. 2019. OXA-1 beta-lactamase and non-susceptibility to penicillin/beta-lactamase
563 inhibitor combinations among ESBL-producing Escherichia coli. J Antimicrob Chemother 74:326-
564 333.
565 18. Noguchi T, Matsumura Y, Kanahashi T, Tanaka M, Tsuchido Y, Matsumura T, Nakano S,
566 Yamamoto M, Nagao M, Ichiyama S. 2019. Role of TEM-1 β-lactamase in the predominance of
567 ampicillin-sulbactam-nonsusceptible Escherichia coli in Japan. Antimicrobial Agents and
568 Chemotherapy 63:e02366-18.
569 19. Waltner-Toews RI, Paterson DL, Qureshi ZA, Sidjabat HE, Adams-Haduch JM, Shutt KA, Jones M,
570 Tian GB, Pasculle AW, Doi Y. 2011. Clinical characteristics of bloodstream infections due to
571 ampicillin-sulbactam-resistant, non-extended- spectrum-beta-lactamase-producing Escherichia
572 coli and the role of TEM-1 hyperproduction. Antimicrob Agents Chemother 55:495-501.
573 20. Bush K, Bradford PA. 2016. β-Lactams and β-Lactamase Inhibitors: An Overview. Cold Spring
574 Harbor Perspectives in Medicine 6:a025247.
575 21. Rodriguez-Villodres A, Gutierrez Linares A, Galvez-Benitez L, Pachon J, Lepe JA, Smani Y. 2021.
576 Semirapid Detection of Piperacillin/Tazobactam Resistance and Extended-Spectrum Resistance
577 to beta-Lactams/beta-Lactamase Inhibitors in Clinical Isolates of Escherichia coli. Microbiol
578 Spectr 9:e0080121.
579 22. Lartigue MF, Leflon-Guibout V, Poirel L, Nordmann P, Nicolas-Chanoine M-H. 2002. Promoters
580 P3, Pa/Pb, P4, and P5 upstream from bla TEM genes and their relationship to β-lactam
581 resistance. Antimicrobial agents and chemotherapy 46:4035-4037.
582 23. Zhou K, Tao Y, Han L, Ni Y, Sun J. 2019. Piperacillin-tazobactam (TZP) resistance in Escherichia
583 coli due to hyperproduction of TEM-1 β-lactamase mediated by the promoter Pa/Pb. Frontiers in
584 microbiology 10:833.
585 24. Peter-Getzlaff S, Polsfuss S, Poledica M, Hombach M, Giger J, Bottger EC, Zbinden R, Bloemberg
586 GV. 2011. Detection of AmpC beta-lactamase in Escherichia coli: comparison of three
587 phenotypic confirmation assays and genetic analysis. J Clin Microbiol 49:2924-32.
588 25. Lee J, Oh CE, Choi EH, Lee HJ. 2013. The impact of the increased use of piperacillin/tazobactam
589 on the selection of antibiotic resistance among invasive Escherichia coli and Klebsiella
590 pneumoniae isolates. International Journal of Infectious Diseases 17:e638-e643.
591 26. Beceiro A, Maharjan S, Gaulton T, Doumith M, Soares NC, Dhanji H, Warner M, Doyle M, Hickey
592 M, Downie G, Bou G, Livermore DM, Woodford N. 2011. False extended-spectrum beta-
593 lactamase phenotype in clinical isolates of Escherichia coli associated with increased expression
594 of OXA-1 or TEM-1 penicillinases and loss of porins. J Antimicrob Chemother 66:2006-10.
595 27. Salverda ML, De Visser JA, Barlow M. 2010. Natural evolution of TEM-1 beta-lactamase:
596 experimental reconstruction and clinical relevance. FEMS Microbiol Rev 34:1015-36.
597 28. Rosenkilde CEH, Munck C, Porse A, Linkevicius M, Andersson DI, Sommer MOA. 2019. Collateral
598 sensitivity constrains resistance evolution of the CTX-M-15 β-lactamase. Nature
599 Communications 10.
600 29. Shropshire WC, Konovalova A, McDaneld P, Gohel M, Strope B, Sahasrabhojane P, Tran CN,
601 Greenberg D, Kim J, Zhan X, Aitken S, Bhatti M, Savidge TC, Treangen TJ, Hanson BM, Arias CA,
602 Shelburne SA. 2022. Systematic Analysis of Mobile Genetic Elements Mediating Beta-Lactamase
603 Gene Amplification in Noncarbapenemase-Producing Carbapenem-Resistant Enterobacterales
604 Bloodstream Infections. Msystems 7:e00476-22.
605 30. Tonkin-Hill G, Lees JA, Bentley SD, Frost SDW, Corander J. 2018. RhierBAPS: An R
606 implementation of the population clustering algorithm hierBAPS. Wellcome Open Res 3:93.

607 31. Clermont O, Christenson JK, Denamur E, Gordon DM. 2013. The Clermont Escherichia coli phylo-
608 typing method revisited: improvement of specificity and detection of new phylo-groups.
609 Environmental Microbiology Reports 5:58-65.
610 32. Oteo J, Gonzalez-Lopez JJ, Ortega A, Quintero-Zarate JN, Bou G, Cercenado E, Conejo MC,
611 Martinez-Martinez L, Navarro F, Oliver A, Bartolome RM, Campos J, Spanish Network for
612 Research in Infectious D. 2014. Inhibitor-resistant TEM- and OXA-1-producing Escherichia coli
613 isolates resistant to amoxicillin-clavulanate are more clonal and possess lower virulence gene
614 content than susceptible clinical isolates. Antimicrob Agents Chemother 58:3874-81.
615 33. Yee R, Simner PJ. 2022. Next-Generation Sequencing Approaches to Predicting Antimicrobial
616 Susceptibility Testing Results. Clin Lab Med 42:557-572.
617 34. Vanstokstraeten R, Piérard D, Crombé F, De Geyter D, Wybo I, Muyldermans A, Seyler L, Caljon
618 B, Janssen T, Demuyser T. 2023. Genotypic resistance determined by whole genome sequencing
619 versus phenotypic resistance in 234 Escherichia coli isolates. Scientific Reports 13.
620 35. Harris PNA, Tambyah PA, Lye DC, Mo Y, Lee TH, Yilmaz M, Alenazi TH, Arabi Y, Falcone M,
621 Bassetti M, Righi E, Rogers BA, Kanj S, Bhally H, Iredell J, Mendelson M, Boyles TH, Looke D,
622 Miyakis S, Walls G, Al Khamis M, Zikri A, Crowe A, Ingram P, Daneman N, Griffin P, Athan E,
623 Lorenc P, Baker P, Roberts L, Beatson SA, Peleg AY, Harris-Brown T, Paterson DL, Investigators
624 MT, the Australasian Society for Infectious Disease Clinical Research N. 2018. Effect of
625 Piperacillin-Tazobactam vs Meropenem on 30-Day Mortality for Patients With E coli or Klebsiella
626 pneumoniae Bloodstream Infection and Ceftriaxone Resistance: A Randomized Clinical Trial.
627 JAMA 320:984-994.
628 36. Ren Y, Chakraborty T, Doijad S, Falgenhauer L, Falgenhauer J, Goesmann A, Hauschild AC,
629 Schwengers O, Heider D. 2022. Prediction of antimicrobial resistance based on whole-genome
630 sequencing and machine learning. Bioinformatics 38:325-334.
631 37. Kim JI, Maguire F, Tsang KK, Gouliouris T, Peacock SJ, McAllister TA, McArthur AG, Beiko RG.
632 2022. Machine learning for antimicrobial resistance prediction: current practice, limitations, and
633 clinical perspective. Clinical Microbiology Reviews 35:e00179-21.
634 38. Avershina E, Sharma P, Taxt AM, Singh H, Frye SA, Paul K, Kapil A, Naseer U, Kaur P, Ahmad R.
635 2021. AMR-Diag: Neural network based genotype-to-phenotype prediction of resistance
636 towards beta-lactams in Escherichia coli and Klebsiella pneumoniae. Comput Struct Biotechnol J
637 19:1896-1906.
638 39. Kim J, Greenberg DE, Pifer R, Jiang S, Xiao G, Shelburne SA, Koh A, Xie Y, Zhan X. 2020. VAMPr:
639 VAriant Mapping and Prediction of antibiotic resistance via explainable features and machine
640 learning. PLoS Comput Biol 16:e1007511.
641 40. Clausen PTLC, Aarestrup FM, Lund O. 2018. Rapid and precise alignment of raw reads against
642 redundant databases with KMA. BMC Bioinformatics 19.
643 41. Clausen PTLC, Zankari E, Aarestrup FM, Lund O. 2016. Benchmarking of methods for
644 identification of antimicrobial resistance genes in bacterial whole genome data. Journal of
645 Antimicrobial Chemotherapy 71:2484-2488.
646 42. Florensa AF, Kaas RS, Clausen PTLC, Aytan-Aktug D, Aarestrup FM. 2022. ResFinder – an open
647 online resource for identification of antimicrobial resistance genes in next-generation
648 sequencing data and prediction of phenotypes from genotypes. Microbial Genomics 8.
649 43. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, Philippon A, Allesoe RL, Rebelo
650 AR, Florensa AF, Fagelhauer L, Chakraborty T, Neumann B, Werner G, Bender JK, Stingl K,
651 Nguyen M, Coppens J, Xavier BB, Malhotra-Kumar S, Westh H, Pinholt M, Anjum MF, Duggett
652 NA, Kempf I, Nykasenoja S, Olkkola S, Wieczorek K, Amaro A, Clemente L, Mossong J, Losch S,
653 Ragimbeau C, Lund O, Aarestrup FM. 2020. ResFinder 4.0 for predictions of phenotypes from
654 genotypes. J Antimicrob Chemother 75:3491-3500.

44. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, 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 19:455-77.
45. Feldgarden M, Brover V, Gonzalez-Escalona N, Frye JG, Haendiges J, Haft DH, Hoffmann M, Pettengill JB, Prasad AB, Tillman GE, Tyson GH, Klimke W. 2021. AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. Scientific Reports 11.
46. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068-9.
47. Tonkin-Hill G, Macalasdair N, Ruis C, Weimann A, Horesh G, Lees JA, Gladstone RA, Lo S, Beaudoin C, Floto RA, Frost SDW, Corander J, Bentley SD, Parkhill J. 2020. Producing polished prokaryotic pangenomes with the Panaroo pipeline. Genome Biology 21.
48. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772-80.
49. Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. Molecular Biology and Evolution 35:518-522.
50. Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nature Methods 14:587-589.
51. Yu G, Smith DK, Zhu H, Guan Y, Lam TTY. 2017. ggtree: package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution 8:28-36.