MOLECULAR EPIDEMIOLOGY OF *SERRATIA MARCESCENS* COMPLEX INFECTIONS

MOLECULAR EPIDEMIOLOGY OF CLINICAL INFECTIONS CAUSED BY *SERRATIA MARCESCENS* COMPLEX IN A TERTIARY CARE HOSPITAL SYSTEM

By ADAM S. KOMOROWSKI, HBSc MBBS

A Thesis Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements for the Degree Master of Science

McMaster University MASTER OF SCIENCE (2025) Hamilton, Ontario (Health Research Methodology)

TITLE: Molecular Epidemiology of Clinical Infections Caused by *Serratia marcescens* Complex in a Tertiary Care Hospital System

AUTHOR: Adam S. Komorowski, HBSc MBBS

SUPERVISOR: Professor Dominik Mertz, MD MSc

NUMBER OF PAGES: xv and 114

Lay Abstract

This study looked at a group of bacteria called the *Serratia marcescens* complex, which can cause infections in hospital patients. To better understand whether these bacteria spread between patients — something that could change infection control policies — we collected samples from infected patients in six Canadian hospitals over two years. Using advanced genetic tests, we found that although standard laboratory methods identified all infections as being caused by a single *Serratia* species, many different species were involved. Most cases did not appear to be spread between patients, but two possible cases of transmission were found. Our findings suggest that current laboratory tools may not be accurate enough to fully identify these bacteria, and that improvements are needed.

Abstract

Background: *Serratia marcescens* is a Gram-negative opportunistic pathogen associated with outbreaks in healthcare settings. It produces an AmpC beta-lactamase, conferring resistance to many antibiotics and contributing to morbidity and mortality, particularly in intensive care units. The extent of *S. marcescens* transmission between hospitalized patients and its implications for infection control are not well understood.

Methods: From January to December 2022 and from January to June 2024, we prospectively identified and enrolled consecutive specimens taken from patients for routine bacterial culture at six hospitals. We collected relevant patient characteristics using retrospective chart review. Only the first isolate of *Serratia marcescens* from a patient that was identified via conventional bacteriologic culture was included, and the identification was confirmed by bioMérieux matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS). We performed whole genome sequencing using the Illumina NextSeq 2000 platform and used Genome Taxonomy Database Toolkit (GTDB-Tk) taxonomies to construct a maximum-likelihood phylogenetic tree. We queried assembled genomes against the Comprehensive Antibiotic Resistance Database (CARD) to predict determinants of antimicrobial resistance. Potential transmission was assessed by calculating average nucleotide identity (ANI) between isolates from different patients on the same ward within one month.

Results: Of 147 identified isolates, 125 met study inclusion criteria. MALDI-TOF MS and genome-based species identification were discordant in 64 (51.2%) cases, indicating the involvement of multiple species within the recently described *S. marcescens* complex in both community- and hospital-associated infections. Two isolate pairs had a putative spatiotemporal link, with one pair meeting the ANI threshold for possible transmission.

Conclusions: Possible transmission of *S. marcescens* complex between hospital inpatients was rare. Current MALDI-TOF MS methods cannot reliably distinguish between *S. marcescens* complex members; laboratory reporting should therefore identify isolates to the complex level.

Improved species-level identification may refine infection prevention efforts by informing decisions on whether to investigate temporally clustered cases as potential outbreaks.

Acknowledgements

A thesis takes a village, and mine is no exception. Heartfelt thanks to Drs. Dominik Mertz and Marek Śmieja, my supervisor and co-supervisor, respectively: both have offered wise counsel, kindness, expertise, and ardent support professionally and personally, and I am indebted to them for their mentorship and friendship. I am also very grateful for my committee member Dr. Andrew McArthur who offered invaluable expertise, guidance, and support; and my external examiner, Dr. Jennie Johnstone. The assistance and expertise of Dr. Michael Surette and his laboratory members Laura Rossi, Dominique Tertigas, and Dr. Shahrokh Shekarriz are greatly appreciated. Thanks are also due to Mark Gaskin for helping identify clinical isolates. This work was supported by funders at the Ontario Ministry of Health, the Research Institute of St. Joseph's Healthcare Hamilton, and the McMaster University Department of Pathology and Molecular Medicine; the analyses, conclusions, opinions, and statements herein are mine alone, and do not necessarily reflect those of the funders.

I stand on the shoulders of many mentors, supervisors, and colleagues who have been instrumental in shaping and encouraging me, among them: Drs. Daniela Leto, Cheryl Main, Philippe El-Helou, Joanna Dionne, Fiona Smaill, Mark Loeb, Paul Sax, Kimberly Hanson, Romney Humphries, Mariska Leeflang, Kevin Woodward, Eva Piessens, Zain Chagla, Todd Lee, Prameet Sheth, Robert Kozak, Annette Plüddemann, Kevin Katz, Andrew Morris, Menaka Pai, Ekkehard Kasper, Roman Jaeschke, Catherine Nix, Des Leddin, Tom Lynch, Rohan D'Souza, Katarzyna Jerzak, and Rossanna Pezo. Special thanks are due to Drs. Helen MacKay, Boris Steipe, John Tuttle, and Jerzy Cichocki; as well as Peter Barley, John Woodger, and Giuliana Donato, who nurtured me out of my youth and inexperience and have left an indelibly positive mark on my life.

Most of all, I would like to acknowledge the patient unconditional love, and unfailing support of my wife, Dr. Wynn Peterson; my parents, Dr. Richard and Maria Giuseppina Komorowski; my family; and my dearest friends. Thank you for believing in me when I didn't, and for putting up with me when you shouldn't have.

This thesis is dedicated to the memory of my grandparents,

Lauretta, Luigi, Aniela, and Stanisław,

who came to Canada in the aftermath of the Second World War

to a country that accepted them with open arms,

without whom and without which none of this work would be possible.

„Unkraut wächst weiter, die Ackerwinden erwürgen die Sträucher, die gelben Wurzeln der Brennesseln kriechen unter der Erde fort, die Klettenstauden überragen einen um Haupteslänge, die Braunfäule und die Milben breiten sich aus, und sogar das Papier, auf dem man mühselig Wörter und Sätze aneinanderreiht, fühlt sich an, als sei es vom Meltau überzogen. Tage- und wochenlang zermartert man sich vergebens den Kopf, wüßte, wenn man danach befragt würde, nicht, ob man weiterschreibt aus Gewohnheit oder aus Geltungssucht, oder weil man nichts anderes gelernt hat, oder aus Verwunderung über das Leben, aus Wahrheitsliebe, aus Verzweiflung oder Empörung, ebensowenig wie man zu sagen vermöchte, ob man durch das Schreiben klüger oder verrückter wird. Vielleicht verliert ein jeder von uns den Überblick genau in dem Maß, in dem er fortbaut am eigenen Werk, und vielleicht neigen wir aus diesem Grund dazu, die zunehmende Komplexität unserer Geisteskonstruktionen zu verwechseln mit einem Fortschritt an Erkenntnis, während wir zugleich schon ahnen, daß wir die Unwägbarkeiten, die in Wahrheit unsere Laufbahn bestimmen, nie werden begreifen können.“

Dr. W.G. Sebald, Kapitel VII, *Die Ringe des Saturn* (1995)

Table of Contents

List of Illustrations, Charts, and Diagrams

List of Tables

List of Abbreviations and Symbols

ABC: ATP-binding cassette

ANI: average nucleotide identity

AST: antimicrobial susceptibility testing

BLASTP: Basic Local Alignment Search Tool - Protein

bp: base pairs

CARD: Comprehensive Antibiotic Resistance Database

CAT: chloramphenicol acetyltransferase

CI: confidence interval

CLSI: Clinical and Laboratory Standards Institute

CPE: carbapenemase-producing Enterobacterales

DNA: deoxyribonucleic acid

dNTP: deoxynucleoside triphosphate

ESBL: extended-spectrum β-lactamase

GC: guanine-cytosine

GTDB-Tk: Genome Taxonomy Database Toolkit

ICU: intensive care unit

IPAC: infection prevention and control

kb: kilobases

kbp: kilobase pair

MALDI-TOF MS: matrix-assisted laser desorption ionization time-of-flight mass spectrometry

MFS: major facilitator superfamily

MIC: minimum inhibitory concentration

MLST: multi-locus sequence typing

mpf: mating pair formation

NAG: *N*-acetylglucosamine

NAM: *N*-acetylmuramic acid

NCBI: National Center for Biotechnology Information

NICU: neonatal intensive care unit

OR: odds ratio

PBP: penicillin binding protein

PFGE: pulsed-field gel electrophoresis

QC: quality control

RND: resistance-nodulation-cell division

SNV: single nucleotide variant

STROBE: Strengthening of Reporting of Observational Studies in Epidemiology

UDP: uridine diphosphate

WGS: whole genome sequencing

Declaration of Academic Achievement

Adam S. Komorowski conceived and designed the study under the supervision of Dr. Dominik Mertz with input from Dr. Marek Śmieja, Dr. Andrew McArthur, and Dr. Michael Surette. Adam S. Komorowski performed the data analysis and wrote the initial draft of the thesis and manuscript. Dr. Dominik Mertz, Dr. Marek Śmieja, and Dr. Andrew McArthur provided revisions to the draft thesis prior to submission.

Chapter 1. Background

1.1 Serratia marcescens

The Venetian pharmacist Bartolomeo Bizio first described *Serratia marcescens* in 1823 as the cause of "bleeding polenta", spoilage of a boiled cornmeal dish ubiquitously eaten in northern Italy (1, 2). A member of the order Enterobacterales, *Serratia marcescens* is a Gram-negative bacillus that is common in the environment and is known to cause opportunistic disease in humans (2). In the environment, *S. marcescens* is frequently found in water, soil, and plant material (3). In veterinary medicine and agriculture, *Serratia marcescens* is known as a common cause of mastitis in dairy cattle and is known to promote plant growth through interactions with rhizomes (4). It is a frequent colonizer of the gastrointestinal, genitourinary, and respiratory mucosae as well as the epidermis in humans (1). *Serratia marcescens* are facultative anaerobic bacteria that are motile by virtue of the peritrichous flagellae present on most strains (1). The bacilli ferment glucose, sucrose, mannitol, and trehalose, and give positive reactions on citrate, ortho-nitrophenyl-β-D-galactoside, and Voges-Proskauer tests (2, 5). *S. marcescens* is among the few members of the order Enterobacterales to produce DNase, gelatinase, and lipase enzymes (2). Many strains produce colonies on agar media that exhibit a characteristic red colour from the production of prodigiosin (2).

The role of *S. marcescens* in causing human disease was not well established until the 1970s, partially due to the misattribution of many infections in the early 20th century to a pigment-producing "*Chromobacterium* group" of organisms: the red pigment of many *S. marcescens* strains likely underlies this errant historical classification as *Chromobacter prodigiosum* (2). Since being identified as a causative agent of human infection, *Serratia marcescens* has been found to be responsible for cases of bacteremia, pneumonia, meningitis, urinary tract infections, and intraabdominal infections. Canadian surveillance data from Calgary, Alberta between 2000 and 2005 showed an annual population incidence of 10.8 per 100,000, with no seasonal pattern to infection acquisition (6). Survey data from the Madigan Army Medical Center in the United States

illustrates that *Serratia marcescens* infections in humans mirror the likely anatomical sites of colonization, with respiratory specimens being the most common site of infection, followed by urinary and skin sources (2). Surveillance data from the late 1990s in North and South American countries shows *Serratia marcescens* to be the 12th most common organism in bloodstream infections (1.4% of isolates overall), and the 7th most commonly isolated pathogen in pneumonia, accounting for 3.5% of all hospitalized cases (7, 8).

Recently, whole genome sequencing data has cast doubt over the cohesion of the historical species *Serratia marcescens*, which may help explain the phenotypic variation observed between different *S. marcescens* isolates such as those lacking in characteristic pigmentation. *S. marcescens* was previously thought to be closely related taxonomically to *S. ureilytica,* first described in 2005 (9), and *S. nematodiphila,* first described in 2009 (10). However, a wider species complex, termed "*Serratia marcescens* complex" has recently been proposed which includes *Serratia bockelmannii, S. marcescens* sensu stricto, *S. nematodiphila, S. nevei,* and *S. ureilytica* (10-13). This re-classification is based on granular whole genome sequencing data which shows the remarkable diversity of the species that were formerly all classified taxonomically simply as *Serratia marcescens.* Genomics data highlights that within the *Serratia marcescens* complex, there is a remarkable diversity of lineages which correspond to host adaptation and ecological niche specialization: certain lineages harbour the majority of environmental isolates, while human-adapted lineages have high plasmid diversity and carry antimicrobial resistance determinants (14). To maintain consistency throughout this manuscript, when referring to the constitutive member of the newly described *Serratia marcescens* complex, we will use the term "*Serratia marcescens* sensu stricto", whereas when referencing historical data which may – or may not – include all members of the newly described *Serratia marcescens* complex, we will use the terminology "*Serratia marcescens*".

1.2 Outbreaks and transmission dynamics

Nosocomial outbreaks remain the most well-described phenomenon of *Serratia marcescens* in the medical literature (15). In a recent review of neonatal intensive care unit (NICU) outbreaks, *S. marcescens* accounted for 12.8% (n=5/39) of outbreaks overall (16). A 2005 outbreak described in a NICU in Gaza City, Palestinian Territories had a 44% mortality rate, demonstrating the potential for this organism to have severe consequences in vulnerable patients in healthcare settings (17). Predominantly environmental sources for NICU outbreaks have been described, with contaminated medical devices (18, 19), disinfectants and soaps (20, 21), sinks (22, 23), and air conditioning units (24) implicated as reservoirs. As early as 1976, studies have implicated the role of healthcare workers in propagating *Serratia* spp. in the healthcare environment, with hand hygiene lapses frequently pointed to as contributors to spread (17, 25-27).

The transmission of microorganisms takes place across both space and time: we will use the term “spatiotemporal dynamics” to describe this phenomenon throughout this thesis. The spatiotemporal dynamics of members of the *Serratia marcescens* complex continue to be poorly understood in medical practice. For example, a 2021 outbreak in Norway was believed to be a common-source outbreak of nosocomial origin, whereas a wider epidemiologic investigation revealed the outbreak had a community origin and was national in its scope (28). The outbreak investigation was eventually halted by the public health authority in this case, without a common source being identified, due to a decrease in reported cases. Most outbreaks occurring on a large geographical scale that have had their common source identified have been related to contaminated medical devices, as was the case with a 2005 multistate American outbreak from contaminated magnesium sulphate solutions (29), and both a 2008 multistate American outbreak and a 2009 outbreak in China traced to contaminated prefilled heparin syringes (30, 31). *Serratia*

marcescens isolates implicated in outbreaks tend to be resistant to many antimicrobials, further exacerbating the severity of outbreaks and their ultimate impact on patients when they occur (2).

1.3 AmpC beta-lactamases

Serratia marcescens is notable for possessing an AmpC β-lactamase enzyme, which is a type of enzyme encoded by Enterobacterales that catalyze β-lactam antimicrobials such that they are rendered ineffective. β-lactam antimicrobials target peptidoglycan, a component of the bacterial cell wall. Peptidoglycan is composed of alternating *N*-acetylglucosamine (NAG) and *N*-acetylmuramic acid (NAM) amino sugars, with a pentapeptide of L-alanine, γ-D-glutamic acid, L-lysine, D-alanine, and D-alanine residues linked to the NAM amino sugar (32). These pentapeptide units are cross-linked during bacterial cell wall growth by a bacterial enzyme known as penicillin binding protein (PBP)(32). PBP recognizes the terminal D-alanine-D-alanine dipeptide of peptidoglycan, cross-linking it to another pentapeptide through its catalytic adenyl-alanine endopeptidase activity. In so doing, PBP reinforces the growing peptidoglycan layer of the bacterial cell wall, strengthening it and conferring osmotic pressure resistance (32). Penicillin, the prototypical β-lactam antimicrobial, possesses a structure similar to the terminal D-alanine-D-alanine dipeptide of peptidoglycan, rendering it functionally indistinguishable to PBP (32). When penicillin is bound by the PBP enzyme, it inhibits cell wall crosslinking, eventually leading to cessation of cell wall synthesis and bacterial cell death from osmotic stress (32).

The Ambler classification is a method of categorizing β-lactamases based on their mechanism of active site catalysis and primary structures, placing enzymes into one of four classes A through D (33). Class A enzymes include narrow- and extended-spectrum β-lactamases with serine active site catalysis; Class B enzymes are broad-spectrum carbapenemase enzymes with zinc ion-mediated catalysis at the active site; Class C enzymes are cephalosporinase enzymes with serine active site catalysis; and Class D enzymes catalyze oxacillin hydrolysis through serine-mediated catalysis (33).

The AmpC β-lactamase of *S. marcescens* contains a serine at its catalytic active site and is an Ambler class C enzyme (34). AmpC enzymes are located in the periplasm between the inner and outer lipid bilayers of the Gram-negative cell membrane. AmpC enzymes are active on penicillins, cephamycins (e.g., cefoxitin, cefotetan), oxyimino cephalosporins (e.g., ceftriaxone, ceftazidime), and monobactams; in addition, they are generally stable to β-lactamase inhibitors such as clavulanic acid, tazobactam, and sulbactam (34, 35). There are three primary mechanisms resulting in AmpC-mediated resistance to the aforementioned antimicrobials: chromosomally encoded *ampC* genes causing inducible resistance phenotypes; plasmid-encoded *ampC* genes; or promoter/regulator mutations or deletions resulting in permanent de-repression of *ampC* (35, 36).

The induction of AmpC-mediated resistance is a complex phenomenon. AmpR, a member of the LysR transcriptional regulator family, typically regulates the basal expression of AmpC, reducing it to low levels when no environmental stressors are present (36, 37). Exposure to a β-lactam antimicrobial disrupts cell wall biosynthesis as previously described, resulting in the accumulation of NAG-1,6-anhydro-NAM moieties that are ultimately broken down into 1,6-anhydro-NAM oligopeptides (35). These 1,6-anhydro-NAM oligopeptides compete with uridine diphosphate (UDP)-NAM oligopeptides for a binding site on AmpR: when UDP-NAM is displaced by the cell wall degradation oligopeptide byproducts, it induces a conformational change in AmpR, resulting in increased *ampC* transcription and, thus, gene expression (35, 36). Other important proteins in AmpC regulation include the AmpG inner membrane permease, which transports the aforementioned oligopeptide byproducts into the cytosol (38). A cytoplasmic *N*-acetyl-muramyl-L-alanine amidase, AmpD, cleaves peptides off NAG-1,6-anhydro-NAM and 1,6-anhydro-NAM, allowing their recycling into the cell wall synthesis pathway while reducing their ability to bind AmpR (35, 39). However, once AmpD is saturated by degradation products in the cell and is no longer able to efficiently cleave them – as would occur with persistent β-lactam stimulus – the remaining oligopeptides in the cytosol are free to exert their effect binding AmpR, thus increasing

ampC gene expression (36). Under typical circumstances, when the β-lactam stimulus is removed, the expression of *ampC* returns to basal levels.

Permanent (also termed “stable”) de-repression of AmpC typically results from mutations or loss of function of one of the regulatory components of AmpC: *ampD, ampR,* or *ampG,* in decreasing order of frequency (39, 40). When permanent de-repression occurs, hyperproduction of AmpC persists even in the absence of a β-lactam trigger – and despite the fitness cost to the bacterium from the accumulation of cell wall degradation byproducts in the cytoplasm (36). AmpC hyperproduction is typical of *ampD* or *ampR* mutants (39, 40). Mutations in the AmpG inner membrane permease may also result in permanent de-repression of AmpC, but typically at a much lower level of gene expression than *ampD* or *ampR* mutants (35).

It is important to discuss how – and which – antimicrobials can act as inducers and substrates of the AmpC pathway to better understand the pathway itself. Being a substrate and/or an inducer are inherent properties of the antibiotic (Table 1). Induction refers to the property of an antimicrobial to act as a substrate for AmpC. A potent inducer is an antimicrobial which significantly upregulates AmpC expression, whereas a weak inducer still upregulates AmpC, but to an appreciably lesser extent. Substrate refers to the property of an antimicrobial to act as a molecule upon which the AmpC enzyme can itself act, facilitating the catalytic reaction that renders the antimicrobial ineffective. A poor substrate is one whose chemical structure renders it relatively stable to hydrolysis by the AmpC β-lactamase, while a good substrate is an antimicrobial that is readily hydrolyzed (35, 36, 41). Examples of potent inducers are ampicillin, the first generation cephalosporins, and cefoxitin; while imipenem is a potent inducer of AmpC, it forms a stable acyl-enzyme complex, thus protecting it from hydrolysis (42). The time to AmpC induction varies significantly based on the antimicrobial exposure: for example, imipenem induces AmpC expression within 180 minutes (43). An *in vitro* model in the Gram-negative bacillus *Escherichia coli* showed that promoter mutations occurred by day 2 of exposure to amoxicillin (44). When AmpC hyperproduction occurs, this results in an increase in the minimum inhibitory

concentration (MIC) for weak inducers, whereas the MICs of strong inducers shows relatively little change in de-repressed isolates (35).

Table 1. Antimicrobials Acting as Inducers and Substrates of the AmpC enzyme

	Potent inducers	**Weak inducers**
Good substrates	Ampicillin* Cefadroxil Cefazolin* Cefotetan* Cefoxitin* Cephalexin*	Aztreonam* Cefotaxime* Ceftazidime* Ceftriaxone* Piperacillin* Ticarcillin
Poor substrates	Imipenem*	Cefepime* Meropenem*

* Denotes antimicrobials with a clinical breakpoint (or for which another agent acts as a surrogate) for Enterobacterales including *S. marcescens* in the Clinical and Laboratory Standards Institute (CLSI) document M100, 35th edition (45).

Clinical data have demonstrated that some Enterobacterales organisms are at higher risk of permanent de-repression of the *ampC* operon than others. For example, *Citrobacter freundii, Enterobacter cloacae,* and *Klebsiella aerogenes* have mutation rates 50- to 150-fold higher than *S. marcescens,* thus conferring a much higher risk of clinically-relevant AmpC permanent de-repression in those organisms (46). AmpC expression rates are also approximately ten-fold lower for de-repressed *S. marcescens* compared to *C. freundii* or *E. cloacae* (36, 46). In a study by Choi and colleagues examining 42 patients with *Serratia marcescens* bloodstream infections which tested susceptible to cephalosporins, only 3 patients (7%) developed resistance to expanded-spectrum cephalosporins on therapy (47). Compared with similar data for *Enterobacter cloacae* showing between 19-67% of isolates demonstrated emergence of resistance on cephalosporin therapy (48-51), the relative risk of inducible AmpC-mediated resistance in *Serratia marcescens* is significantly lower. This creates two divergent categories amongst Enterobacterales with chromosomally encoded *ampC* genes: the “low risk” and “high risk” species, of which *Serratia marcescens* may be categorized among the former. Recent clinical guidance from the Infectious Diseases Society of America states that antimicrobial selection for the treatment of infection in *S. marcescens* should be selected based on antimicrobial susceptibility testing results, and suggests against the selection of ampicillin and first- and second-generation cephalosporins due to basal AmpC production rendering *S. marcescens* intrinsically resistant to these agents (52).

Antimicrobial susceptibility testing profiles differ for *Serratia marcescens* based on the level of AmpC β-lactamase expression. Antimicrobial susceptibility testing in the clinical laboratory is typically performed using a variety of different techniques, with the most common approach being broth microdilution. In broth microdilution, a known number of bacteria are incubated in a microtitre plate well containing a fixed volume of nutrient broth media with an antimicrobial solution in increasing geometric concentrations (53). After a standardized incubation period, each microtitre well is checked for bacterial growth. The lowest concentration of antimicrobial required to achieve bacterial killing of the isolate *in vitro* and is associated with a high likelihood of clinical therapeutic success is termed the minimum inhibitory concentration (MIC) (54). A clinical

breakpoint, set by a national or international laboratory standards organization, is then used to determine whether an isolate is susceptible to the antimicrobial based on the measured MIC value. Setting a clinical breakpoint requires pharmacokinetic and pharmacodynamic knowledge of the antimicrobial in question, data showing the wild-type normal distribution of MICs in the population, and clinical outcome data for patients in whom the antimicrobial has been used (54). Antimicrobial susceptibility testing with and without the presence of an inhibitory substance such as cloxacillin or boronic acid can be used to confirm AmpC production (36). Phenotypic testing, however, cannot provide information regarding whether constitutive AmpC expression is a result of chromosomal de-repression or a plasmid-borne AmpC operon (36). Antimicrobials for which *S. marcescens* has clinical breakpoints that have been determined by the Clinical and Laboratory Standards Institute (CLSI) are highlighted in Table 1.

A recent analysis by Aracil-Gisbert and colleagues defined three antimicrobial susceptibility phenotypes for *Serratia marcescens* complex based on AmpC production: in “basal-type” AmpC expression, which would correspond to the wild-type in absence of an external antimicrobial stimulus, amoxicillin-clavulanate, cefoxitin, and a third-generation cephalosporin test susceptible; in “inducible” AmpC expression, third-generation cephalosporins remain susceptible while amoxicillin-clavulanate and cefoxitin test resistant; lastly, de-repressed isolates test resistant to amoxicillin-clavulanate, cefoxitin, third-generation and fourth generation cephalosporins, with variable susceptibility to carbapenems (55). These susceptibility patterns differ from other ESBL-producing organisms because AmpC producers are consistently resistant *in vitro* to cephamycins such as cefoxitin and cefotetan, whereas ESBL producers are not.

1.4 Contrasting AmpC and extended-spectrum β-lactamase infection control practices

Gram-negative bacilli of the order Enterobacterales are common intestinal flora; in some patients, ESBL-producing Enterobacterales colonize the intestine (56). Clinical infections resulting from ESBL producers, as with AmpC producers, carry significant morbidity and mortality because

these infections require broader spectrum antimicrobials for clinical cure (57, 58). Many ESBL genetic elements are mobilized on plasmids and constitutively expressed, in contrast to AmpC operons, which tend to be chromosomally encoded and more commonly expressed in an inducible manner, as previously described. The transmission of ESBL-producing organisms in the hospital setting is well established in the literature both via direct or indirect contact between patients colonized with these organisms (59-61). Direct contact occurs when an individual that is not colonized comes into contact with fecal material from an ESBL-colonized individual, whereas indirect contact occurs via transfer of ESBL-producing Enterobacterales from colonized individuals to a susceptible host via fomites or healthcare workers (62).

Wastewater drainage systems have become increasingly recognized as an important environmental reservoir in the hospital environment for ESBL- and carbapenemase-producing Enterobacterales (CPE) (63-66). The use of sinks and showers in patient rooms as a means of disposal for human effluent contributes to the seeding of plumbing systems with bacterial biofilms: if human effluent contains ESBL producers or CPE, this serves as an important environmental reservoir for ESBL/CPE spread (67). As most ESBL and carbapenemase enzymes are mobilizable, this allows for the efficient spread of resistance among different constituent members of the biofilm. After the establishment of a bacterial biofilm, subsequent use of the sink or shower can result in the generation of droplets containing resistant bacteria, which then transfer onto humans using the sink or shower, thus allowing propagation of the bacterium throughout the hospital environment (68, 69).

A number of practices are implicated in the prevention of transmission of ESBL and CPE organisms, namely hand hygiene, environmental cleaning, and standardized administrative policies and procedures (62). The universal use of additional precautions such as the use of gowns and gloves – termed “contact precautions” – for all contact with the patient environment, particularly for ESBL producers, remains controversial. In a systematic review by Hagiya and Otsuka, they found 18 studies looking at the impact of discontinuing contact precautions on ESBL producers (70). While this review lacked a sufficient quantity of studies to extend their findings to

pediatric units, geriatric units, or ICUs, it showed that discontinuing contact precautions for ESBL-producing Enterobacterales had a minimal effect on the incidence of ESBL acquisition (70).

The use of contact precautions for patients colonized with CPE is somewhat less controversial, in part because evidence shows these controls decrease colonization rates and in part based on the precautionary principle, since treatment options for CPE are limited and mortality rates with invasive infection have been reported to be as high as 56.7% in a study performed in India (71, 72). In the province of Ontario, Canada, the Provincial Infectious Diseases Advisory Committee on Infection Prevention and Control (PIDAC-IPC) provides evidence-based advice to Public Health Ontario, a Crown corporation involved in the provision of healthcare at a provincial level. The PIDAC-IPC recommendations for the use of precautions for ESBL producers and CPE reflect these data, advising that contact precautions *may* be initiated for ESBL producers but *should* be initiated when CPE are isolated (62).

In contrast to ESBL producers and CPE, the role – if any – of contact precautions in AmpC producing Enterobacterales is incompletely defined. One study in Lebanon assessed the use of contact precautions to control the spread of plasmid-borne AmpC Enterobacterales and showed that the introduction of contact precautions significantly reduced the transmission of plasmid-mediated AmpC resistance (73). Chromosomal-mediated AmpC resistance is not conventionally thought to transmit readily between patients because it is not located on a mobilizable genetic element. As a result, while the use of routine practices in the hierarchy of controls is encouraged, such as environmental cleaning, hand hygiene, and standardized administrative policies and procedures for AmpC producers, guidance stops short of recommending routine gown and glove use for all contact with the patient and the patient environment (62). Until recently, however, little attention has been paid in the literature to systematically studying whether significant transmission of AmpC-producing Enterobacterales exists in the hospital environment.

1.5 Studying bacterial transmission in the hospital setting

If multiple cases of an organism are identified within a discrete geographical unit (e.g., a hospital ward, a community, etc.) over a period of time above an established baseline rate of cases, the possibility of an outbreak may be raised. In order to assess whether an outbreak is occurring, it is necessary to establish a case definition (i.e., relating to period, subject, and setting of interest) which can then be used to perform descriptive epidemiology by identifying and confirming cases that meet the definition to create a line list, as well as gathering demographic, risk factor, and clinical data (74). The time of onset of illness or disease may then be used to graphically summarize case data in an epidemic curve, which in turn provides information regarding the possible exposure, incubation period, and transmission mode of the infection (74). In the hospital context, dot density plots can also be used to plot infections in physical space – such as a hospital ward – to help understand the evolving infectious outbreak. Analytical epidemiology is then applied by generating and testing hypothesis about the exposures that lead to disease, typically using cohort or case-control designs to answer research questions (74).

When an outbreak is suspected within the hospital environment, a number of technologies may be used to help determine the genetic relatedness of organisms isolated from individual patients being investigated for colonization or infection with the organism of interest. The reference standard method is pulsed-field gel electrophoresis (PFGE), which creates deoxyribonucleic acid (DNA) “fingerprints” by harnessing the ability of restriction enzymes to cleave genomic DNA at specific sites, using an alternating electrical field pulsed through an agarose gel to separate out DNA fragments into specific banding patterns that can be visualized using fluorescent dyes (75). In recent years, the use of nucleic acid sequence-based techniques have largely supplanted PFGE as the methodologies of choice, since PFGE is a low-resolution method that is labour-intensive and has a prolonged turnaround time relative to newer techniques. One such technique is multi-locus sequence typing (MLST), which uses the genome sequences of specific “housekeeping” genes and the allelic differences between organisms to establish

sequence types, which are then compared between patient isolates (75). Finally, whole genome sequencing (WGS) has been recognized as an important tool in outbreaks, particularly as its cost per genome has significantly decreased over the past decade, rendering it a viable alternative to PFGE and MLST. WGS has the added advantage of being the highest resolution method to discriminate between isolates, allowing comparison of single nucleotide variants (SNVs) across the entire genome.

A study that systematically approaches invasive infections with *Serratia marcescens* could contribute to the evidence in a meaningful way by analyzing the genetic relatedness of invasive infections using WGS. In the event that chromosomal AmpC organisms transmit between patients via direct or indirect contact, this would result in the identification of multiple patients whose isolates were found to be genetically related to each other, but which had not been uncovered by the current surveillance techniques employed by IPAC practitioners at the institutions in question. In the event that chromosomal AmpC organisms do not appreciably transmit between patients via direct or indirect contact, none – or a small number – of patients would be found to have genetically related bacterial isolates: this finding would add to the evidence base for the current use of routine practices without the use of gowns and gloves during interaction with the patient infected or colonized with *Serratia marcescens* and/or their hospital environment.

1.6 Research question

We conducted a multicentre cohort study to answer the following research question: in patients with clinical infections caused by *Serratia marcescens,* using whole genome sequencing and subsequent determination of genomic relatedness of bacterial isolates, can we identify whether transmission of *S. marcescens* occurs between patients?

Chapter 2. Study Methods

2.1 Methodological Considerations

After formulating hypotheses regarding a possible infectious disease outbreak, there are multiple study designs that may be used in order to perform analytical epidemiology and formally test said hypotheses (76). While these strategies were briefly outlined in Chapter 1, we will further discuss the rationale for the design chosen for our study here. In analytical epidemiology, the most common study designs for hypothesis testing are case-control and cohort designs. In a case-control design, a relevant health outcome is identified by the investigators, such as the presence of a disease state. Within the paradigm of the present study, this health outcome would correspond to the presence of *Serratia marcescens* recovered from the bacterial culture of a patient. Two groups of study participants are then identified: those who have the outcome, which are termed “cases”, and those who do not have the outcome, which are termed “controls” (76). The case definition used to assign patients must be sufficiently stringent and detailed with respect to the population, intervention (where applicable), setting, and time period of interest so as to ensure no incorrect categorization of participants occurs in the study, as this would introduce a classification bias (77). A case-control design then obtains data retrospectively on both cases and controls to determine whether a relevant exposure occurs more frequently in cases versus controls (76). Within the paradigm of the present study, this exposure could include cases having been exposed to a specific intervention (i.e., surgery performed in the same operating theatre) or setting (i.e., being admitted to the same hospital ward) that puts them at higher risk than the controls of achieving the outcome of interest.

While case-control studies are typically simple and efficient to perform, they come with a number of distinct disadvantages, in particular their inability to study disease incidence, their inherent risk of biased findings, and the possibility of confounders distorting the association between the exposure and outcome of interest (76). A confounder is a variable, either measured or unmeasured, which acts on both the exposure and the outcome and is causally related to neither entity (78). This stands in contrast to a mediator, which is a variable on the causal

pathway between exposure and outcome: *viz.*, the exposure causes the mediator, which then causes the outcome (78). Confounders are important to identify in case-control studies: without adequately controlling for a confounder, the researcher may come to spurious and incorrect conclusions in their study (76).

We decided that a case-control design for this study would be inappropriate for a number of reasons. The goal of the present study was to determine whether transmission of *S. marcescens* occurs outside of known outbreak settings, which renders a case-control design fruitless. Furthermore, establishing a case definition would be difficult and render the conduct of this study unlikely, as this would require *a priori* knowledge of groups of patients in which *S. marcescens* had transmitted, in order to be appropriately compared to those that did not. In such a study, the primary goal would be establishing risk factors for transmission, which would have required a design in which patients were regularly swabbed for colonization. One group of patients that could fulfil this criterion are those identified as part of a *Serratia marcescens* outbreak. As previously described, such outbreaks are known to occasionally occur in hospital settings. *Post hoc,* it is possible to confirm that this case-control design would have been unfeasible to complete during the study period 2022-2024, as no detected *Serratia* outbreaks occurred over the given period.

As such, we performed a hybrid cohort study design with both prospective and retrospective elements. In a retrospective cohort study, the outcome of interest has already occurred, so patient data is extracted and reviewed after the fact to determine whether there is a common exposure in the study population (79). In contrast, in a prospective study, the exposure is determined at study enrolment and patients are followed to determine if they achieve the outcome of interest (79). A prospective study typically has a lower risk of biased data collection, since researchers have greater control over how variables of interest are measured and collected (79). However, prospective studies require prospective patient enrolment, thus increasing the effort expended and – as a consequence – costs. Increased costs can lead to budgetary

constraints insofar as the number of patients one is able to enrol in the study in question. We therefore chose a retrospective cohort design to minimize costs and maximize patient recruitment.

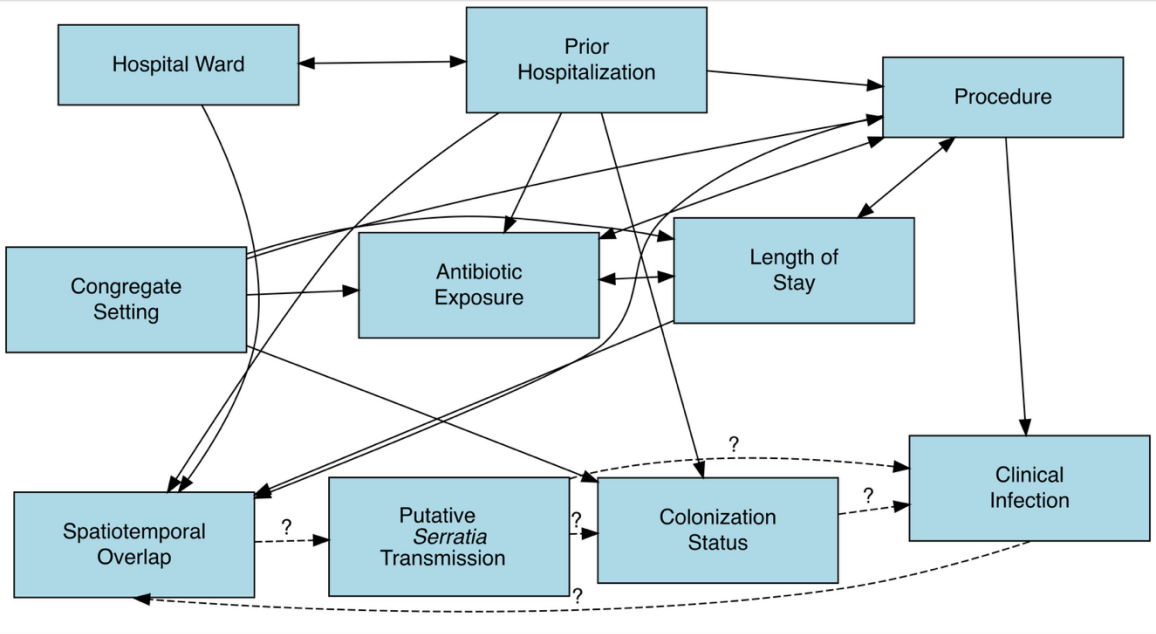
One component of data acquisition, however, was required to be prospective for practical purposes: specimen identification and banking. Clinical microbiological specimens may be classified as retrievable or irretrievable based on the specimen type, the method of collection, or the timing of collection (80). For example, a cerebrospinal fluid collected before antimicrobial administration in a suspected *Serratia marcescens* meningitis would be considered irretrievable because the specimen requires an invasive procedure to retrieve the specimen; repetition of the invasive procedure may expose the patient to undue harm; and specimens collected after effective antimicrobials have been administered may not isolate the organism of interest. In contrast, retrievable specimens are those which can be easily recollected or were taken from typically non-sterile sites (e.g., skin, nasopharyngeal swabs, sputum, urine, blood). Clinical microbiology laboratories operate with environmental constraints on specimen storage capabilities, leading to the institution of retention times for various specimen types based on their retrievability prior to their disposal in accordance with biosafety rules and regulations.

Since the present study was not limited to a specific specimen type, instead allowing any invasive infection to be included, the eligible specimen types included both retrievable and irretrievable specimens with varying retention times. If the patient specimens were identified retrospectively in the same manner as the patient data was extracted, it is likely that the majority of eligible patients would have had their *Serratia marcescens* isolates discarded, thus rendering them excluded from the study. Therefore, the only feasible way in which the study could proceed was by prospectively identifying patients with appropriate specimen types positive for *Serratia marcescens* growth in the laboratory, retaining these specimens until the end of the study period, and then retrospectively collecting patient data from their charts as well as analyzing the whole genome sequences of the retained isolates.

As previously described, confounders are important variables to control for in non-randomized study designs. The non-randomized nature of the present study allows for

imbalances in prognostic factors between exposed and unexposed patients (81). There are a number of possible confounders identified in this study, including hospital ward, patient length of stay, congregate living, prior hospitalization, and antimicrobial exposure. A directed acyclic graph is a method that can be used to easily visualize possible confounders when estimating causal effects in a pathway (82). A directed acyclic graph of the possible confounders and how they may relate to putative *Serratia marcescens* transmission is found in Figure 1. The geographical unit of the hospital ward is a possible confounder because it is a unit of natural geographical clustering that may result in patients sharing environmental exposures that affect both the presence of *Serratia marcescens* in the patient and the risk of developing invasive *Serratia marcescens* infection. Similarly, length of stay is a form of time-modified confounding, as it affects the possible exposure time to the hospital environment, increasing the likelihood of both exposure and outcome (83). Congregate living, such as occurs in long-term care facilities or prisons, can affect the risk of colonization with a pathogen and introduces the possibility of shared environmental exposures that can confound the association between exposure and outcome.

Figure 1. Directed Acyclic Graph Illustrating the Possible Relationships between Confounders and Putative *Serratia marcescens* Transmission



Solid lines indicate likely relationships, either unidirectional or bidirectional, between exposures or risk factors. Dashed lines with question marks indicate possible unilateral relationships on a putative causal pathway from spatiotemporal overlap to clinical *Serratia marcescens* infection, investigated by the present study.

Confounding can be addressed in multiple ways in observational studies, at the design or the analysis stage (81). In the design stage, restriction or matching can be used as methodological strategies (79, 84). In applying restriction, the researcher uses specific inclusion criteria to limit the characteristics of included patients to minimize the variability of factors that may confound the association between exposure and outcome. In our study, we felt that this approach was feasible, and we restricted isolates to the first *Serratia marcescens* identified on culture during a given hospitalization. For example, if a patient was hospitalized for a long period of time such as one year and had two *Serratia marcescens* invasive infections over that time period, it is entirely possible that changes in the patient's hospital ward or cumulative antimicrobial exposure may confound the possible association being studied. As a disadvantage, however, restriction decreases the likelihood that the findings are generalizable. Restriction to the first isolate reduces the risk of confounding by avoiding within-patient correlations, thus ensuring only independent observations are measured and analyzed, as well as reducing time-varying confounding. Matching was not felt to be feasible for the present study because there are many possible confounders identified, and this technique can only be done for a limited number of variables.

At the analysis stage, statistical techniques including stratification, regression, or propensity score matching may be used to help control the influence confounding has on study results (81, 84). Of these approaches, stratification was the most likely to be operationally feasible. However, with a number of possible confounders and a low number of included patients, introducing stratification would decrease statistical power and lead to an increased risk of biased results. Multivariable regression modelling was felt to be a suboptimal approach because the number of putative transmission events was expected to be rare and having fewer than ten events per variable included in a model leads to biased parameter estimates (85). Finally, propensity score matching was also felt to be inappropriate to address the study question. Propensity score matching is an attempt to balance observed covariates among exposed and unexposed groups by computing a score, typically using multivariable logistic regression

modelling, by which study participants may be adjusted (86). As the covariates which we would be adjusting for are not necessarily known to be true or potential confounders – since the underlying research question is to assess *whether* transmission is occurring at all – this would be difficult to implement in practice.

Yet another important methodological aspect to discuss with respect to the present study relates to how transmissions were defined. The terminology “spatiotemporal dynamics” introduced in the previous chapter contains two discrete elements: transmission across space and across time. The discrete unit of space is much more easily and intuitively defined in this expression by the geographical “hospital unit” in which patients are housed and share communal exposures such as equipment, washrooms, healthcare staff, etc. This stands in contrast to the definition of transmission across time: in examining a heretofore unstudied hypothesis, one is required to make an assumption regarding the period of time over which transmission of an organism between patients is likely. Since there is no evidence upon which this assumption is grounded, this methodological choice by the researcher may introduce bias into the findings. One way in which the researcher can further explore whether this methodological choice influenced the overall study conclusions is through the use of sensitivity analysis. A sensitivity analysis is one which seeks to determine whether research findings are robust to changes in the analysis methods or assumptions made – *viz.,* whether altering assumptions changes the underlying study conclusions (87). If the overall study conclusions are congruent between the primary and sensitivity analyses, the researcher may be reassured that their methodological choices were unlikely to influence the underlying results and, thus, the conclusions drawn from them (87).

As discussed in chapter 1.5, there are multiple technologies that can be used to study microbial isolates that are suspected of being transmitted in hospital. The typical methods used contemporarily include Sanger sequencing, MLST, PFGE, and whole genome sequencing. Sanger sequencing uses a mixture of normal deoxynucleotides and chain-terminating dideoxynucleotides to sequence a DNA strand complimentary to the template strand (88). The DNA sequence being interrogated is separated into four separate reactions, containing all

deoxynucleotides but only one of each dideoxynucleotide. Over multiple rounds of DNA template extension, this creates fragments of varying sizes because the chain-terminating dideoxynucleotide is added at varying places during extension. Resulting DNA fragments are separated by capillary gel electrophoresis, and a fluorophore moiety attached to the dideoxynucleotides is excited by a light source to determine which nucleotide represents a given position on the DNA strand. While Sanger sequencing has high base accuracy ≥99.0%, long read lengths, and is cost effective for small numbers of targets or samples (89), it is labour intensive and low throughput, rendering it impractical for a large study of over 100 isolates such as ours. MLST provides some distinct advantages over Sanger sequencing, including the ability to support phylogenetic analysis with its moderate resolution using housekeeping genes and the ability to determine sequence types for included organisms. However, a recent analysis of *Salmonella enterica* serovar Schwarzengrund isolates showed MLST has poor discriminatory power in clonal outbreaks or when clonally dominant endemic strains are present (90). Another distinct disadvantage of MLST is that it analyzes highly conserved genes and does not offer any information regarding antimicrobial resistance determinants, precluding it from being useful in our study. PFGE is the historical reference standard for outbreak investigations, which examines genome-wide restriction enzyme polymorphisms with well-established interpretation criteria first published by Tenover and colleagues in 1995 (91). Distinct disadvantages of PFGE include its laborious nature, high cost, and limited resolution for closely-related isolates (90). Whole genome sequencing offers unparalleled resolution at a single nucleotide level, also allowing the detection of resistance determinants that was necessary for this study's primary outcome.

Multiple platforms and sequencing chemistries exist with which to perform whole genome sequencing. The choice of a sequencing platform or chemistry may be pragmatic in nature – it can often be partially based on the infrastructure already available in a laboratory, given the high start-up costs to bring a platform online. However, in general, whole genome sequencing may be split into short- and long-read sequencing, based on the read lengths. Whereas short-read platforms typically sequence lengths of up to 1000bp, long-read platforms can sequence from

10kb to 1Mb lengths of DNA (88). Long-read sequencing has higher per-read error rates than a short-read platform such as the Illumina NextSeq 2000, with typically more complex library preparation and, thus, costs (88, 92). However, long-read sequencing is also thought to be better able to ensure accurate plasmid assembly because they result in higher overall assembly scores, likely because most assembled plasmids from long-read platforms contain single contigs (93). A long-read sequencing platform was considered for this study, because of the ability to better resolve plasmids in the event that resistance determinants were primarily plasmid-borne; however, short-read sequencing was ultimately chosen for its ease-of-use, lower cost, lower error rates, and the availability of ubiquitous and well-established bioinformatics pipelines for quality control and sequence alignment.

An important methodological assumption that was made during the study was with respect to spatiotemporal dynamics. To be considered a possible transmission event, isolates from different patients had to be isolates within the same hospital ward and within one month of each other to be considered possibly linked. Furthermore, the nucleotide sequences needed to be sufficiently similar when compared between patients. To our knowledge, there is no scientific evidence base with which to support this spatiotemporal assumption: as a result, while making this assumption was necessary to perform the research at hand, it may have the unintended consequence of biasing our conclusions if a sensitivity analysis was not undertaken. As outlined later in this chapter, we decided to perform a sensitivity analysis using two- and three-month time cutoffs in order to determine whether our definition of transmission across time may have influenced the study results. If the number of possible transmissions between patients is similar amongst the analyses where the time cut point was set at one, two, and three months, this provides confidence that this methodological choice did not unduly influence the study's ultimate conclusion.

2.2 Study Design

We conducted a cohort study whereby patient specimens with growth of *Serratia marcescens* in bacterial culture were prospectively identified and stored until the cut-off dates at the end of each study period. Observational patient-level data was then extracted retrospectively via chart review. This study was reported according to the Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) guidelines (94).

2.3 Patient Population

The study population consisted of adults and children, with no age restriction, who had clinical specimens submitted to the microbiology laboratory at participating hospitals for the diagnosis of active infection which were ultimately identified as *Serratia marcescens* using conventional microbiological techniques as defined in laboratory standard operating procedures (further elaborated in Section 2.7 to follow).

2.4 Study Setting

We conducted the study at three adult tertiary care sites, two secondary care sites, and one pediatric tertiary care site in Hamilton, Ontario, Canada. The participating hospitals included: Hamilton General Hospital, Juravinski Hospital, the McMaster University Medical Centre, St. Joseph's Healthcare Hamilton, St. Peter's Hospital, and West Lincoln Memorial Hospital. These hospitals represent a total of 2800 beds in the Greater Toronto-Hamilton geographical area in southwestern Ontario. Specimens were prospectively identified over two time periods: January to December 2022, and January to June 2024. These time periods were chosen to ensure that the study spanned multiple years in the event that a possible outbreak of invasive *Serratia marcescens* infections was identified, and in order to account for the possibility of time-modifying covariates as discussed previously in Section 2.1. Each time period represented consecutive patients.

2.5 Study Inclusion and Exclusion Criteria

Specimens were included if they were submitted to the microbiology laboratory in order to diagnose active infection in a patient over the study period. Specimens were included from any anatomical site, from both hospital inpatients and outpatients at the eligible hospital sites. In order to minimize possible biases in the study design phase (see Section 2.1 for a methodological explanation), we restricted included isolates to the first isolate from a given hospital admission. If a patient was admitted twice to the same or different hospitals over the study period and met all other inclusion criteria, only the first isolate from each admission would be included.

Specimens were excluded if they were submitted for infection prevention and control surveillance purposes to diagnose rectal colonization of *Serratia marcescens* or for point-prevalence surveys, as would be typical for rectal swabs.

2.6 Study Outcome

The study's primary outcome was to identify whether transmission of *S. marcescens* occurs between patients with clinical infections caused by *Serratia marcescens,* using whole genome sequencing and subsequent determination of genomic relatedness of bacterial isolates. The study was insufficiently powered to examine clinical outcomes such as associations of AmpC resistance mechanisms with clinical outcomes, or associations between carbapenem use and clinical outcomes.

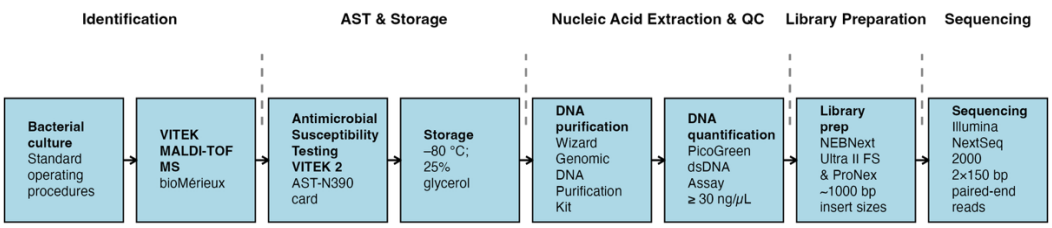
2.7 Laboratory Methods

The laboratory methods and workflow are summarized in Figure 2. Upon receipt in the microbiology laboratory, specimens were accessioned, processed, inoculated onto solid agar or

liquid growth media, and incubated as per standard operating procedures inherent to the specimen type submitted by the requesting physician. Once bacterial colonies were identified on nutrient media after incubation under appropriate conditions, multiple representative colonies with the same morphology – i.e., size, colour, shape, and margin – were selected from a pure culture using a sterile wooden applicator and applied to a target plate. The colonies are mixed with an organic, proprietary compound known as matrix, which co-crystallizes the bacterial sample within it upon drying. Isolates then underwent confirmatory identification using MALDI-TOF MS on a VITEK MS® instrument (bioMérieux, Marcy-l'Étoile, France) using the VITEK MS® V.3.2 database to assign species identifications.

The MALDI-TOF MS functions by using a fixed wavelength laser beam to ionize the sample-matrix complex, leading to desorption of the analyte from the target plate and the generation of cationic protein fragments (95). The protonated ions are then accelerated through a vacuum flight tube at a fixed potential towards a detector and the ions separate in a predictable way based on their mass-to-charge (*m/z*) ratio (95). A time-of-flight analyzer determines an ion's *m/z* ratio based on the time it took to travel from the target plate to the detector (95). Once all protonated molecules have reached the detector, it is able to generate a peptide mass fingerprint which is then queried against a database of microbial spectra of known isolates (95). If the database returns a match of sufficiently high similarity between query sequence and the database, the MALDI-TOF MS assigns a genus and/or species name to the analyte alongside a quality score.

Figure 2. Laboratory Methods Workflow Diagram



AST = antimicrobial susceptibility testing; MALDI-TOF MS = matrix-assisted laser desorption ionization time-of-flight mass spectrometry; QC = quality control

We used the VITEK® 2 system (bioMérieux, Marcy-l'Étoile, France) to perform antimicrobial susceptibility testing (AST) on the isolates after identification. The VITEK® 2 system is an automated system using closed-system incubation and is based off a miniaturized broth microdilution platform (96). Isolates are tested using cards that contain increasing concentrations of antibiotics, monitoring growth throughout incubation using broth turbidity to measure bacterial growth by monitoring percent change in raw transmittance units (96). Isolate MIC data is determined algorithmically by comparing the growth of the test isolate against that of an organism with a known MIC (96). We used the VITEK® 2 system's AST-N390 card specific to Gram-negative pathogens to test all study isolates. Interpretative breakpoints were assigned for the MICs in accordance with the CLSI M100 guideline used in the clinical laboratory at the time of specimen analysis for clinical care (45). After clinical testing was completed for the purposes of patient care, samples were prospectively identified by study staff and frozen at -80°C in a 25% glycerol stock solution.

We used the Wizard Genomic DNA Purification Kit (Promega, Madison, United States of America) to purify genomic DNA according to the manufacturer's product insert. We subsequently quantified the genomic DNA to ensure that a minimum concentration of 30 ng/μL was present for downstream applications. We quantified the DNA using the PicoGreen dsDNA Assay (ThermoFisher Scientific, Waltham, United States of America). We then prepared individually barcoded short-read libraries using the NEBNext Ultra II FS DNA Library Prep Kit (New England Biolabs, Ipswich, United States of America) and enriched for insert sizes of approximately 1000 base-pairs (bp) using the ProNex Size-Selective Purification System (Promega, Madison, United States of America). We used the Illumina NextSeq 2000 to sequence the genomic DNA library, generating 2x150bp paired-end reads from an average of 3.2 million clusters (range: 0.6-5.6 million clusters).

Illumina platforms are an example of sequencing-by-synthesis technologies. Small fragments of genomic DNA that have had molecular adaptors ligated onto the 5' and 3' ends of the DNA strands (97). Regularly spaced DNA sequences complementary to the ligated genomic

DNA adaptors are immobilized on a glass slide, allowing genomic DNA to bind to the flow cell (97). This arrangement allows DNA synthesis to occur in many millions of clusters over a flow cell in parallel. A proprietary labelled deoxynucleoside triphosphate (dNTP) molecule is added to the growing nucleic acid chain in a synthetic reaction to create a complementary DNA sequence to that which is immobilized on the flow cell (97). The dNTP molecules serve as chain terminators for DNA polymerization, allowing a fluorophore label to emit fluorescence that can be detected after each base incorporation (97). Base calling – the process by which the incorporated nucleotide is determined to represent adenine, guanine, thymine, or cytosine – occurs based on the relative fluorescence, after which the fluorescent moiety is cleaved, effectively "unblocking" the complementary DNA molecule for a further round of DNA synthesis via base incorporation and calling (97). This recurs until the entire DNA strand is sequenced. The resulting genomic DNA sequences can then be used for downstream bioinformatics analyses.

2.8 Bioinformatics

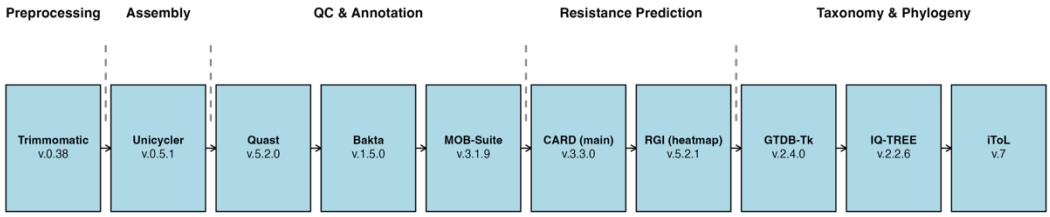
The bioinformatics workflow is summarized in Figure 3. We used Trimmomatic v.0.38 (98) to filter and trim adapter sequences from the raw Illumina platform sequencing reads; we specified a minimum read-length of 36bp and a quality score cut-off of 3. We then used Unicycler v.0.5.1 (99) to assemble the filtered reads, and used Quast v.5.2.0 (100) to evaluate genome assembly quality and calculate whole genome guanine-cytosine (GC) content. Quast aligns whole genome assemblies to a reference genome, allowing contig size and other functional genomic elements to be evaluated. A contig is a set of DNA sequences that have been aligned such that there is sufficient overlap to assemble a contiguous stretch of DNA without any gaps (101). A typical representation of the genome assembly quality is the N50 statistic, which is a weighted median of the contig lengths (102). The N50 is computed by ordering contigs by length, then summing the contig length until the cumulative length reaches 50% of the total assembled nucleotide sequence length: the shortest contig length amongst those summed is the N50 (102).

A high N50 is suggestive of fewer gaps in the assembled genome, which is desirable because it reflects a more “complete” assembly. The ”correctness” of the genome assembly is also assessed by Quast via detection of misassemblies or errors in the consensus alignment (100). We then used Bakta v.1.5.0 (103) to annotate contigs in order to identify relevant genes. To predict plasmids, we used the Solu cloud-based platform v.1.0.251 (104) to run MOB-Suite v.3.1.9 (105). MOB-Suite is a software tool that allows the reconstruction and typing of plasmids from whole genome sequencing data and allows visualization of putative plasmids (105).

After genes have been annotated in the assembled contigs, the resistome was predicted by querying the Comprehensive Antibiotic Resistance Database (CARD, v.3.3.0)’s Resistance Gene Identifier (RGI) tool (106, 107). We allowed only “Perfect” and “Strict” hits in the database, where a perfect match is one that is 100% identical to the wild-type reference protein sequence along its entire length, while a strict match has a Basic Local Alignment Search Tool Protein (BLASTP) expectation value used to infer homology which is above the database cut-off but which may or may not contain a relevant mutation (106). We used the -include nudge argument in RGI main so that any “Loose” hits in the database with ≥95% sequence identity are “nudged” into the “Strict” hit category (106). RGI v.5.2.1 was used to summarize the output using rgi heatmap (106).

Finally, we focussed on taxonomic assignment and inference of phylogeny. We determined an isolate’s taxonomic assignment using the Genome Database Taxonomy (GTDB)’s GTDB-Tk software v. 2.4.0 (108), which assigns each genome to a domain-specific reference tree and ascribes taxonomy based on the genome’s relative evolutionary divergence and average nucleotide identity. We then used IQ-TREE v.2.2.6 (109) to infer a maximum-likelihood phylogenetic tree, visualizing said tree using iToL v.7 (110).

Figure 3. Bioinformatics Methods Workflow Diagram



CARD = Comprehensive Antibiotic Resistance Database; GTDB-Tk = Genome Taxonomy Database Toolkit; iToL = Interactive Tree of Life; QC = quality control; RGI = Resistance Gene Identifier.

2.9 Data Collection

We collected patient-level data at two times during the study period, at the end of each prospective sample identification period – *i.e.,* in December 2022 and July 2024. We performed a retrospective chart review to gather patient demographic and microbial isolate data. We piloted the chart review process on the first five included patients to ensure all relevant characteristics were being collected and also in order to minimize the risk of a high proportion of missing or incomplete data undermining the study's validity.

We collected the following patient-level data from the medical chart: patient age; patient sex; hospital; ward; whether the patient was a resident of a congregate setting; the number of hospitalizations in the past year; the principal admission diagnosis; the length of stay at time of specimen collection; whether the patient was receiving antibiotics at specimen collection; which antibiotics were being received at specimen collection (if any); which antibiotics were used to treat the infection; treatment duration; patient outcome; and whether hospital discharge summary included infection as a cause of death, if relevant. Many of the aforementioned variables were collected as they are both epidemiologically relevant and because they may be possible confounders or mediators of the association between exposure and outcome, as discussed in greater detail in Section 2.1 previously.

In terms of microbiologically relevant data from the medical chart, we extracted the following variables: specimen site, specimen collection date, species identified by MALDI TOF-MS, percent certainty of MALDI TOF-MS identification >95%, minimum inhibitory concentrations (MIC) and interpretative breakpoints for all antimicrobials tested.

2.10 Epidemiological and Sensitivity Analyses

We combined whole genome sequences and conventional epidemiological approaches to determine if *Serratia marcescens* complex isolates were spatiotemporally linked with each other. We considered there to be a possible epidemiological link between patients if a pair of patients' *Serratia marcescens* complex isolates visually clustered together on the maximum-likelihood phylogenetic tree and the following criteria were met: patients overlapped temporally by one month, using the specimen collection date as an indicator of this, and; patients overlapped spatially (i.e., were admitted to the same hospital ward), and/or; patients shared an exposure to the same procedure (e.g., bronchoscopy, colonoscopy, lumbar puncture).

If patients fulfilled the above criteria, we then calculated pairwise average nucleotide identity (ANI) using OrthoANIu (111). We considered an ANI value ≥99.0% to be significant. While an ANI >95% is typically sufficient for species-level identification, recent evidence suggests an ANI value ≥99.0% is required for sufficiently granular intra-species strain determination and differentiation (112).

As discussed in Section 2.1 previously, the decision to define temporal overlap as one month in length was an arbitrary one, since there is no literature to inform the choice of a threshold. As a result, we performed a sensitivity analysis by re-evaluating the primary outcome using temporal thresholds of two and three months to determine whether the arbitrary choice of threshold affected our primary analysis.

2.11 Sample Size

Given the *a priori* assumption that chromosomal ampC-producing Enterobacterales are not hospital-transmissible pathogens, an event rate of 0 is presumed. For the purposes of sample size calculation, as an event rate of 0 is not mathematically permissible, a correction of 0.5% is applied. Assuming 5% of isolates identified are found to be spatiotemporally linked and whole

genome sequencing analysis shows possible relatedness, with an α of 0.05 and a β of 0.2, we determined that 51 samples were required to be adequately powered to determine whether unrecognized transmission was occurring between patients in hospital. With a projected attrition rate of 10% for contaminated or improperly stored samples, we calculated that 60 included samples were required for statistical validity.

2.12 Statistical Analysis

We used descriptive statistics to summarize the characteristics of patients and their *Serratia marcescens* complex isolates. To calculate relevant statistics, we used command-line tools for genome annotation and R, version 4.4.2 (R Foundation for Statistical Computing, Vienna). Figures and data visualizations for this study were created using the packages ggplot2 (113), dplyr (114), diagrammeR (115), and ComplexHeatmap (116).

2.13 Data Transparency

Genomic DNA sequences and their linked antimicrobial MIC data are available under the NCBI BioProject #PRJNA1232785 (https://www.ncbi.nlm.nih.gov/bioproject/).

2.14 Ethics Approval

This study received ethical approval from the Hamilton Integrated Research Ethics Board under project number 13873. Individual patient consent was not required for this study, as it utilized residual specimen from testing ordered by attending physicians for routine clinical care that would otherwise have been discarded. The study underwent one protocol amendment to narrow the scope of organisms from *Serratia, Enterobacter,* and *Citrobacter* species to *Serratia*

marcescens only, and to modify the whole genome sequencing methodology and analytical techniques to reflect contemporary *in silico* tools.

2.15 Funding

This thesis work was supported by operational funding for reagents, laboratory personnel, and laboratory testing from the Department of Pathology and Molecular Medicine at McMaster University, Award #PMM 2021-01. Salary grants were received from the Government of Ontario's Ministry of Health for a Clinician-Investigator Program Fellowship, as well as the Master of Science Studentship Award from the Research Institute of St. Joseph's Healthcare Hamilton.

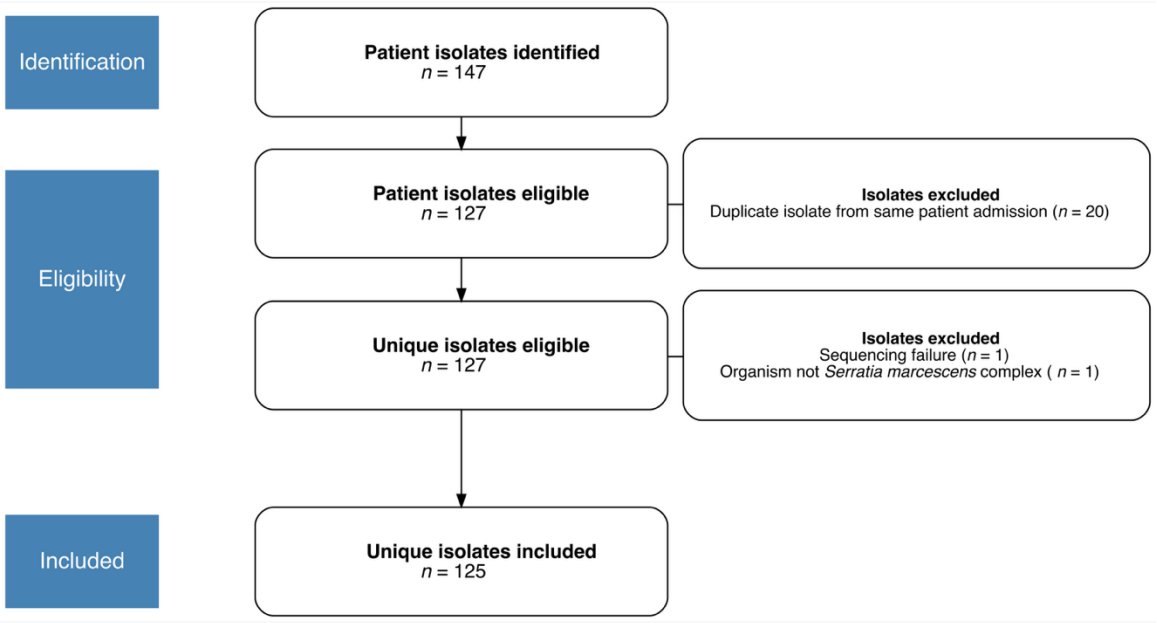
The funders did not have any role in the design, data acquisition, or data analysis performed in this work.

Chapter 3. Results

3.1 Patient Characteristics

During the periods January to December 2022 and January to June 2024, we identified 147 eligible patients with invasive *Serratia marcescens* infections from participating hospital sites. During the study period, no outbreaks of *Serratia marcescens* complex were detected by standard of care IPAC surveillance or were otherwise known to have occurred. We excluded 22 eligible patient isolates from the study: 20 identified isolates were subsequent – i.e., duplicate – isolates from an included patient's same hospital admission; one isolate failed whole genome sequencing; and one isolate was in fact *Aeromonas caviae.* The *A. caviae* isolate was erroneously frozen during specimen banking from a patient with a mixed *Serratia marcescens* and *A. caviae* infection. Figure 4 illustrates the study flow with respect to participants.

Figure 4. Participant Flow Diagram



Patient demographic and clinical data are summarized in Table 2. The majority of included patients were male (n=88/125, 70.4%), with the median age among study patients being 69.0 years (interquartile range (IQR): 57.0-76.0). Figure 5 shows the distribution of cases of *Serratia marcescens* complex infections, categorized by sex and age in deciles. Peak incidence among both males and females occurred between age 70-79. In males, there was a higher incidence of cases in the age deciles 0-9 and 10-19 years, when compared to females. Eleven of 125 (8.9%) included patients lived in a congregate setting.

Roughly one-third of isolates (32.0%, n=40/125) likely represented microbial colonization at the time of specimen collection, as patients did not receive documented antimicrobial treatment for their *Serratia marcescens* complex culture result. While only 4.0% of patients (n = 5/125) received empiric carbapenem therapy, 28.8% (n = 36/125) received a carbapenem as definitive therapy for their *Serratia marcescens* infection. Of those patients who received antimicrobials, the median duration of treatment for all indications was 7.0 days (IQR: 7.0-14.0). Seventeen of 125 patients (13.6%) died during the study period, with the majority of those deaths (n = 11/17, 64.7%) attributable to *Serratia marcescens* infection upon review of hospital discharge summary documentation.

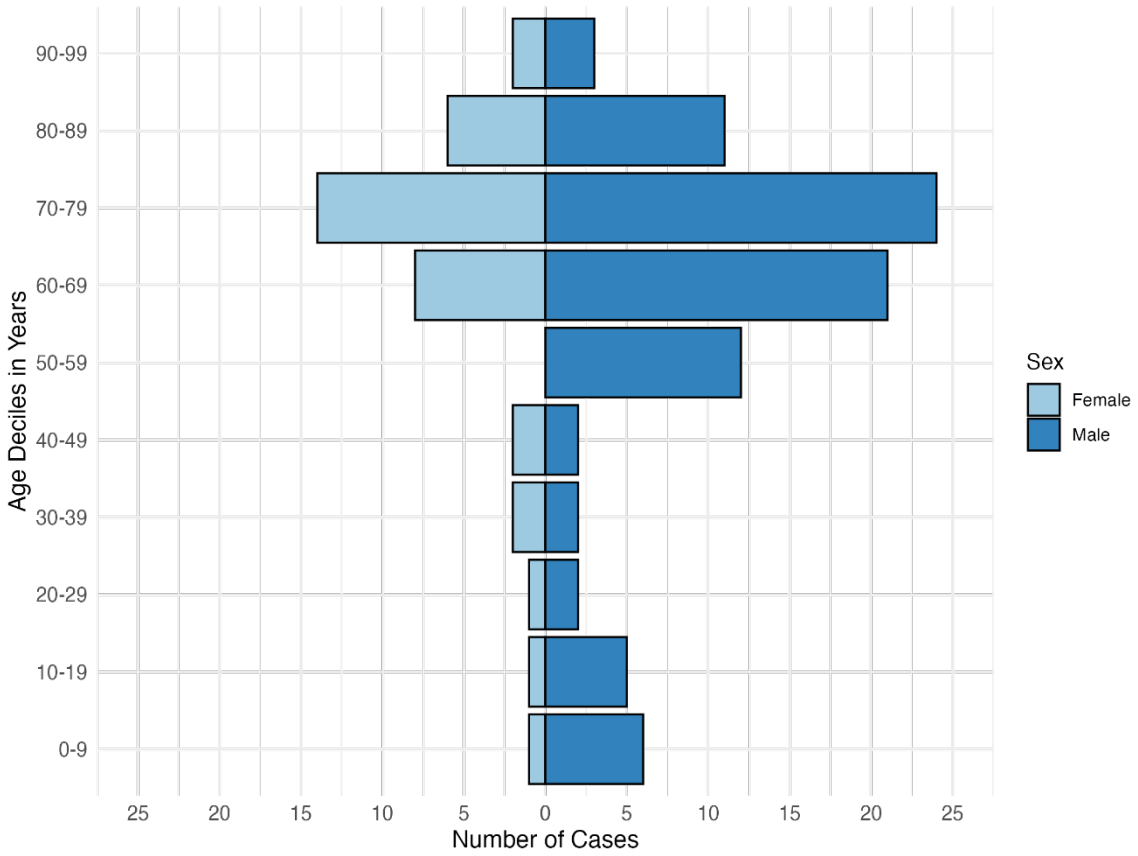
Stratification of patient demographics by study year revealed key differences in these subpopulations with respect to the use of carbapenem antimicrobials and the proportion of patients who did not receive treatment for the organisms isolated. Interestingly, the proportion of patients receiving empiric therapy with a carbapenem was higher in the 2022 cohort compared to the 2024 cohort (3.2% vs. 1.8%, respectively), but among those who did receive treatment, definitive therapy with a carbapenem was less common in 2022 than in 2024 (23.9% vs. 35.2%, respectively).

Table 2. Demographics of Included Patients with *Serratia marcescens* complex Infections

Patient characteristic	All included patients (*n*=125), *n* (%)	Patients identified in 2022 (*n*=71), *n* (%)	Patients identified in 2024 (*n*=54), *n* (%)
Median age in years, *n* (IQR)	69.0 (57.0-76.0)	67.0 (59.5-76.0)	69.5 (51.5-78.5)
Female, *n* (%)	37 (29.6)	21 (29.6)	16 (29.6)
Congregate setting, *n* (%)	11 (8.9)	5 (7.0)	6 (11.1)
Median hospitalizations in past year, *n* (IQR)	0 (0-1.0)	1.0 (0-1.0)	0 (0-1.0)
Hospital length of stay at specimen collection, median *n* days (IQR)	1.0 (0-10.0)	1.0 (0-7.0)	2.0 (1.0-20.00)
Receipt of antimicrobials prior to specimen collection, *n* (%)	62 (49.6)	36 (50.7)	26 (48.1)
Receipt of carbapenem as empiric treatment, *n* (%)	5 (4.0)	4 (3.2)	1 (1.8)
Receipt of carbapenem as definitive treatment, *n* (%)	36 (28.8)	17 (23.9)	19 (35.2)
Median treatment duration, *n* days (IQR)	7.0 (7.0-14.0)	7.0 (7.0-14.0)	7.0 (7.0-10.7)
Patients not treated for positive culture result, *n* (%)	40 (32.0)	16 (22.5)	24 (44.4)

Patient deceased, *n* (%)	17 (13.6)	12 (16.9)	5 (9.3)
Of deceased, infection as attributable cause of death, *n* (%)	11 (64.7)	6 (50.0)	5 (100.0)

Figure 5. Distribution of *Serratia marcescens* Cases by Patient Age and Sex



Patient cases of invasive *Serratia marcescens* infection at six Hamilton area hospitals identified from January to December 2022 and January to June 2024 by routine bacteriological culture and confirmed with MALDI-TOF MS, categorized according to patient age and sex.

3.2 Specimen and Microbial Characteristics

Serratia marcescens was isolated from multiple anatomical sites in included patients, with an itemized accounting of the specimen sources found in Table 3. The three most common specimen sources were urine (*n* = 49, 39.2%), sputum (*n* = 22, 17.6%), and blood (*n* = 18, 14.4%), corresponding with the most common sites of mucosal *Serratia marcescens* colonization discussed previously in Chapter 1.1. The median hospital length of stay at time of specimen collection was 1.0 days (IQR: 0-10.0 days). Figure 6 shows the number of days from hospitalization to specimen acquisition, stratified by prior hospitalization status. Approximately half (*n* = 62/125) of included patients had received antimicrobials at the time of specimen collection.

All patient isolates underwent MALDI-TOF MS confirmatory identification using the VITEK MS®, as described in Section 2.7 previously. All (*n* = 125/125) isolates were identified as *Serratia marcescens* using VITEK MS®.

AST performed on the microbial isolates suggested two major AmpC-mediated resistance profiles: resistance to amoxicillin-clavulanate and cefoxitin, with susceptibility to ceftriaxone, suggestive of inducible AmpC expression; and resistance to all three aforementioned agents, suggestive of de-repression. The proportion of isolates with AST suggestive of AmpC-mediated resistance is shown in Table 4. The MIC distributions and interpretative breakpoints for all antimicrobials tested are summarized for the study population in Figures 7 and 8, respectively.

Table 3. Site of Specimen Collection for Patient Isolates with *Serratia marcescens* complex Infections

Specimen source	Patients, *n* (%)
Blood	18 (14.4)
Body fluid	8 (6.4)
Bronchoalveolar lavage fluid	6 (4.8)
Sputum	22 (17.6)
Tissue	5 (4.0)
Urine	49 (39.2)
Wound swab	17 (13.6)

Figure 6. Days from Hospitalization to Sample Acquisition

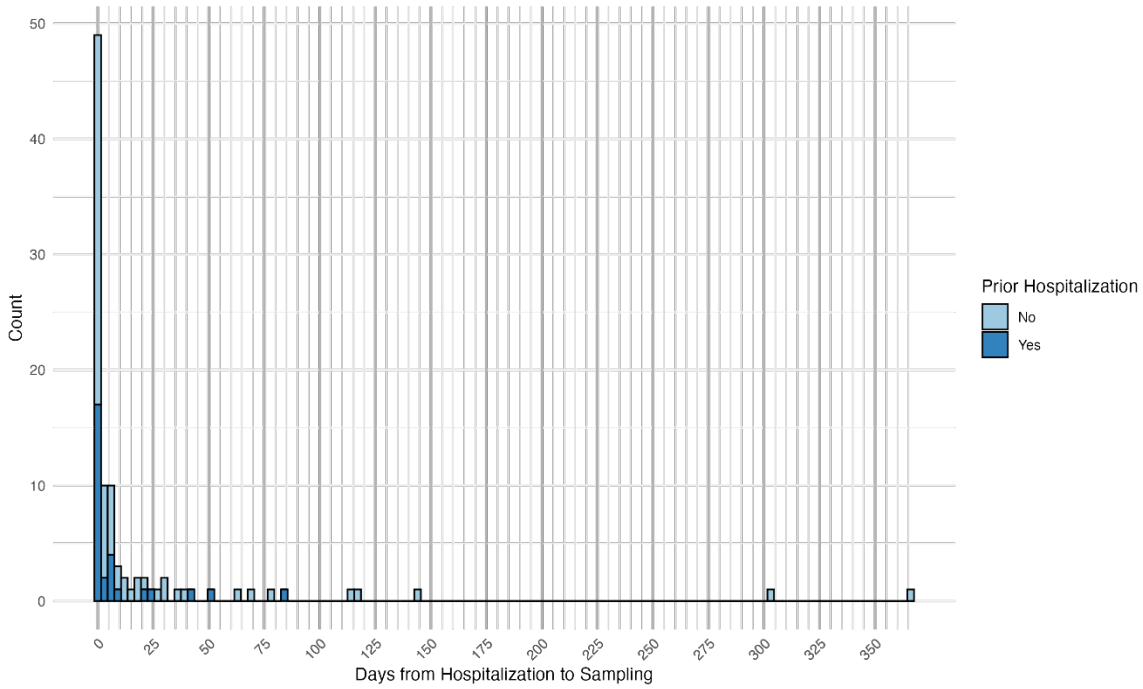


Table 4. Antimicrobial susceptibility patterns of *Serratia marcescens* complex isolates suggestive of AmpC-mediated resistance.

Phenotypic expression profile	Expected antimicrobial resistance	Number of specimens, *n* (%)
Basal (wild-type)	Amoxicillin-clavulanate **susceptible**	0 (0)
	Cefoxitin **susceptible**	
	Ceftriaxone **susceptible**	
Inducible	Amoxicillin-clavulanate **resistant**	109 (87.2)
	Cefoxitin **resistant**	
	Ceftriaxone **susceptible**	
De-repressed	Amoxicillin-clavulanate **resistant**	12 (9.6)
	Cefoxitin **resistant**	
	Ceftriaxone **resistant**	
Not applicable	One or more of amoxicillin-clavulanate, cefoxitin, or ceftriaxone not tested; *or*	4 (3.2)
	One or more of amoxicillin-clavulanate, cefoxitin, or ceftriaxone tested intermediate	

Figure 7. Distribution of Antimicrobial Minimum Inhibitory Concentrations for *Serratia marcescens* complex isolates.

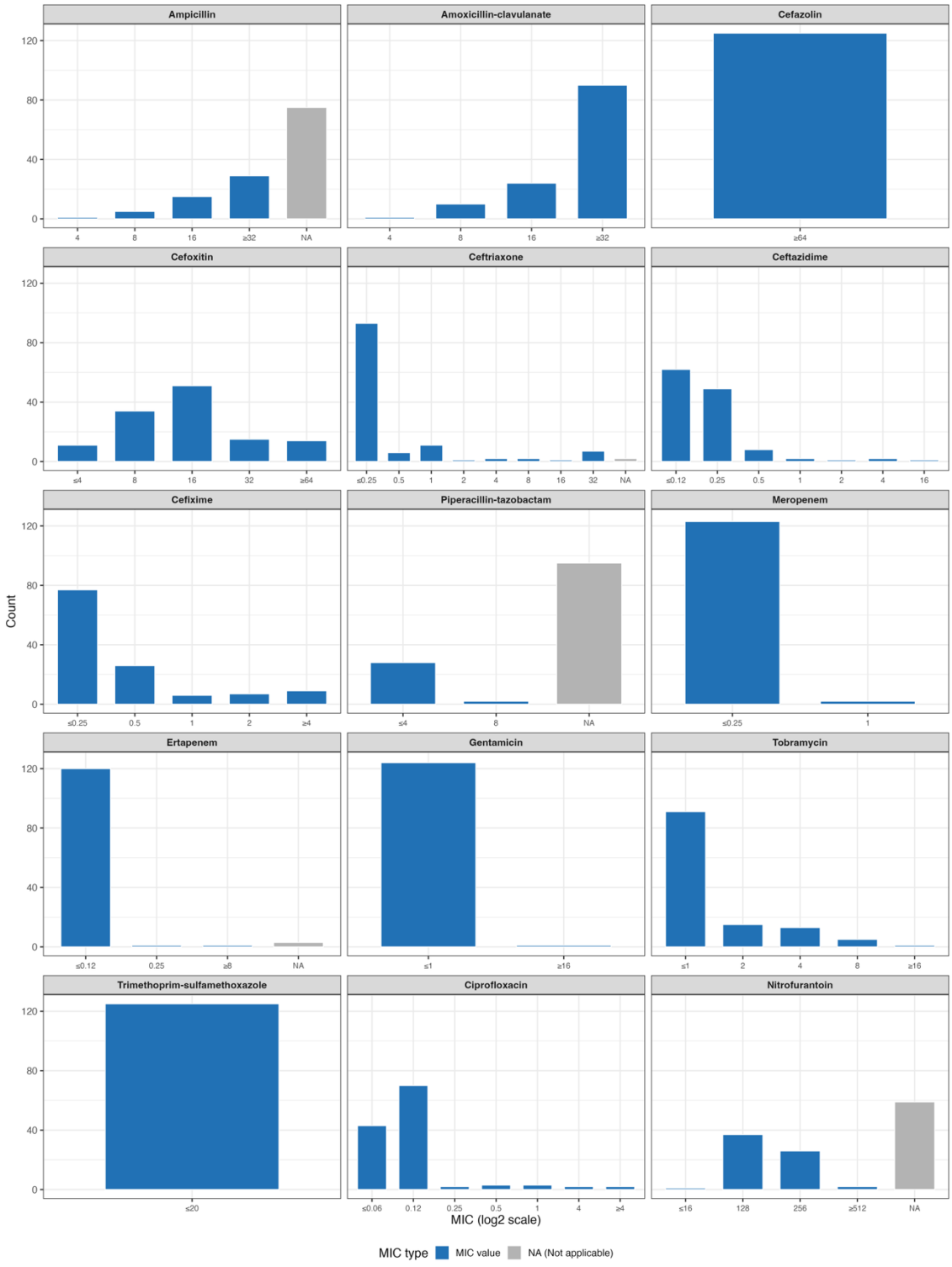


Figure 8. Distribution of Antimicrobial Interpretative Breakpoints for *Serratia marcescens* complex isolates.

3.3 Whole Genome Sequencing

We performed whole-genome sequencing on all 125 unique clinical isolates. The median genome size of the sequenced *Smc* isolates was 5.10 megabases (range: 4.34-5.74 Mb) with a median guanine-cytosine content of 59.8% (range: 58.5-60.3%). The median contig number was 36 (range: 17-1501), with a median largest contig length per genome of 760 kilobases (kb) (range: 38-2380) and a median N50 statistic of 317 kb. The average sequencing coverage per base was 197-fold (range: 37-330).

Table 5 illustrates the discrepancies between species-level identification using the VITEK MS® compared to those using whole genome sequencing with GTDB-Tk-assigned taxonomy. When comparing the species-level identification provided by the MALDI-TOF MS used routinely in the clinical laboratory to those assigned by sequencing-informed taxonomic assignment, we noted stark differences: roughly one-half (n=64/126, 51.2%) of the isolates were not *Serratia marcescens* sensu stricto. In fact, we found 19 to be *S. bockelmannii,* 4 to be *S. nematodiphila,* 19 to be *S. nevei,* and 22 to be *S. ureilytica.*

Table 5. Species-level Identification of *Serratia marcescens* complex by matrix-assisted laser desorption-ionization time-of-flight mass spectrometry compared to whole genome sequencing

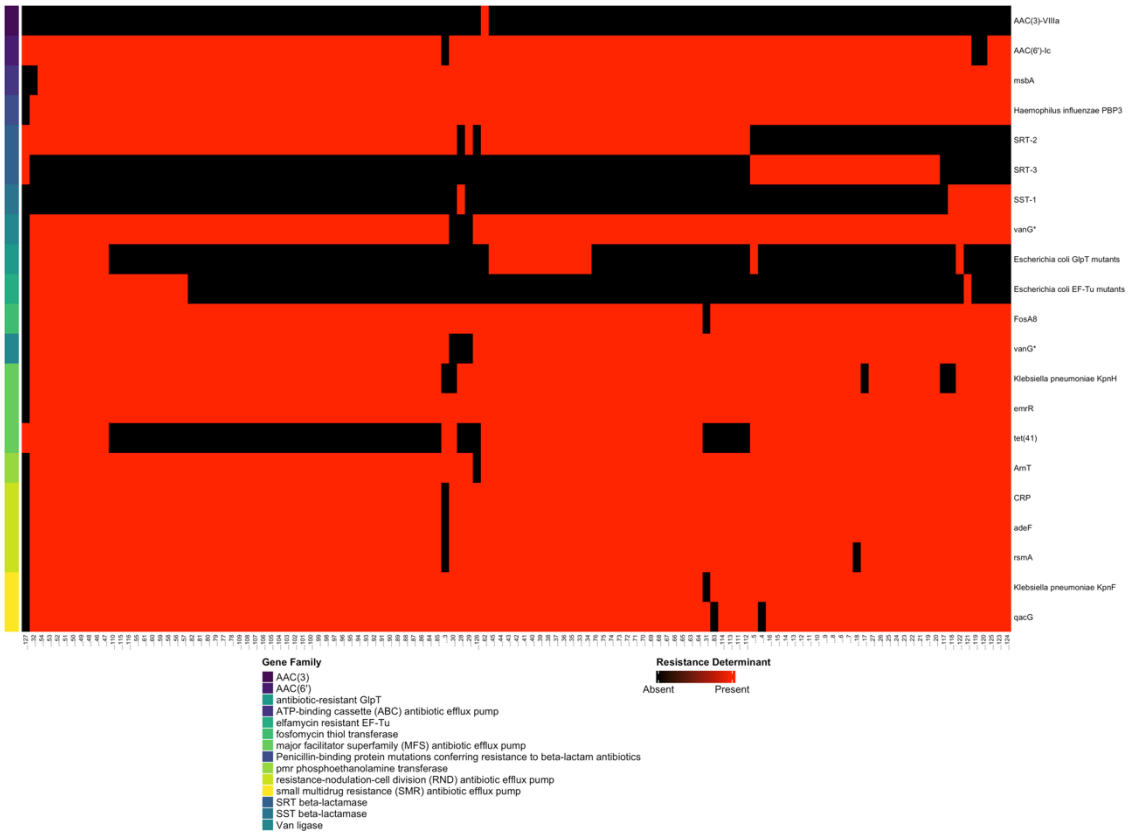
Identification technique	Species identified	Number of specimens, *n* (%)
Matrix-assisted laser desorption-ionization time-of-flight mass spectrometry (MALDI-TOF MS)	*Serratia marcescens*	125 (100.0)
Whole genome sequencing (WGS)	*Serratia bockelmannii*	19 (15.2)
	Serratia marcescens sensu stricto	61 (49.2)
	Serratia nematodiphila	4 (3.2)
	Serratia nevei	19 (15.2)
	Serratia ureilytica	22 (17.6)

3.4 Resistome Prediction

We identified 21 different predicted resistance determinants in the 125 sequenced isolates. Fifteen genes were found in a majority of sequenced isolates: *emrR*, a fluoroquinolone efflux pump (n=124/125, 99.2%); PBP3, a β-lactamase (n=124/125, 99.2%); *adeF*, a fluoroquinolone and tetracycline efflux pump (n=123/125, 98.4%); *ArnT,* a phosphoethanolamine transferase affecting peptide antimicrobials (n=123/125, 98.4%); *CRP*, a macrolide, fluoroquinolone, and penam antimicrobial efflux pump (n=123/125, 98.4%); *FosA8,* a fosfomycin thiol transferase (n=123/125, 98.4%); the nitroimidazole efflux pump *msbA* (n=123/125, 98.4%); *AAC(6')-Ic,* which produces the aminoglycoside N-acetyltransferase enzyme (n=122/125, 97.6%); *KpnF,* a multidrug efflux pump (n=123/125, 98.4%); *qacG,* an antiseptic efflux pump (n=122/125, 97.6%); *rsmA,* a fluoroquinolone, diaminopyrimidine, and phenicol antibiotic efflux pump (n=122/125, 97.6%); *vanG,* which confers vancomycin resistance by altering peptidoglycan precursor structure (n=121/125, 96.8%); *KpnH,* a multidrug efflux pump (n=119/125, 95.2%); *SRT-2,* a β-lactamase (n=90/125, 72.0%); and *tet(41),* a tetracycline efflux pump (n=74/125, 59.2%). We found a predicted AmpC β-lactamase in 124 of 125 sequenced isolates (99.2%). All isolates except isolate 138 harboured either a predicted SST-1, SRT-2, or SRT-3 class C β-lactamase. Notably, this isolate still retained an antimicrobial susceptibility pattern suggestive of induced AmpC expression, suggesting that an AmpC β-lactamase was present but not predicted using RGI criteria.

Figure 8 shows a heat map with the presence or absence of resistance genes in each isolate clustered by CARD RGI gene family ontology.

Figure 8. Resistance Gene Identifier Heat Map of Predicted Resistomes



Resistance determinants are mapped on the heat map as follows: the left-hand column shows the gene family according to the CARD RGI ontology; the right-hand axis labels show the gene names corresponding to the heatmap rows. The binary heat map shows either the presence of a resistance determinant for a given isolate in red, or the absence of said resistance determinant in black. The numerical isolate number is found on the x-axis of the heatmap.

3.5 Plasmid Prediction

In total, 42 (33.6%) of 125 isolates harboured 63 total plasmids; among those with plasmids, the median plasmid number per isolate was 1 (IQR: 1-2). The most common plasmid identified belonged to MOB cluster AB960 and is a non-mobilizable *Serratia marcescens* plasmid with a Mash nearest neighbour being NCBI Reference Sequence #CP042515.1 isolated in Australia in 2017. This plasmid occurred in 20 of 42 isolates (47.6%).

We found evidence of 13 antimicrobial resistance genes that were plasmid-borne in our patient population. Notably, there was no evidence of plasmid-borne AmpC operons or carbapenemase genes amongst predicted plasmid assemblies in our patient population. The most common resistance determinant found was *adeF,* a multidrug efflux pump affecting fluoroquinolones and tetracyclines, in 6 of 63 plasmids (9.5%). Other antimicrobial resistance genes that were found in multiple isolates include: *soxS,* the global regulator of *E. coli acrAB* multidrug efflux genes (n=3/63, 4.7%); *Mycobacterium tuberculosis* ndh mutant G313R conferring isoniazid resistance (n=3/63, 4.7%); *M. tuberculosis* kasA mutant D66N conferring isoniazid resistance (n=3/63, 4.7%); *tet(35),* a tetracycline efflux pump (n=2/63, 3.2%); *oqxA,* a fluoroquinolone efflux pump (n=2/63, 3.2%). Predicted resistance genes that only occurred once (n=1/63, 1.6%) in the patient population include: *tetA(58),* a tetracycline efflux pump; *evgA* and *evgS,* which regulate mdtEF and emrKY efflux pump expression; *mdtG,* which increases fosfomycin resistance when overexpressed; *vanH,* which decreases vancomycin binding affinity to peptidoglycan, and its transcriptional regulator *vanR;* and H-NS, a histone-like protein involved in the repression of RND-type transporters.

A heatmap of the collective set of plasmids in the study population, or the plasmidome, is found in Figure 9. We identified four plasmids that have not been previously described in the literature amongst our sequenced clinical isolates. Plasmid maps showing the reconstructed plasmids with gene annotations visible are found in Figures 10 through 13.

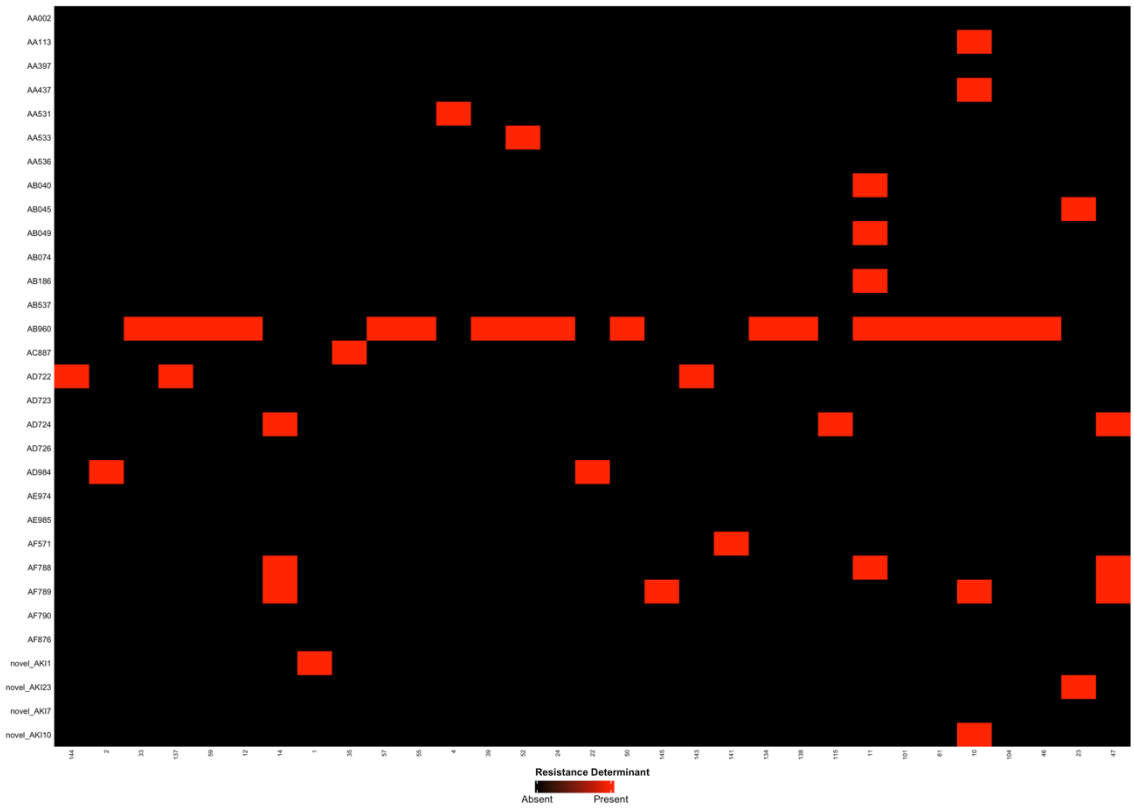
Novel plasmid AKI1 (NCBI Biosample SAMN47405310) is a 14-kilobase pair (kbp) conjugative plasmid found in a *Serratia bockelmannii* isolate (Figure 10). Plasmid AKI1 is a replicon type IncI1, relaxase type MOBF, and mating pair formation (mpf) type MPF_T plasmid. Its Mash nearest neighbour is NCBI Reference Sequence #NC_015972, a *Serratia marcescens* plasmid isolated from environmental sampling in Russia.

Novel plasmid AKI7 (NCBI Biosample SAMN47405316) is a 5.2-kbp mobilizable plasmid found in a *Serratia nevei* isolate (Figure 11). Plasmid AKI7 is a replicon type Col(Ye4449), relaxase type MOBP plasmid. Its Mash nearest neighbour is Genbank #CP040890.1, a *Leclercia adecarboxylata* plasmid isolated from patient stool in Hangzhou Province, China.

Novel plasmid AKI10 (NCBI Biosample SAMN47405319) is a 24-kbp conjugative plasmid found in a *Serratia marcescens* sensu stricto isolate (Figure 12). Plasmid AKI10 is a relaxase type MOBC, mpf type MPF_T plasmid. Its Mash nearest neighbour is Genbank #AP012552.1, a plasmid from a symbiont of the stink bug *Plautia stali* isolated in Japan.

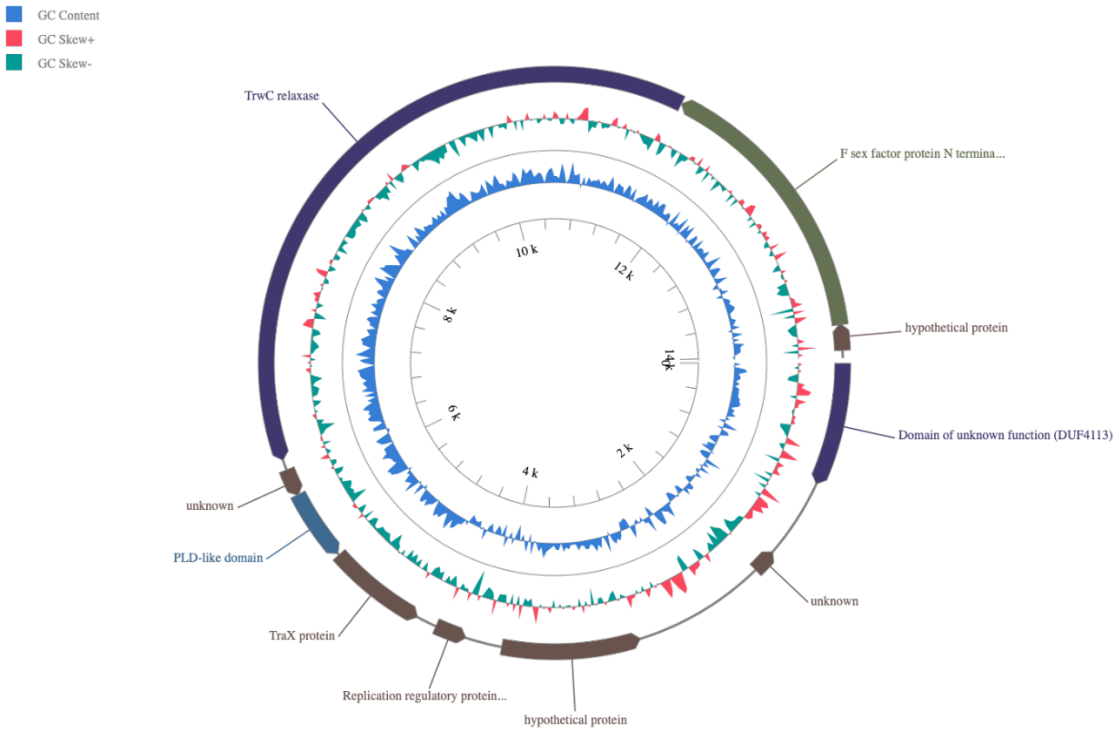
Finally, novel plasmid AKI23 (NCBI Biosample SAMN47405332) is a 58-kbp conjugative plasmid found in a *Serratia marcescens* sensu stricto isolate (Figure 13). Plasmid AKI23 is a replicon type IncI1, relaxase type MOBF, and mating pair formation (mpf) type MPF_F plasmid. Its Mash nearest neighbour is Genbank #CP027301.1, a plasmid isolated from environmental sampling in Singapore.

Figure 9. Heat Map of Predicted Plasmidome.



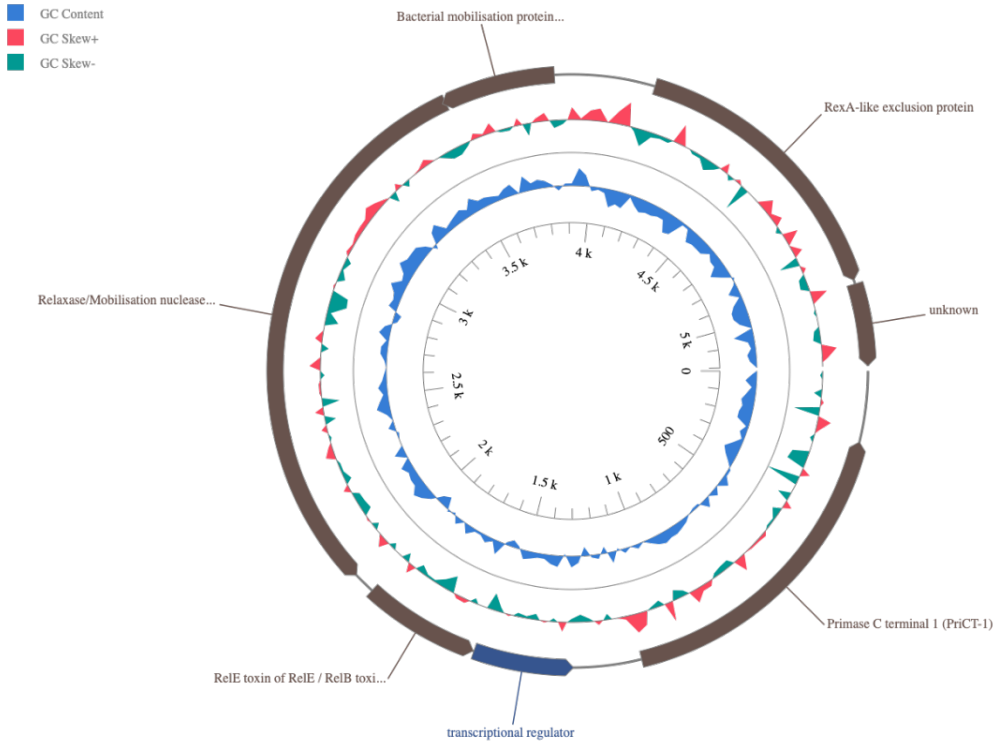
Predicted plasmids are mapped on the heat map as follows: the y-axis shows the MOB type of the predicted plasmids; the x-axis labels show the numerical isolate number for those isolates which had predicted plasmids, only. The binary heat map shows either the presence of a predicted plasmid for a given isolate in red, or the absence of said predicted plasmid in black.

Figure 10. Plasmid map of novel plasmid AKI1 identified from *Serratia bockelmannii* isolate



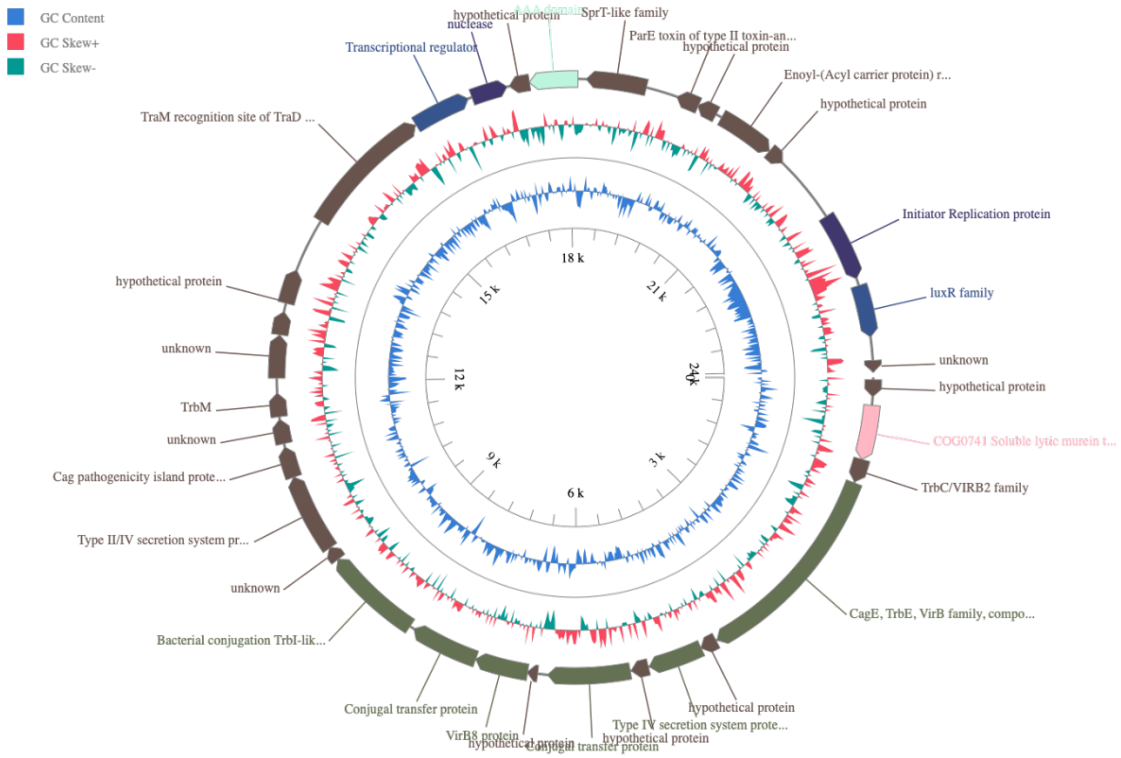
Plasmid map showing novel 14-kilobase pair conjugative plasmid sequenced from isolate 1, found on whole genome sequencing to be *Serratia bockelmannii.* Open reading frames labelled using arrows, with guanine-cytosine (GC) content labelled in blue, positive GC skew labelled in red, and negative GC skew labelled in green. Sequence data available under NCBI Bioproject PRJNA1232785, Biosample SAMN47405310.

Figure 11. Plasmid map of novel plasmid AKI7 identified from *Serratia nevei* isolate



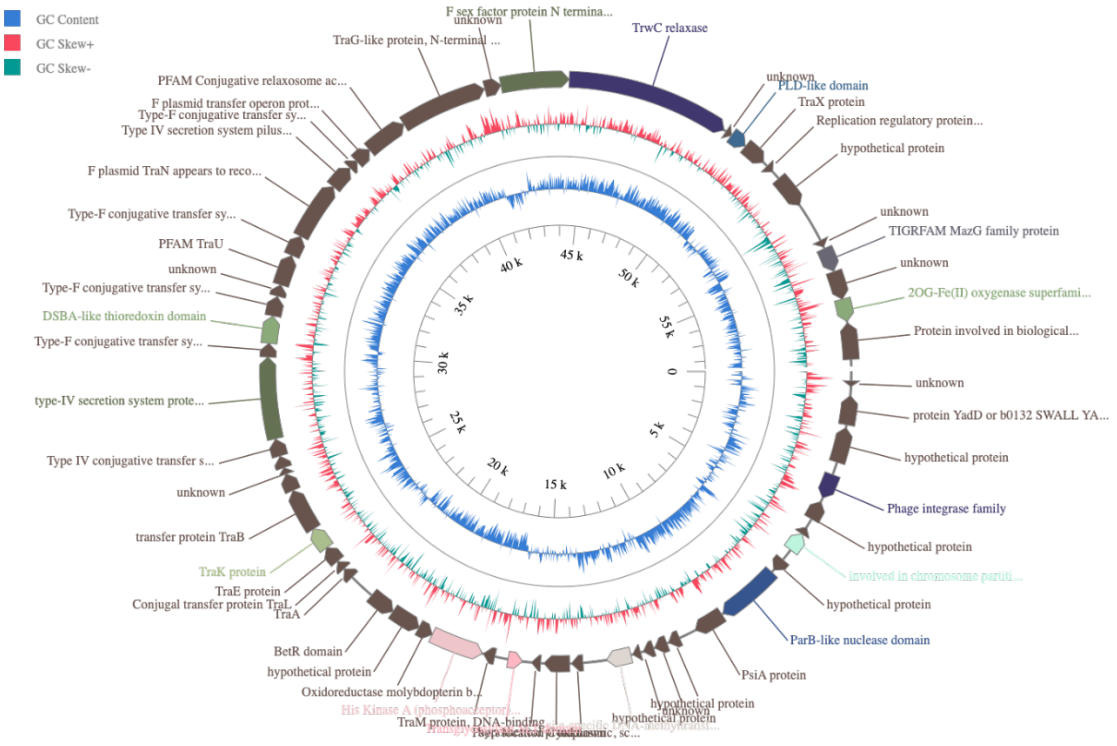
Plasmid map showing novel 5.2-kilobase pair mobilizable plasmid sequenced from isolate 7, found on whole genome sequencing to be *Serratia nevei.* Open reading frames labelled using arrows, with guanine-cytosine (GC) content labelled in blue, positive GC skew labelled in red, and negative GC skew labelled in green. Sequence data available under NCBI Bioproject PRJNA1232785, Biosample SAMN47405316.

Figure 12. Plasmid map of novel plasmid AKI10 identified from *Serratia marcescens* sensu stricto isolate



Plasmid map showing novel 24-kilobase pair conjugative plasmid sequenced from isolate 10, confirmed by whole genome sequencing to be *Serratia marcescens* sensu stricto. Open reading frames labelled using arrows, with guanine-cytosine (GC) content labelled in blue, positive GC skew labelled in red, and negative GC skew labelled in green. Sequence data available under NCBI Bioproject PRJNA1232785, Biosample SAMN47405319.

Figure 13. Plasmid map of novel plasmid AKI23 identified from *Serratia marcescens* sensu stricto isolate



Plasmid map showing novel 58-kilobase pair conjugative plasmid sequenced from isolate 23, confirmed by whole genome sequencing to be *Serratia marcescens* sensu stricto. Open reading frames labelled using arrows, with guanine-cytosine (GC) content labelled in blue, positive GC skew labelled in red, and negative GC skew labelled in green. Sequence data available under NCBI Bioproject PRJNA1232785, Biosample SAMN47405332.

3.6 Epidemiological and Phylogenetic Analyses

We plotted an epidemic curve to help examine and determine whether undetected transmission and/or an unrecognized outbreak of *Serratia marcescens* complex had occurred (Figure 14). We stratified the epidemic curve by specimen type and year to determine whether there were any differences between these variables. The baseline monthly case rate was approximately 2-3 cases/month, with increases above the baseline particularly notable in the spring months in both 2022 and 2024. Based on the shape of the epidemic curve, there was no point or propagated source. There are multiple peaks present during the periods of study, suggesting likely intermittent exposures or periods of risk. The use of phylogenetic analysis can further help differentiate whether these intermittent exposures are independent or possibly linked.

Figure 15 shows the inferred maximum-likelihood phylogenetic tree. Of 125 isolates, 16 pairs (n=32, 25.6%) were collected from the same hospital ward within 30 days of each other and were identified as *Serratia marcescens* by MALDI-TOF MS. When using whole genome sequencing-informed taxonomic identification, however, only two isolate pairs met the spatiotemporal linkage definition. Thus, most isolates did not, in fact, share a spatiotemporal link. For example: isolates 54 and 93 were clustered together phylogenetically and were spatially linked but not temporally linked, so we did not consider a putative transmission event to have occurred between these patients because the present study did not include environmental sampling to help inform said linkage. Taken together, this suggests that the possible intermittent exposure pattern to the epidemic curve is a result of independent exposures rather than a disease outbreak.

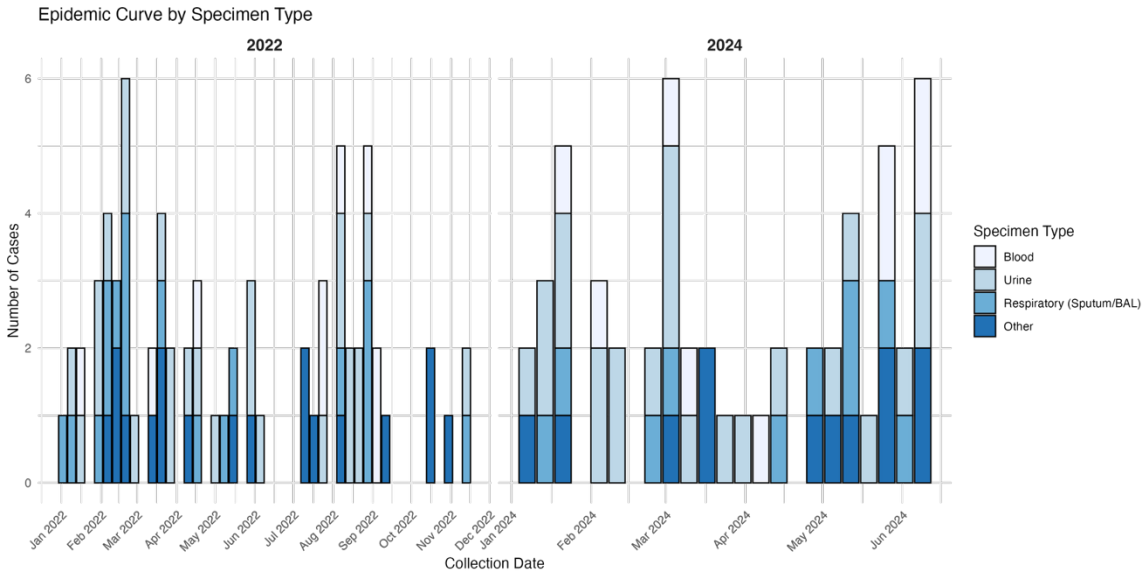
As mentioned above, isolate 50 and 56, as well as isolate 106 and 113 appeared to cluster together phylogenetically and overlap spatiotemporally.

For isolates 50 and 56, both of these *Serratia nevei* isolates were collected within one month of each other in July and August 2024, respectively. Both isolates were collected on post-admission day one. We performed a pairwise USEARCH alignment using the OrthoANIu tool,

which showed 98.80% ANI over an average aligned length of 3.74 Mb (isolate 50 genome coverage 70.41%, isolate 56 genome coverage 68.64%). Notably, this isolate pair did not therefore meet the prespecified ANI cutoff of ≥99.0%. Isolate 56 had an additional predicted resistance determinant, *EF-Tu,* conferring elfamycin resistance, compared to isolate 50 (Figure 8). The phenotypic AST data for both isolates is identical, save for piperacillin-tazobactam (isolate 50 MIC = 8 mg/L, isolate 56 MIC ≤ 4 mg/L) and ciprofloxacin (isolate 50 MIC = 0.12 mg/L, isolate 56 MIC ≤ 0.06 mg/L).

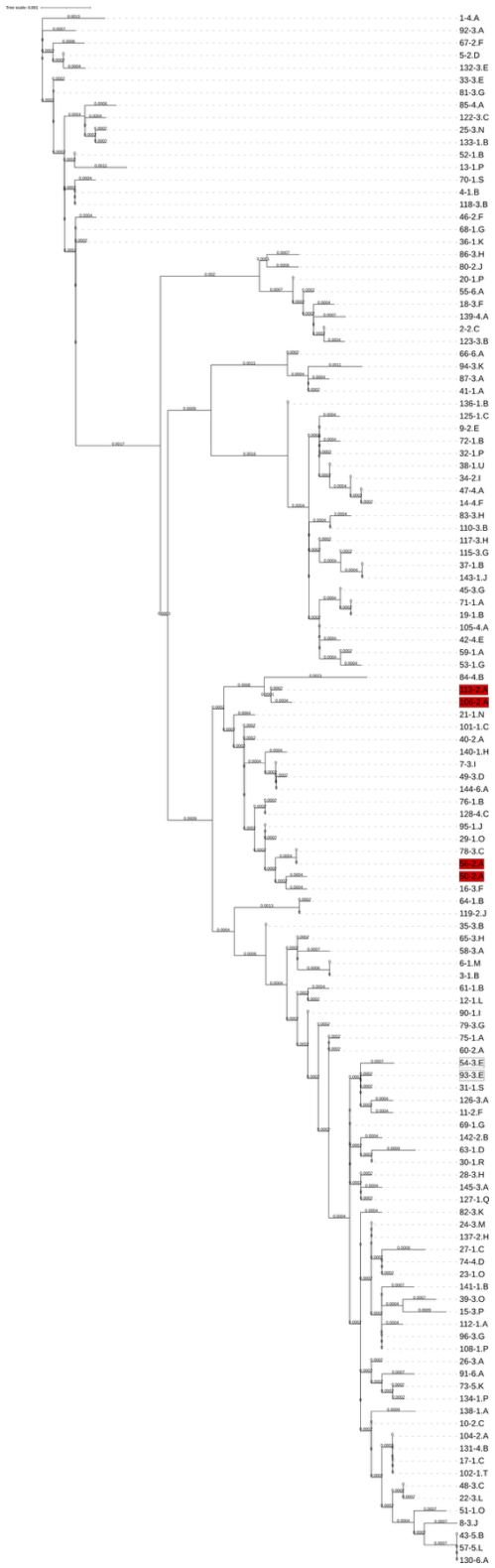
For isolates 106 and 113, both of these *Serratia nevei* isolates were collected in March 2024 within 10 days of each other, with each isolate collected on post-admission day one. We performed a pairwise USEARCH alignment using the OrthoANIu tool, which showed 99.27% ANI over an average aligned length of 3.77 Mb (isolate 106 genome coverage 75.53%, isolate 113 genome coverage 76.29%). Both isolates shared an identical predicted resistome (Figure 8). The phenotypic AST data for both isolates is identical, save for cefoxitin (isolate 106 MIC = 16 mg/L, isolate 113 MIC ≤4 mg/L) and cefixime (isolate 106 MIC = 0.5 mg/L, isolate 113 MIC ≤ 0.25 mg/L).

Figure 14. Epidemic Curve of *Serratia marcescens* Invasive Infections



Epidemic curve showing the number of new invasive *Serratia marcescens* cases identified on a weekly basis, stratified by year of occurrence as well as by specimen type.

Figure 15. Phylogenetic tree of sequenced *Serratia marcescens* complex clinical isolates



Phylogenetic tree visualized using iToL, v. 7 (110). Each tree node has an associated length showing the distance from its nearest neighbour. Adjacent nodes are labelled in red if patients had spatiotemporal overlap. Each node has a unique specimen code assigned to it in the format “x-y.z”, where x is the numeric study identification number, y is the anonymized numeric code for the hospital where the specimen originated, and z is the anonymized letter code for the ward where the specimen originated.

3.7 Sensitivity Analysis

We performed a sensitivity analysis by expanding the definition of temporal overlap between *Serratia marcescens* complex isolates from one month to two and three months to determine whether setting this arbitrary cut-off had an unintended effect on our primary analysis. On visual inspection of the phylogenetic tree (Figure 15), we identified the following isolate pairs that clustered closely together with spatial overlap: isolates 3 and 61; 33 and 132; and 54 and 93. None of the three isolate pairs identified met the definition of spatiotemporal overlap in either sensitivity analysis, suggesting that the methodological choice to use one month as the temporal cutoff to define possible transmission was unlikely to have had an impact on the study's overall conclusions.

Chapter 4. Discussion

4.1 Summary of findings

In this retrospective cohort study with prospective specimen identification, we evaluated the molecular epidemiology of clinical infections caused by *Serratia marcescens* complex in a large hospital system using whole genome sequencing and bioinformatics analytical techniques. This multi-year study of Canadian patients over two periods from 2022 to 2024 in six hospitals including adult and pediatric tertiary and secondary care centres sheds light on numerous facets of medical care, including the identification and reporting of *Serratia marcescens* complex infections in the clinical microbiology laboratory, and infection prevention and control practices relevant to this species complex.

This study represents – to our knowledge – the largest multi-year, multi-hospital survey of *Serratia marcescens* complex infections that uses whole genome sequencing to analyze clinical isolates. A key finding of our study from a clinical laboratory perspective was the significant species-level discordance between traditional identification methods using MALDI-TOF MS and whole genome sequencing with GTDB-Tk-assigned taxonomic identification. Approximately half (51.2%) of isolates initially identified by MALDI-TOF MS as *Serratia marcescens* were found to be other species of the genus *Serratia* belonging to the newly described *Serratia marcescens* complex, which includes *S. bockelmannii, S. marcescens* sensu stricto, *S. nematodiphila, S. nevei,* and *S. ureilytica* (10-13). This is in spite of MALDI-TOF MS identification showing a high (≥95%) certainty of species-level identification to *Serratia marcescens* for all included study isolates. The VITEK® MALDI- TOF MS United States Food and Drug Administration 501(k)-cleared database, version 3.2, does not contain *S. bockelmannii, S. nematodiphila, S. nevei,* or *S. ureilytica,* which explains the misidentification of multiple isolates as *Serratia marcescens* (117). This discrepancy highlights a substantial limitation in standard clinical laboratory identification techniques, as the current mass spectrum databases used by MALDI-TOF MS

instruments do not adequately reflect contemporary knowledge of the genetic diversity of the *Serratia marcescens* complex.

Our study also confirms the presence and coexistence of multiple members of the *Serratia marcescens* complex in hospitalized patients. This finding has significant implications for the investigation of possible outbreaks of *Serratia marcescens* complex in hospital settings. If the clinical microbiological techniques used to identify *Serratia marcescens* complex in the laboratory are insufficiently granular to differentiate amongst constituent members of the complex, it may lead isolates identified by MALDI-TOF MS during outbreak investigations to be erroneously included in line lists.

For example, suppose a point prevalence survey is conducted during a period where infection control practitioners noted a higher than baseline incidence of invasive *Serratia marcescens* infections in a NICU. During the survey, two patients are identified as being colonized with *Serratia marcescens* based on chromogenic agar identification and subsequent MALDI-TOF MS confirmation and added to a line list during the possible outbreak investigation; these patients would be subject to enhanced control measures including environmental cleaning and contact precautions *inter alia* while the outbreak is being investigated. If accurate MALDI-TOF MS databases with sufficient granularity were available in the laboratory and showed that these patients were colonized with different organisms – for example, *Serratia nevei* and *Serratia bockelmannii* – it would obviate the need for significant infection prevent and control measures to be enacted, possibly reducing hospitalization costs, length of stay, *etc.* MALDI-TOF MS manufacturers must urgently update their mass spectra databases to reflect the taxonomic shifts in *Serratia marcescens* complex to reflect these realities and prevent undue downstream consequences.

Another finding from our study which is germane to infection prevention and control practices is that putative transmission events between patients are likely rare. Only two pairs of isolates showed spatiotemporal overlap, which just one pair meeting the stringent ANI threshold we set to determine possible transmission events. While there were slight differences in

phenotypic AST profiles in the single pair of isolates with ≥99.0% ANI, we note the isolates had an identical CARD RGI resistome prediction profile. This suggests that possible patient-to-patient transmission of *Serratia marcescens* complex infections in hospitals is likely uncommon, and most infections likely represent sporadic, independent events. Further evidence to support this conclusion is the fact that the sole pair of isolates with spatiotemporal overlap and ≥99.0% ANI were sampled on post-admission day one, making nosocomial transmission less likely. Furthermore, these findings provide an evidence base for the current use of routine infection control practices without gowns and gloves during interaction with the patient infected or colonized with *Serratia marcescens* complex organisms and/or their hospital environment. Given the common underlying mechanism of resistance shared with other AmpC-producing Enterobacterales, this finding is likely generalizable to *Enterobacter* spp., *Citrobacter freundii, Morganella morganii,* and *Providencia* spp.

4.2 Comparison to prior literature

The question of whether appreciable nosocomial transmission of AmpC-producing Enterobacterales exists has not been significantly investigated in the medical literature. Recently, Aracil-Gisbert and colleagues reported a study of 1432 environmental isolates of *Serratia* spp. from environmental sampling of ICU sinks as well as 99 clinical isolates from a large hospital in Spain (55). They noted multiple unique findings in their isolates, including a basal AmpC expression pattern amongst many environmental isolates which also lacked prodigiosin pigmentation, carrying $bla_{\text{VIM-1}}$or $bla_{\text{OXA-48}}$ on IncL/pB77-CPsm plasmids persistent since 2017, suggesting that environmental reservoirs serve as milieux for the propagation and dissemination of carbapenemases amongst *Serratia marcescens* isolates (55). These findings help explain the likely mechanism of propagation of AmpC-producing Enterobacterales in hospital outbreaks through the use of a source-sink dynamic model, where the contact with the static environmental reservoir “source” allows the transmission of the organism to the transient patient “sink” (55).

Contrary to the findings of Aracil-Gisbert and colleagues, we did not identify a subpopulation of *Serratia marcescens* complex organisms in our study displaying a basal AmpC expression profile with susceptibility to amoxicillin-clavulanate, cefoxitin, and a third-generation cephalosporin. We posit this may be due to a lack of environmental sampling in our study, since there is less externally applied selective pressure on the *Serratia marcescens* complex organisms in the environment than in the hospitalized human host. It is also possible that AmpC expression plays a role in adaptation to non-host environments and altered gene expression may lead to a basal-type susceptibility pattern (55). In our present study, we only found evidence of inducible or de-repressed AmpC expression, based on the characteristic AST patterns of included isolates described previously in section 3.2. We found only 9.6% of our isolates exhibited AmpC de-repression, in line with Infectious Diseases Society of America guidelines that designate *Serratia marcescens* a “low risk” AmpC Enterobacterales organism with respect to de-repression potential (52). Of note, in contrast to Aracil-Gisbert and colleagues, we also did not identify any predicted carbapenemases amongst our study isolates.

With respect to laboratory identification of members of the *Serratia marcescens* complex, our findings corroborate those of Harch *et al.,* who examined the similarly poor correlation of the other major MALDI-TOF MS platform marketed by Bruker Corporation (Billerica, United States of America) with whole genome sequencing-based identification of *Serratia marcescens* complex constituent members (118). In their study, Harch and colleagues compared 19 *Serratia marcescens* isolates identified as such by MALDI-TOF MS with whole genome sequencing-based identification followed by a comparison of Kraken2 and ribosomal MLST tools for species assignment (118). They found that ribosomal MLST showed a superior ability to assign species-level identification, likely due to the lack of inclusion of *Serratia bockelmannii* in the database software version used, and the difficulty of assigning species identifications from the lack of insufficiently unique *k*-mers – sequences of consecutive nucleotides of length *k* – amongst the constituent species of the *Serratia marcescens* complex (118).

4.3 Study strengths

This study has several strengths. A major methodological strength of this study is its evaluation of over 100 clinical isolates, which allowed a determination of whether appreciable *Serratia marcescens* transmission occurred in the hospital setting. The study population consisted of both adult and pediatric patients, with predominant specimen types mirroring the most common sites of colonization of *Serratia marcescens* complex in the human body. Unique among current literature, this study was also a multicentre analysis over multiple years, with two discrete sampling periods. The use of a robust bioinformatics pipeline, including sophisticated tools such as GTDB-Tk for accurate taxonomic assignment and CARD for detailed resistome prediction helped provide high-quality genomic data analysis.

From a methodological perspective, the study used consecutive patient samples to reduce the risk of bias by including only independent observations. There was low risk of attrition bias, as only two non-duplicate specimens were excluded from the study. In addition, the study minimized the risk of incomplete data by using a pre-piloted data extraction process that ensured the reliability and robustness of the data collection process.

4.4 Study limitations

This study has several limitations worth mentioning. The most notable limitation of the study's methodology was the lack of concomitant and prospective environmental sampling, which severely restricts the ability to establish definitive causal relationships for samples with possible spatiotemporal overlap. Without environmental data, the present study was unable to identify potential reservoirs or vectors, thus limiting interpretations to putative transmission events rather than conclusive mechanistic findings. This approach using only clinical isolates also significantly limits the ability to fully understand possible outbreaks if transmission had indeed occurred.

Furthermore, the fact that the median time to specimen collection was on day 1 post-admission suggests that many isolates were likely community-acquired infections, rather than nosocomial.

As a retrospective cohort study with a chart review component, this study is also prone to misclassification bias and is susceptible to confounding by unmeasured variables. We used a pre-piloted data collection form and identified consecutive patients over multiple time periods to minimize this. However, in spite of these efforts, residual confounding may persist. In addition, it was not feasible to perform statistical techniques including stratification, multivariable logistic regression, or propensity score matching, to adjust for possible confounders; stratification would have resulted in a decrease in statistical power, whereas multivariable logistic regression may have rendered imprecise estimates with low event numbers, as seen in the study, and propensity score matching would not be appropriate given it is unclear which measured variables are true or potential confounders.

An additional limitation of the study is the study was powered to detect transmission between patients and was not powered to assess clinical outcomes. As a result, the study cannot assess whether certain AmpC expression profiles are associated with risk of putative transmission, whether specific species within the *Serratia marcescens* complex or indeed whether certain AmpC expression profiles are associated with increased risk of morbidity and/or mortality.

Finally, the study's definition of a one-month temporal threshold to define putative transmission events was arbitrary in nature, as no literature-guided timeframe existed for *Serratia marcescens* complex. We attempted to limit the introduction of bias into our study from this methodological choice by performing a sensitivity analysis, which showed no increase in the number of putative transmission events when a threshold of two or three months was used to define temporal overlap between cases.

4.5 Study implications

The study findings hold significant implications for infection prevention and control practices: the infrequent observations of putative transmission events in the hospital environment suggests that widespread implementation of stringent contact precautions and enhanced environmental cleaning is unlikely to be warranted for *Serratia marcescens* complex infections, with the important exception of suspected or confirmed outbreaks. These findings thus add empirical evidence to a practice that was heretofore largely defined by a low theoretical risk of transmission due to the lack of abundance of resistance determinants on mobile genetic elements. While we noted a number of isolates with plasmids harbouring resistance determinants, we did not find any plasmid AmpC operons or carbapenemase genes amongst our patient cohort.

Our results suggest that current mass spectrum databases for both major manufacturers do not offer sufficient granularity within the *Serratia marcescens* complex to provide a species-level identification. While the clinical relevance of the constituent species of the *Serratia marcescens* complex continue to be elucidated, laboratory reporting and MALDI-TOF MS databases should identify isolates only to the level of "*Serratia marcescens* complex". Until such time as MALDI-TOF MS databases are accurately updated, to reflect current taxonomic understanding, it may be necessary to rely on costlier whole genome sequencing or MLST to characterize isolates in a suspected clinical outbreak rather than relying on conventional clinical species-level identification provided by MALDI-TOF MS. Continued reliance on traditional MALDI-TOF MS may result in misattribution of infections to the incorrect species: in the event that infection rates are above a perceived baseline threshold, this may trigger additional infection prevention and control investigations and interventions that would be otherwise unneeded if updated taxonomic identification was able to be applied clinically. Therefore, whole genome sequencing may be necessary in infection prevention and control investigations of possible *Serratia marcescens* complex outbreaks until such time as other techniques are able to provide the appropriate resolution to serve this clinical purpose.

4.6 Directions for future research

Future studies of *Serratia marcescens* complex and its transmission dynamics in the nosocomial setting should focus on the generation of large datasets that are prospectively collected and followed. Such studies should include environmental sampling and may also wish to be powered to assess clinical outcomes, such as whether AmpC expression levels alter the risk of transmission in the rare occasions that it occurs, and whether specific species within the *Serratia marcescens* complex are associated with increased risk of morbidity and/or mortality.

While anticipating higher quality prospective studies with sufficient power to analyze the aforementioned outcomes, our present study adds to the literature by showing that current laboratory techniques offer insufficient species-level granularity within the *Serratia marcescens* complex, and – while limited by the lack of environmental sampling – suggests that transmission of *Serratia marcescens* complex infections between hospital patients is rare.

Chapter 5. References

1. Cooney S, O'Brien S, Iversen C, Fanning S. Bacteria: Other Pathogenic Enterobacteriaceae – Enterobacter and Other Genera. In: Motarjemi Y, editor. Encyclopedia of Food Safety. Waltham: Academic Press; 2014. p. 433-41.
2. Mahlen SD. Serratia infections: from military experiments to current practice. Clin Microbiol Rev. 2011;24(4):755-91.
3. Grimont PA, Grimont F. The genus Serratia. Annu Rev Microbiol. 1978;32:221-48.
4. Abreo E, Altier N. Pangenome of Serratia marcescens strains from nosocomial and environmental origins reveals different populations and the links between them. Sci Rep. 2019;9(1):46.
5. Lecca-Caballero D, Vega-Moreno E, Cabanillas-Chirinos L, Diaz Del Aguila K, Rojas-Villacorta W, Salvatierra-Espinola W, et al. An In Vitro Study of the Effects of Temperature and pH on Lead Bioremoval Using Serratia marcescens. Sustainability. 2023;15(19):14048.
6. Laupland KB, Parkins MD, Gregson DB, Church DL, Ross T, Pitout JD. Population-based laboratory surveillance for Serratia species isolates in a large Canadian health region. Eur J Clin Microbiol Infect Dis. 2008;27(2):89-95.
7. Diekema DJ, Pfaller MA, Jones RN, Doern GV, Winokur PL, Gales AC, et al. Survey of bloodstream infections due to gram-negative bacilli: frequency of occurrence and antimicrobial susceptibility of isolates collected in the United States, Canada, and Latin America for the SENTRY Antimicrobial Surveillance Program, 1997. Clin Infect Dis. 1999;29(3):595-607.
8. Jones RN. Microbial etiologies of hospital-acquired bacterial pneumonia and ventilator-associated bacterial pneumonia. Clin Infect Dis. 2010;51 Suppl 1:S81-7.
9. Bhadra B, Roy P, Chakraborty R. Serratia ureilytica sp. nov., a novel urea-utilizing species. Int J Syst Evol Microbiol. 2005;55(Pt 5):2155-8.

10. Zhang CX, Yang SY, Xu MX, Sun J, Liu H, Liu JR, et al. Serratia nematodiphila sp. nov., associated symbiotically with the entomopathogenic nematode Heterorhabditidoides chongmingensis (Rhabditida: Rhabditidae). Int J Syst Evol Microbiol. 2009;59(Pt 7):1603-8.
11. Williams DJ, Grimont PAD, Cazares A, Grimont F, Ageron E, Pettigrew KA, et al. The genus Serratia revisited by genomics. Nat Commun. 2022;13(1):5195.
12. Cho GS, Stein M, Brinks E, Rathje J, Lee W, Suh SH, et al. Serratia nevei sp. nov. and Serratia bockelmannii sp. nov., isolated from fresh produce in Germany and reclassification of Serratia marcescens subsp. sakuensis Ajithkumar et al. 2003 as a later heterotypic synonym of Serratia marcescens subsp. marcescens. Syst Appl Microbiol. 2020;43(2):126055.
13. Ono T, Taniguchi I, Nakamura K, Nagano DS, Nishida R, Gotoh Y, et al. Global population structure of the Serratia marcescens complex and identification of hospital-adapted lineages in the complex. Microb Genom. 2022;8(3).
14. Matteoli FP, Pedrosa-Silva F, Dutra-Silva L, Giachini AJ. The global population structure and beta-lactamase repertoire of the opportunistic pathogen Serratia marcescens. Genomics. 2021;113(6):3523-32.
15. Dessì A, Puddu M, Testa M, Marcialis MA, Pintus MC, Fanos V. Serratia marcescens infections and outbreaks in neonatal intensive care units. J Chemother. 2009;21(5):493-9.
16. Martineau C, Li X, Lalancette C, Perreault T, Fournier E, Tremblay J, et al. Serratia marcescens Outbreak in a Neonatal Intensive Care Unit: New Insights from Next-Generation Sequencing Applications. Journal of Clinical Microbiology. 2018;56(9):10.1128/jcm.00235-18.
17. Al Jarousha AM, El Qouqa IA, El Jadba AH, Al Afifi AS. An outbreak of Serratia marcescens septicaemia in neonatal intensive care unit in Gaza City, Palestine. J Hosp Infect. 2008;70(2):119-26.

18. Cullen MM, Trail A, Robinson M, Keaney M, Chadwick PR. Serratia marcescens outbreak in a neonatal intensive care unit prompting review of decontamination of laryngoscopes. J Hosp Infect. 2005;59(1):68-70.
19. Fleisch F, Zimmermann-Baer U, Zbinden R, Bischoff G, Arlettaz R, Waldvogel K, et al. Three consecutive outbreaks of Serratia marcescens in a neonatal intensive care unit. Clin Infect Dis. 2002;34(6):767-73.
20. Villari P, Crispino M, Salvadori A, Scarcella A. Molecular epidemiology of an outbreak of Serratia marcescens in a neonatal intensive care unit. Infect Control Hosp Epidemiol. 2001;22(10):630-4.
21. Madani TA, Alsaedi S, James L, Eldeek BS, Jiman-Fatani AA, Alawi MM, et al. Serratia marcescens-contaminated baby shampoo causing an outbreak among newborns at King Abdulaziz University Hospital, Jeddah, Saudi Arabia. J Hosp Infect. 2011;78(1):16-9.
22. Lai KK, Baker SP, Fontecchio SA. Rapid eradication of a cluster of Serratia marcescens in a neonatal intensive care unit: use of epidemiologic chromosome profiling by pulsed-field gel electrophoresis. Infect Control Hosp Epidemiol. 2004;25(9):730-4.
23. Bourdin T, Benoit M, Monnier A, Bédard E, Prévost M, Charron D, et al. Serratia marcescens Colonization in a Neonatal Intensive Care Unit Has Multiple Sources, with Sink Drains as a Major Reservoir. Appl Environ Microbiol. 2023;89(5):e0010523.
24. Uduman SA, Farrukh AS, Nath KN, Zuhair MY, Ifrah A, Khawla AD, et al. An outbreak of Serratia marcescens infection in a special-care baby unit of a community hospital in United Arab Emirates: the importance of the air conditioner duct as a nosocomial reservoir. J Hosp Infect. 2002;52(3):175-80.
25. Schaberg DR, Alford RH, Anderson R, Farmer JJ, 3rd, Melly MA, Schaffner W. An outbreak of nosocomial infection due to multiply resistant Serratia marcescens: evidence of interhospital spread. J Infect Dis. 1976;134(2):181-8.

26. Schaberg DR, Weinstein RA, Stamm WE. Epidemics of nosocomial urinary tract infection caused by multiply resistant gram-negative bacilli: epidemiology and control. J Infect Dis. 1976;133(3):363-6.
27. de Vries JJ, Baas WH, van der Ploeg K, Heesink A, Degener JE, Arends JP. Outbreak of Serratia marcescens colonization and infection traced to a healthcare worker with long-term carriage on the hands. Infect Control Hosp Epidemiol. 2006;27(11):1153-8.
28. Taxt AM, Eldholm V, Kols NI, Haugan MS, Raffelsberger N, Asfeldt AM, et al. A national outbreak of Serratia marcescens complex: investigation reveals genomic population structure but no source, Norway, June 2021 to February 2023. Euro Surveill. 2025;30(5).
29. Sunenshine RH, Tan ET, Terashita DM, Jensen BJ, Kacica MA, Sickbert-Bennett EE, et al. A multistate outbreak of Serratia marcescens bloodstream infection associated with contaminated intravenous magnesium sulfate from a compounding pharmacy. Clin Infect Dis. 2007;45(5):527-33.
30. Blossom D, Noble-Wang J, Su J, Pur S, Chemaly R, Shams A, et al. Multistate outbreak of Serratia marcescens bloodstream infections caused by contamination of prefilled heparin and isotonic sodium chloride solution syringes. Arch Intern Med. 2009;169(18):1705-11.
31. Liu D, Zhang LP, Huang SF, Wang Z, Chen P, Wang H, et al. Outbreak of Serratia marcescens infection due to contamination of multiple-dose vial of heparin-saline solution used to flush deep venous catheters or peripheral trocars. J Hosp Infect. 2011;77(2):175-6.
32. Sawa T, Kooguchi K, Moriyama K. Molecular diversity of extended-spectrum β-lactamases and carbapenemases, and antimicrobial resistance. Journal of Intensive Care. 2020;8(1):13.
33. Philippon A, Slama P, Dény P, Labia R. A Structure-Based Classification of Class A β-Lactamases, a Broadly Diverse Family of Enzymes. Clin Microbiol Rev. 2016;29(1):29-57.
34. Philippon A, Arlet G, Labia R, Iorga BI. Class C β-Lactamases: Molecular Characteristics. Clin Microbiol Rev. 2022;35(3):e0015021.

35. Jacoby GA. AmpC beta-lactamases. Clin Microbiol Rev. 2009;22(1):161-82, Table of Contents.

36. Tamma PD, Doi Y, Bonomo RA, Johnson JK, Simner PJ. A Primer on AmpC β-Lactamases: Necessary Knowledge for an Increasingly Multidrug-resistant World. Clin Infect Dis. 2019;69(8):1446-55.

37. Honoré N, Nicolas MH, Cole ST. Inducible cephalosporinase production in clinical isolates of Enterobacter cloacae is controlled by a regulatory gene that has been deleted from Escherichia coli. Embo j. 1986;5(13):3709-14.

38. Sverak HE, Yaeger LN, Worrall LJ, Vacariu CM, Glenwright AJ, Vuckovic M, et al. Cryo-EM characterization of the anydromuropeptide permease AmpG central to bacterial fitness and β-lactam antibiotic resistance. Nature Communications. 2024;15(1):9936.

39. Schmidtke AJ, Hanson ND. Model system to evaluate the effect of ampD mutations on AmpC-mediated beta-lactam resistance. Antimicrob Agents Chemother. 2006;50(6):2030-7.

40. Kaneko K, Okamoto R, Nakano R, Kawakami S, Inoue M. Gene mutations responsible for overexpression of AmpC beta-lactamase in some clinical isolates of Enterobacter cloacae. J Clin Microbiol. 2005;43(6):2955-8.

41. Macdougall C. Beyond Susceptible and Resistant, Part I: Treatment of Infections Due to Gram-Negative Organisms With Inducible β-Lactamases. J Pediatr Pharmacol Ther. 2011;16(1):23-30.

42. Sanders CC, Bradford PA, Ehrhardt AF, Bush K, Young KD, Henderson TA, et al. Penicillin-binding proteins and induction of AmpC beta-lactamase. Antimicrob Agents Chemother. 1997;41(9):2013-5.

43. López-Argüello S, Montaner M, Oliver A, Moya B. Molecular Basis of AmpC β-Lactamase Induction by Avibactam in Pseudomonas aeruginosa: PBP Occupancy, Live Cell Binding Dynamics and Impact on Resistant Clinical Isolates Harboring PDC-X Variants. Int J Mol Sci. 2021;22(6).

44. Nong L, Jonker M, de Leeuw W, Wortel MT, Ter Kuile B. Progression of ampC amplification during de novo amoxicillin resistance development in E. coli. mBio. 2025;16(2):e0298224.
45. CLSI. Performance Standards for Antimicrobial Susceptibility Testing. CLSI Supplement M100. 35th. ed. Wayne, PA, USA: Clinical and Laboratory Standards Institute.
46. Kohlmann R, Bähr T, Gatermann SG. Species-specific mutation rates for ampC derepression in Enterobacterales with chromosomally encoded inducible AmpC β-lactamase. J Antimicrob Chemother. 2018;73(6):1530-6.
47. Choi SH, Lee JE, Park SJ, Kim MN, Choo EJ, Kwak YG, et al. Prevalence, microbiology, and clinical characteristics of extended-spectrum beta-lactamase-producing Enterobacter spp., Serratia marcescens, Citrobacter freundii, and Morganella morganii in Korea. Eur J Clin Microbiol Infect Dis. 2007;26(8):557-61.
48. Hilty M, Sendi P, Seiffert SN, Droz S, Perreten V, Hujer AM, et al. Characterisation and clinical features of Enterobacter cloacae bloodstream infections occurring at a tertiary care university hospital in Switzerland: is cefepime adequate therapy? Int J Antimicrob Agents. 2013;41(3):236-49.
49. Chow JW, Fine MJ, Shlaes DM, Quinn JP, Hooper DC, Johnson MP, et al. Enterobacter bacteremia: clinical features and emergence of antibiotic resistance during therapy. Ann Intern Med. 1991;115(8):585-90.
50. Kaye KS, Cosgrove S, Harris A, Eliopoulos GM, Carmeli Y. Risk factors for emergence of resistance to broad-spectrum cephalosporins among Enterobacter spp. Antimicrob Agents Chemother. 2001;45(9):2628-30.
51. Jacobson KL, Cohen SH, Inciardi JF, King JH, Lippert WE, Iglesias T, et al. The relationship between antecedent antibiotic use and resistance to extended-spectrum cephalosporins in group I beta-lactamase-producing organisms. Clin Infect Dis. 1995;21(5):1107-13.
52. Tamma PD, Heil EL, Justo JA, Mathers AJ, Satlin MJ, Bonomo RA. Infectious Diseases Society of America 2024 Guidance on the Treatment of Antimicrobial-Resistant Gram-Negative Infections. Clinical Infectious Diseases. 2024:ciae403.

53. European Committee for Antimicrobial Susceptibility Testing. Determination of minimum inhibitory concentrations (MICs) of antibacterial agents by broth dilution. Clinical Microbiology and Infection. 2003;9(8):ix-xv.

54. Turnidge J, Paterson DL. Setting and revising antibacterial susceptibility breakpoints. Clin Microbiol Rev. 2007;20(3):391-408, table of contents.

55. Aracil-Gisbert S, Fernández-De-Bobadilla Miguel D, Guerra-Pinto N, Serrano-Calleja S, Pérez-Cobas Ana E, Soriano C, et al. The ICU environment contributes to the endemicity of the “Serratia marcescens complex” in the hospital setting. mBio. 2024;15(5):e03054-23.

56. Karanika S, Karantanos T, Arvanitis M, Grigoras C, Mylonakis E. Fecal Colonization With Extended-spectrum Beta-lactamase–Producing Enterobacteriaceae and Risk Factors Among Healthy Individuals: A Systematic Review and Metaanalysis. Clinical Infectious Diseases. 2016;63(3):310-8.

57. Namikawa H, Imoto W, Yamada K, Tochino Y, Kaneko Y, Kakeya H, et al. Predictors of mortality from extended-spectrum beta-lactamase-producing Enterobacteriaceae bacteremia. Emerg Microbes Infect. 2023;12(1):2217951.

58. Abubakar U, Tangiisuran B, Elnaem MH, Sulaiman SAS, Khan FU. Mortality and its predictors among hospitalized patients with infections due to extended spectrum beta-lactamase (ESBL) Enterobacteriaceae in Malaysia: a retrospective observational study. Future Journal of Pharmaceutical Sciences. 2022;8(1):17.

59. Nguyen MN, Gladstone BP, De Angelis G, Biggel M, Xavier BB, Lammens C, et al. Tracing carriage, acquisition, and transmission of ESBL-producing Escherichia coli over two years in a tertiary care hospital. Genome Medicine. 2024;16(1):151.

60. Harris AD, Kotetishvili M, Shurland S, Johnson JA, Morris JG, Nemoy LL, et al. How important is patient-to-patient transmission in extended-spectrum beta-lactamase Escherichia coli acquisition. Am J Infect Control. 2007;35(2):97-101.

61. Harris AD, Perencevich EN, Johnson JK, Paterson DL, Morris JG, Strauss SM, et al. Patient-to-patient transmission is important in extended-spectrum beta-lactamase-producing Klebsiella pneumoniae acquisition. Clin Infect Dis. 2007;45(10):1347-50.

62. Ontario Agency for Health Protection and Promotion (Public Health Ontario), Provincial Infectious Diseases Advisory Committee. Routine practices and additional precautions in all health care settings. 3rd ed., 4th revision ed. Toronto, ON: King's Printer for Ontario; 2025.

63. Jamal AJ, Mataseje LF, Williams V, Leis JA, Tijet N, Zittermann S, et al. Genomic Epidemiology of Carbapenemase-Producing Enterobacterales at a Hospital System in Toronto, Ontario, Canada, 2007 to 2018. Antimicrob Agents Chemother. 2021;65(8):e0036021.

64. Carling PC. Wastewater drains: epidemiology and interventions in 23 carbapenem-resistant organism outbreaks. Infect Control Hosp Epidemiol. 2018;39(8):972-9.

65. Kizny Gordon AE, Mathers AJ, Cheong EYL, Gottlieb T, Kotay S, Walker AS, et al. The Hospital Water Environment as a Reservoir for Carbapenem-Resistant Organisms Causing Hospital-Acquired Infections-A Systematic Review of the Literature. Clin Infect Dis. 2017;64(10):1435-44.

66. De Geyter D, Blommaert L, Verbraeken N, Sevenois M, Huyghens L, Martini H, et al. The sink as a potential source of transmission of carbapenemase-producing Enterobacteriaceae in the intensive care unit. Antimicrobial Resistance & Infection Control. 2017;6(1):24.

67. Parkes LO, Hota SS. Sink-Related Outbreaks and Mitigation Strategies in Healthcare Facilities. Curr Infect Dis Rep. 2018;20(10):42.

68. Kotay S, Chai W, Guilford W, Barry K, Mathers AJ. Spread from the Sink to the Patient: In Situ Study Using Green Fluorescent Protein (GFP)-Expressing Escherichia coli To Model Bacterial Dispersion from Hand-Washing Sink-Trap Reservoirs. Appl Environ Microbiol. 2017;83(8).

69. Kotay SM, Donlan RM, Ganim C, Barry K, Christensen BE, Mathers AJ. Droplet- Rather than Aerosol-Mediated Dispersion Is the Primary Mechanism of Bacterial Transmission from Contaminated Hand-Washing Sink Traps. Appl Environ Microbiol. 2019;85(2).

70. Hagiya H, Otsuka F. Increased evidence for no benefit of contact precautions in preventing extended-spectrum β-lactamases-producing Enterobacteriaceae: Systematic scoping review. American Journal of Infection Control. 2023;51(9):1056-62.

71. Mariappan S, Sekar U, Kamalanathan A. Carbapenemase-producing Enterobacteriaceae: Risk factors for infection and impact of resistance on outcomes. Int J Appl Basic Med Res. 2017;7(1):32-9.

72. Toth DJA, Khader K, Beams A, Samore MH. Model-based Assessment of the Effect of Contact Precautions Applied to Surveillance-detected Carriers of Carbapenemase-producing Enterobacteriaceae in Long-term Acute Care Hospitals. Clin Infect Dis. 2019;69(Suppl 3):S206-s13.

73. Fleifel M, Machmouchi A, Alameddine O, Hoyek K, Melki D, Hallab E, et al. The Spread of Plasmidic AmpC in a General Lebanese Hospital Over Nine Consecutive Years and the Relationship With Restricted Isolation Protocols. Front Med (Lausanne). 2021;8:633783.

74. Reintjes R, Zanuzdana A. Outbreak Investigations. Modern Infectious Disease Epidemiology. 2010;28:159-76.

75. Simar SR, Hanson BM, Arias CA. Techniques in bacterial strain typing: past, present, and future. Curr Opin Infect Dis. 2021;34(4):339-45.

76. Lewallen S, Courtright P. Epidemiology in practice: case-control studies. Community Eye Health. 1998;11(28):57-8.

77. Chyou PH. Patterns of bias due to differential misclassification by case-control status in a case-control study. Eur J Epidemiol. 2007;22(1):7-17.

78. Corraini P, Olsen M, Pedersen L, Dekkers OM, Vandenbroucke JP. Effect modification, interaction and mediation: an overview of theoretical insights for clinical investigators. Clin Epidemiol. 2017;9:331-8.

79. Wang X, Kattan MW. Cohort Studies: Design, Analysis, and Reporting. Chest. 2020;158(1s):S72-s8.

80. Gannon V, Albright T, Wells T, Cahak C, Harrington AM, Buchan B, et al. Error Reduction in Specimen Processing of Irretrievable Body Fluids. American Journal of Clinical Pathology. 2020;154(Supplement_1):S130-S1.

81. Miroshnychenko A, Zeraatkar D, Phillips MR, Bakri SJ, Thabane L, Bhandari M, et al. Cohort studies investigating the effects of exposures: key principles that impact the credibility of the results. Eye. 2022;36(5):905-6.

82. Tennant PWG, Murray EJ, Arnold KF, Berrie L, Fox MP, Gadd SC, et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. International Journal of Epidemiology. 2021;50(2):620-32.

83. Platt RW, Schisterman EF, Cole SR. Time-modified confounding. Am J Epidemiol. 2009;170(6):687-94.

84. Kyriacou DN, Lewis RJ. Confounding by Indication in Clinical Research. Jama. 2016;316(17):1818-9.

85. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. 1996;49(12):1373-9.

86. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behav Res. 2011;46(3):399-424.

87. Thabane L, Mbuagbaw L, Zhang S, Samaan Z, Marcucci M, Ye C, et al. A tutorial on sensitivity analyses in clinical trials: the what, why, when and how. BMC Med Res Methodol. 2013;13:92.

88. Hu T, Chitnis N, Monos D, Dinh A. Next-generation sequencing technologies: An overview. Hum Immunol. 2021;82(11):801-11.

89. Kozińska A, Seweryn P, Sitkiewicz I. A crash course in sequencing for a microbiologist. J Appl Genet. 2019;60(1):103-11.

90. Li IC, Wu R, Hu CW, Wu KM, Chen ZW, Chou CH. Comparison of Conventional Molecular and Whole-Genome Sequencing Methods for Differentiating Salmonella enterica Serovar Schwarzengrund Isolates Obtained from Food and Animal Sources. Microorganisms. 2021;9(10).

91. Tenover FC, Arbeit RD, Goering RV, Mickelsen PA, Murray BE, Persing DH, et al. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. J Clin Microbiol. 1995;33(9):2233-9.

92. Oehler JB, Wright H, Stark Z, Mallett AJ, Schmitz U. The application of long-read sequencing in clinical settings. Human Genomics. 2023;17(1):73.

93. Hernandez SI, Berezin CT, Miller KM, Peccoud SJ, Peccoud J. Sequencing Strategy to Ensure Accurate Plasmid Assembly. ACS Synth Biol. 2024;13(12):4099-109.

94. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol. 2008;61(4):344-9.

95. Singhal N, Kumar M, Kanaujia PK, Virdi JS. MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. Front Microbiol. 2015;6:791.

96. Bobenchik AM, Hindler JA, Giltner CL, Saeki S, Humphries RM. Performance of Vitek 2 for antimicrobial susceptibility testing of Staphylococcus spp. and Enterococcus spp. J Clin Microbiol. 2014;52(2):392-7.

97. Slatko BE, Gardner AF, Ausubel FM. Overview of Next-Generation Sequencing Technologies. Curr Protoc Mol Biol. 2018;122(1):e59.

98. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20.

99. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol. 2017;13(6):e1005595.

100. Mikheenko A, Saveliev V, Hirsch P, Gurevich A. WebQUAST: online evaluation of genome assemblies. Nucleic Acids Research. 2023;51(W1):W601-W6.
101. Lewin HA, Graves JAM, Ryder OA, Graphodatsky AS, O'Brien SJ. Precision nomenclature for the new genomics. Gigascience. 2019;8(8).
102. Alhakami H, Mirebrahim H, Lonardi S. A comparative evaluation of genome assembly reconciliation tools. Genome Biol. 2017;18(1):93.
103. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microb Genom. 2021;7(11).
104. Saratto T, Visuri K, Lehtinen J, Ortega-Sanz I, Steenwyk JL, Sihvonen S. Solu: a cloud platform for real-time genomic pathogen surveillance. BMC Bioinformatics. 2025;26(1):12.
105. Robertson J, Nash JHE. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. Microb Genom. 2018;4(8).
106. Alcock BP, Huynh W, Chalil R, Smith KW, Raphenya AR, Wlodarski MA, et al. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. Nucleic Acids Res. 2023;51(D1):D690-d9.
107. McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, et al. The comprehensive antibiotic resistance database. Antimicrob Agents Chemother. 2013;57(7):3348-57.
108. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics. 2020;36(6):1925-7.
109. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Molecular Biology and Evolution. 2015;32(1):268-74.
110. Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Research. 2024;52(W1):W78-W82.

111. Yoon SH, Ha SM, Lim J, Kwon S, Chun J. A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie Van Leeuwenhoek. 2017;110(10):1281-6.

112. Rodriguez RL, Conrad RE, Viver T, Feistel DJ, Lindner BG, Venter SN, et al. An ANI gap within bacterial species that advances the definitions of intra-species units. mBio. 2024;15(1):e0269623.

113. Wickham H. ggplot2: Elegant Graphics for Data Analysis: Springer-Verlag New York; 2016.

114. Wickham H, François R, Henry L, Müller K, Vaughan D. dplyr: A Grammar of Data Manipulation. 2025.

115. Ianone R, Roy O. DiagrammeR: Graph/Network Visualization. R package version 1011. 2024.

116. Gu Z. Complex heatmap visualization. Imeta. 2022;1(3):e43.

117. bioMérieux. VITEK MS. Indications for Use Statement - Form 3881: United States Department of Health and Human Services, Food and Drug Administration; [Available from: https://go.biomerieux.com/us_fda_510_k_organisms_vitek_ms.

118. Harch Susan AJ, Jenkins F, Farhat R, van Hal Sebastiaan J. Complexities in species identification for Serratia marcescens complex for the modern microbiology laboratory. Microbiology Spectrum. 2024;13(1):e01361-24.

Chapter 6. Appendix

Published Journal Article

This manuscript reproduced and included *verbatim* below has been published in the Journal of Infectious Diseases on 2025-08-19 under the Digital Object Identifier 10.1093/infdis/jiaf392. As per the publisher, Oxford University Press, authors retain the right to include the article in full or in part in a thesis, provided that it is not published commercially (Journal of Infectious Diseases editorial staff, personal communication, 2025-6-12).

Molecular epidemiology of clinical infections caused by Serratia marcescens complex in a tertiary care hospital system: insights from whole genome sequencing

Adam S. Komorowski, MD[1-5,*]; Michael G. Surette, PhD[4,6,7]; Laura Rossi, MSc[6,7]; Dominique Tertigas, BSc[6]; Mark Gaskin, MLT[2]; Shahrokh Shekarriz, PhD[6]; Andrew G. McArthur, PhD[4,6]; Marek Smieja, MD PhD[1-5,7]; Dominik Mertz, MD MSc[3,4,8]

[1] Department of Pathology and Molecular Medicine, McMaster University, Hamilton, ON, Canada

[2] Department of Laboratory Medicine, Hamilton Health Sciences and St. Joseph's Healthcare, Hamilton, ON, Canada

[3] Department of Health Research Methodology, Evidence, and Impact, Faculty of Health Sciences, McMaster University

[4] Michael G. DeGroote Institute for Infectious Disease Research, McMaster University

[5] Research Institute of St. Joseph's Healthcare Hamilton, Hamilton, ON, Canada

[6] Department of Biochemistry and Biomedical Sciences, Faculty of Health Sciences, McMaster University

[7] Division of Gastroenterology, Department of Medicine, McMaster University

[8] Division of Infectious Diseases, Department of Medicine, McMaster University

* Corresponding author

ORCiD

Adam S. Komorowski: 0000-0003-2101-7701

Michael G. Surette: 0000-0003-0086-4995

Laura Rossi: 0000-0002-1963-2484

Dominique Tertigas: 0009-0008-3245-1355

Andrew G. McArthur: 0000-0002-1142-3063

Marek Smieja: 0000-0003-1655-3761

Dominik Mertz: 0000-0003-4337-1613

Corresponding author

Adam S. Komorowski, MD FRCPC

Hamilton General Hospital, Microbiology Lab

237 Barton St E, Room W-1 1-88, Hamilton, ON, Canada L8L 2X2

Phone: +1 905 521 2100

Email: komorowa@mcmaster.ca

Keywords

Gram-negative bacteria; Enterobacterales; Serratia marcescens; Infection Control; Whole Genome Sequencing; Drug Resistance, Microbial

Abstract

Background

Serratia marcescens is an opportunistic AmpC beta-lactamase-producing Enterobacterales associated with intensive care unit outbreaks causing high morbidity and mortality. The spatiotemporal dynamics of Serratia species and their implications for infection prevention and control practices in hospital remain understudied.

Methods

We prospectively identified patient culture specimens in a multi-hospital academic healthcare system from 2022-2024. We included first-time isolates of Serratia marcescens identified via conventional culture and confirmed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS). Isolates underwent whole genome sequencing on the Illumina NextSeq 2000 platform. We queried assembled genomes using the Comprehensive Antibiotic Resistance Database to identify resistance genes and predict resistomes. We constructed a maximum-likelihood phylogenetic tree using GTDB-Tk-assigned taxonomies. We determined possible links between patient isolates if there was a spatiotemporal overlap between

patients and the average nucleotide identity (ANI) of their Serratia genome sequences was >99.0%. We used retrospective chart review to collect relevant patient characteristics and analyzed Serratia infections using descriptive statistics.

Results

We identified 147 patient isolates and included 125. Phenotypic testing revealed either inducible or de-repressed AmpC expression in all isolates. Whole genome sequencing found discordant species-level identification compared to MALDI-TOF MS in 64 (51.2%) isolates, suggesting the presence of multiple members of the recently described Serratia marcescens complex (Smc) causing hospital- or community-associated infections. We identified one isolate pair with a spatiotemporal link and ANI >99.0%, only.

Conclusions

Current MALDI-TOF MS-based identification methods are insufficient to identify Smc and laboratory reporting should be modified to report only to the level of the complex. Taxonomic changes within Smc may affect infection prevention and control efforts.

Background

Serratia species are opportunistic Gram-negative bacterial pathogens of the order Enterobacterales (1). Among Serratia species, S. marcescens is particularly well-known as a cause of significant morbidity and mortality, particularly in neonatal intensive care unit (NICU) outbreaks (2, 3). It is a frequent colonizer of the epidermis, as well as the gastrointestinal, urinary, and respiratory tracts in humans, and has been found to persist on medical equipment and sinks in the hospital environment (4-6). Recent evidence suggests that Serratia marcescens, while causing most human infections within the genus, is but one constituent species of a wider Serratia marcescens complex (Smc). S. bockelmannii, S. marcescens sensu stricto, S. nematodiphila, S. nevei, and S. ureilytica have been proposed as constituent Smc members (7, 8).

Smc species possess an Ambler class C β-lactamase, known as AmpC enzymes, which catalyze most β-lactams, cephalosporins, and β-lactam/β-lactamase inhibitor combinations through serine-

mediated hydrolysis of the β-lactam ring (9). AmpC activity can occur either through chromosomally encoded inducible resistance, stable de-repression via mutations in regulatory genes, or constitutive expression on mobile genetic elements such as plasmids (9-11). Other Enterobacterales such as Enterobacter cloacae, Klebsiella aerogenes, or Citrobacter freundii are at significantly higher risk of inducible AmpC production, with mutation rates favouring AmpC de-repression being 50 to 150-fold higher among these “high-risk” species than in S. marcescens (12). Inducible chromosomal AmpC expression results from the use of aminopenicillins, cephalosporin, and cephamycin antimicrobials, thus liberating oligopeptide residues during bacterial cell lysis that compete with uridine disphosphate-N-acetylmuramic acid to bind AmpR, the negative regulator of AmpC expression (13). Conformational changes in AmpR caused by oligopeptide binding result in increasing AmpC production, leading to risk of treatment failure with the use of β-lactams, cephalosporins, and β-lactam/β-lactamase inhibitors (9).

Enterobacterales with constitutively expressed, plasmid-borne extended-spectrum β-lactamases (ESBLs) are widely recognised as significant causes of morbidity and mortality that are transmitted in healthcare settings, primarily through person-to-person contact and environmental persistence (14, 15). Gowns and gloves may be used during contact with the patients and/or their environments for those known to be colonized with ESBLs and enhanced environmental cleaning may be instituted to prevent colonization and subsequent infection of other patients (16). Such infection prevention and control recommendations are not routinely in place for AmpC-producing Enterobacterales like Smc. The recent work of Aracil-Gisbert and colleagues suggested, however, that environmental reservoirs may play a role in the endemicity of AmpC-producing Smc isolates in an intensive care unit environment in Spain (8). For example, they showed that environmental reservoirs of Smc can serve as milieux for the propagation of plasmids containing carbapenemase genes (8). It is conceivable that Smc may propagate and persist in the wider hospital environment, particularly given its predilection for water reservoirs (17), and contribute to patient infections. In this study, we aimed to assess whether unrecognized transmission of Smc may be occurring within our hospital system by determining the genomic relatedness of patient

specimens that tested positive for Serratia marcescens by conventional clinical microbiological identification techniques.

Methods

Study design

We prospectively identified clinical specimens with Serratia marcescens from three adult tertiary care sites, two secondary care sites, and one pediatric tertiary care site (a total of 2,800 beds) from one city in Ontario, Canada, from January to December 2022, and again from January to June 2024. These specimens with Serratia marcescens identified from culture were confirmed using the VITEK® matrix-assisted laser desorption ionization time-of-flight mass spectrometer (MALDI-TOF MS, bioMérieux, Marcy-l'Étoile, France). The time periods in the study represented a convenience sample of consecutive patients.

Specimens were included if they were submitted to diagnose active infection; those submitted to the laboratory to diagnose microbial colonization (i.e., rectal swabs) for the purposes of infection prevention and outbreak control were excluded from the analysis. Specimens sampled from any anatomical site, from hospital inpatients and outpatients at the eligible hospital sites were included. Only the first Serratia marcescens isolate from each hospital admission was included. At the end of each prospective specimen identification period (e.g., December 2022, July 2024), a retrospective chart review was undertaken to identify patient characteristics and provide critical epidemiologic information. The data extraction process was piloted on the first five identified isolates to ensure all relevant variables were being collected. Microbiologic data extracted from the patient chart were as follows: specimen site, specimen collection date, species identified by MALDI TOF-MS, percent certainty of MALDI TOF-MS identification >95%, minimum inhibitory concentrations (MIC) and interpretative breakpoints for all antimicrobials tested. Epidemiological data extracted from patient medical charts included: patient age; patient sex; hospital; ward; whether the patient was a resident of a congregate setting; the number of hospitalizations in the past year; the principal admission diagnosis; the length of stay at time of specimen collection; whether the patient was receiving antibiotics at specimen collection; which antibiotics were being

received at specimen collection (if any); which antibiotics were used to treat the infection; treatment duration; patient outcome; and whether hospital discharge summary included infection as a cause of death, if relevant.

The primary objective of the study was to determine, using whole genome sequencing combined with epidemiological data, whether clinically unrecognized clusters of transmission of Serratia marcescens are occurring in the hospital setting.

Laboratory methods

In the clinical microbiology laboratory, specimens were identified as per local standard operating procedures using conventional bacterial media and incubation conditions appropriate to the various specimen types submitted by requesting physicians. All isolates underwent final confirmatory identification using the VITEK® MALDI TOF-MS (bioMérieux, Marcy-l'Étoile, France) using the V.3.2 database to assign species identifications, and antimicrobial susceptibility testing (AST) using the VITEK® 2 system (bioMérieux, Marcy-l'Étoile, France), which performs miniaturized broth microdilution MIC testing using an automated Gram-negative AST-N390 card. MICs were interpreted by the VITEK® 2 system according to the Clinical and Laboratory Standards Institute (CLSI) M100 guideline currently used in the laboratory at the time of testing (18). Bacterial isolates were frozen at –80°C in a 25% glycerol stock solution and transported for whole genome sequencing.

We purified genomic DNA (gDNA) using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI) as per manufacturer instructions and subsequently quantified using the PicoGreen dsDNA Assay (ThermoFisher) to ensure a minimum gDNA concentration of 30 ng/µL was available for downstream applications. Individually-barcoded short-read libraries were enriched for ~1000bp insert sizes using the NEBNext Ultra II FS DNA Library Prep Kit (NEB, Ipswich, MA) and the ProNex Size-Selective Purification System (Promega, Madison, WI). The library was sequenced on the Illumina NextSeq 2000 to generate 2x150bp paired-end reads from an average of 3.2 million clusters (range: 0.6-5.6 million).

Bioinformatics

To filter and trim adapter sequences from the raw reads, we used Trimmomatic v.0.38 (19) specifying a quality score cut-off of 3 and a minimum read-length of 36 base pairs. We then assembled filtered reads using Unicycler v.0.5.1 (20). We evaluated genome assembly quality and calculated whole-genome guanine-cytosine content using Quast v.5.2.0 (21). To annotate contigs to identify relevant genes, we used Bakta v.1.5.0 (22). We then queried the Comprehensive Antibiotic Resistance Database (CARD, v.3.3.0)'s Resistance Gene Identifier (RGI) tool (23, 24) to predict the resistomes of patient isolates from the assembled contigs using the -include nudge argument in rgi main. RGI v.5.2.1 was used to summarize the output using rgi heatmap.

We used the Genome Database Taxonomy (GTDB)'s GTDB-Tk software v. 2.4.0 (25) to assign each genome to a domain-specific reference tree and subsequently use each genome's relative evolutionary divergence and average nucleotide identity to determine its taxonomic assignment. We inferred a maximum-likelihood phylogenetic tree using IQ-TREE v.2.2.6 (26) and visualized phylogenetic relationships using iToL v.7 (27). We used the Solu cloud-based platform v.1.0.251 (28) to run MOB-Suite v.3.1.9 (29) for typing, reconstruction, and visualization of predicted plasmids.

Analysis

We performed a retrospective chart review to extract aforementioned patient and microbiological variables of interest. We used conventional epidemiology combined with WGS data to determine if isolates from different patients were spatiotemporally linked. If patients overlapped within one month and/or space (i.e., admitted to same ward) and/or exposure to procedures and their Serratia isolates clustered together on visual inspection of the phylogenetic tree, we considered there to be a possible epidemiological link between patients. We then calculated pairwise average nucleotide identity (ANI) for possibly linked specimens using OrthoANIu (30); we considered an ANI ≥99.0% to be significant. We used descriptive statistics to summarize the characteristics of patients and their Smc isolates. We used command-line tools for genome annotation and R, version 4.4.2 (R Foundation for Statistical Computing, Vienna) to calculate relevant statistics. We

used the packages ggplot2 (31), dplyr (32), diagrammeR (33), and ComplexHeatmap (34) to create data visualizations.

Sequencing data

Whole genome sequences for all Smc isolates included in the study and their linked antimicrobial MIC data are publicly available under NCBI BioProject PRJNA1232785.

Results

We identified 147 microbial isolates as Serratia marcescens by our clinical microbiology laboratory during the study period. No suspected nosocomial Smc outbreaks occurred during the study period at the participating institutions. Of these 147 isolates, 125 unique isolates were included in the analysis (Figure 1). 22 patient isolates were excluded for the following reasons: 20 isolates were duplicates, one isolate failed WGS, and one sequenced isolate was identified as Aeromonas caviae (the initial patient specimen was identified as a mixture of Serratia marcescens and Aeromonas caviae, and the incorrect isolate was frozen for study inclusion). Most included patients were male (n=88/125, 70.4%) and the majority were community-dwelling (n=114/125, 90.5%). Patient demographic data are further described in Table 1. Of note, 49.6% (n=62/125) of the included patients were receiving empiric antimicrobials at the time of specimen collection; however, 32.0% (n=40/125) did not receive a definitive treatment course after cultures had resulted, suggesting that these isolates were assessed by the treating clinician to be indicative of microbial colonization. Seventeen patients died during their hospital admission, with mortality attributable to Smc infection in 64.7% (n=11/17). For all individual patient-level demographic and microbial isolate data that were extracted as part of the study, including MICs for all tested antimicrobials, please see data freely accessible via the Open Science Framework (35).

We identified S. marcescens from various specimen types, including urines, tissues, fluids, swabs, and sputa (Table 2). All (n=125/125) included microbial isolates were identified as S. marcescens by the clinical microbiology laboratory's VITEK® MALDI-TOF MS. AST using the VITEK® 2 instrument demonstrated two major phenotypic patterns of AmpC expression in our clinical Smc isolates: firstly, resistance to amoxicillin-clavulanate and cefoxitin, with susceptibility

to ceftriaxone, suggestive of inducible AmpC expression; secondly, resistance to all three aforementioned agents, suggestive of de-repression. Most isolates (n=109/125, 87.2%) had an inducible AmpC expression, whereas only 9.6% (n=12/125) exhibited a constitutive de-repression phenotype using the definition of Aracil-Gisbert et al (8). No included isolates exhibited a “basal-like” AmpC phenotype with susceptibility to all three agents. The distributions of MIC values and CLSI interpretative breakpoints for all tested antimicrobials are seen in Figures 3 and 4.

We performed whole genome sequencing on all 125 unique clinical isolates. The median genome size of the sequenced Smc isolates was 5.10 Mb (range: 4.34-5.74 Mb) with a median guanine-cytosine content of 59.8% (range: 58.5-60.3%). The median contig number was 36 (range: 17-1501), with a median largest contig length per genome of 760 kb (range: 38-2380) and a median N50 statistic of 317 kb. The average sequencing coverage per base was 197-fold (range: 37-330).

When comparing the species-level identification between the VITEK® MALDI-TOF MS used in the clinical laboratory and those assigned based on the whole genome sequences using GTDB-Tk, we found roughly one-half (n=64/125, 51.2%) to be discordant. Of the discordant identifications, all were found to be members of the Smc, with 19 S. bockelmannii, 4 S. nematodiphila, 19 S. nevei, and 22 S. ureilytica isolates.

The inferred maximum-likelihood phylogenetic tree is shown in Figure 2. While many isolates clustered together phylogenetically (e.g., isolates 43, 57, 130), most did not share a spatiotemporal link. In instances where isolates clustered together phylogenetically and were isolated from patients on the same hospital ward, we did not consider them putatively linked if there was no temporal overlap (e.g., isolates 54 and 93) because the study did not include environmental sampling to help inform a putative link. However, two pairs of isolates appeared to cluster together phylogenetically and overlap spatiotemporally: isolate 50 and 56; and isolate 106 and 113.

For isolates 50 and 56, both of these Serratia nevei isolates were collected within one month of each other in July and August 2024, respectively, on post-admission day one. We performed a

pairwise USEARCH alignment using the OrthoANIu tool, which showed 98.80% ANI over an average aligned length of 3.74 Mb (isolate 50 genome coverage 70.41%, isolate 56 genome coverage 68.64%). This isolate pair did not therefore meet the prespecified ANI cutoff of ≥99.0%. For isolates 106 and 113, both of these Serratia nevei isolates were collected in March 2024 within 10 days of each other, with each isolate collected on post-admission day one. A pairwise USEARCH alignment performed using the OrthoANIu tool showed 99.27% ANI over an average aligned length of 3.77 Mb (isolate 106 genome coverage 75.53%, isolate 113 genome coverage 76.29%). A comparison of the phenotypic AST data for isolates 106 and 113 shows identical phenotypic MICs for all antimicrobials tested, save for cefoxitin (isolate 106 MIC = 16 mg/L, isolate 113 MIC ≤4 mg/L) and cefixime (isolate 106 MIC = 0.5 mg/L, isolate 113 MIC ≤ 0.25 mg/L).
We identified 21 different predicted resistance determinants in the 125 sequenced isolates. The following genes were found in a majority of sequenced isolates: emrR, a fluoroquinolone efflux pump (n=124/125, 99.2%); PBP3, a β-lactamase (n=124/125, 99.2%); adeF, a fluoroquinolone and tetracycline efflux pump (n=123/125, 98.4%); ArnT, a phosphoethanolamine transferase affecting peptide antimicrobials (n=123/125, 98.4%); CRP, a macrolide, fluoroquinolone, and penam antimicrobial efflux pump (n=123/125, 98.4%); FosA8, a fosfomycin thiol transferase (n=123/125, 98.4%); the nitroimidazole efflux pump msbA (n=123/125, 98.4%); AAC(6')-Ic, which produces the aminoglycoside N-acetyltransferase enzyme (n=122/125, 97.6%); KpnF, a multidrug efflux pump (n=123/125, 98.4%); qacG, an antiseptic efflux pump (n=122/125, 97.6%); rsmA, a fluoroquinolone, diaminopyrimidine, and phenicol antibiotic efflux pump (n=122/125, 97.6%); vanG, which confers vancomycin resistance by altering peptidoglycan precursor structure (n=121/125, 96.8%); KpnH, a multidrug efflux pump (n=119/125, 95.2%); SRT-2, a β-lactamase (n=90/125, 72.0%); and tet(41), a tetracycline efflux pump (n=74/125, 59.2%).
A heat-map of CARD RGI predictions is found in Figure 5A for all included isolates. The majority of isolates (n=114, 91.2%) harboured AAC(6')-Ic, which inactivates aminoglycoside antibiotics; however, notably, only one isolate (0.8%) tested resistant to gentamicin and tobramycin phenotypically. Similarly, there was near-universal presence of multiple multidrug efflux pumps

with fluoroquinolone targets; however, only 8 isolates (6.4%) demonstrated phenotypic resistance to ciprofloxacin. Of note, isolates 106 and 113 shared an identical predicted resistome (Figure 5A).

A heat-map of those isolates which had predicted plasmids found in Figure 5B In total, 42 isolates (33.6%) of 125 harboured plasmids; among those with plasmids, the median plasmid number per isolate was 1 (IQR: 1-2). The most common plasmid identified belonged to MOB cluster AB960 and is a non-mobilizable Serratia marcescens plasmid occurring in 20 of 42 isolates (47.6%). We identified four novel plasmids (NCBI Biosample SAMN47405310, SAMN47405316, SAMN47405319, and SAMN47405332), of which three were found to be conjugative. Plasmid maps for the novel predicted plasmids are found on the Open Science Framework (35). Notably, no predicted plasmids harboured AmpC operons.

Discussion

This multi-year study of patients at six Canadian hospitals from 2022-2024 with invasive specimens identified as Serratia marcescens by the clinical microbiology laboratory sheds light on the epidemiology of Smc infections in the hospital context. This study represents the largest known multi-year, multi-hospital survey of Smc populations to use whole genome sequencing in its analysis. We found significant genotypic and phenotypic heterogeneity amongst Smc isolates, corroborating recent findings (7, 8, 36). Our study confirms the presence and coexistence of multiple members of the Smc in hospitalized patients. Our findings have significant implications for the laboratory diagnosis and clinical prevention of infections with Smc organisms.

We found that VITEK® MALDI-TOF MS provides poor species-level discrimination of Smc isolates despite a high (≥95%) certainty of species-level identification in all (n=125, 100%) included isolates, when compared to the WGS taxonomic identification. The VITEK® MALDI-TOF MS United States Food and Drug Administration 501(k)-cleared database, version 3.2, (37) does not contain S. bockelmannii, S. nematodiphila, S. nevei, or S. ureilytica, which explains the misidentification of multiple isolates as Serratia marcescens. This corroborates a recent report by Harch et al. showing the similarly poor correlation of the Bruker MALDI-TOF MS platform's

identification with WGS for Smc (36). The failure of multiple commercial MALDI-TOF MS platforms to accurately call Smc isolates to the species level suggests that, while the clinical relevance of each species within Smc is being further elucidated, laboratory reporting and MALDI-TOF MS databases should identify isolates only to the level of “Serratia marcescens complex”. This finding is likely due to the mass spectrum databases used in commercial MALDI-TOF MS being insufficiently granular: while the number of strains incorporated in these commercial databases are not publicly available, the current 501(k)-cleared bioMérieux database only includes Serratia marcescens amongst the Smc. It may also have significant relevance for infection prevention and control efforts, as it may be necessary to rely on costlier whole genome sequencing or multi-locus sequence typing to characterize Smc isolates in a suspected outbreak rather than the conventional clinical species-level identification provided by MALDI-TOF MS.

Interestingly, we found one pair of isolates over the study period that originated from the same hospital and ward within 10 days of each other, both sampled on post-admission day 1. While there were slight differences in phenotypic AST profiles, we note the isolates had an identical CARD RGI resistome prediction profile. Without concomitant environmental sampling, we are unable to establish a higher-certainty causal link. However, the presence of >99.0% ANI and spatiotemporal linkage suggests that the hospital environment or patient-to-patient transmission may have contributed to, or played a clinically unrecognized role in, these particular infections. While one other isolate pair had spatiotemporal linkage, it did not have a sufficiently similar ANI value to be termed a putative transmission event. No further cases with spatiotemporal linkage or sufficiently similar ANI were identified over the remainder of the study period originating from the same hospital ward, suggesting that in-hospital transmission remains rare overall.

Contrary to the recent findings of Aracil-Gisbert and colleagues, we did not identify a subpopulation of Smc organisms in our study displaying a basal AmpC expression profile with susceptibility to amoxicillin-clavulanate, cefoxitin, and a third-generation cephalosporin (8). Our study population revealed susceptibility profiles consistent with inducible and de-repressed AmpC expression only, though this difference may be explained by our lack of environmental isolates. It

is possible that AmpC expression plays a role in adaptation to non-host environments and altered gene expression may lead to a basal-type susceptibility pattern (8). In line with current Infectious Diseases Society of America clinical guidelines denoting Smc a “low-risk” AmpC-producing Enterobacterales (38), only 9.6% of our isolates had a susceptibility profile consistent with de-repression. Our microbial isolates also lacked any predicted carbapenemases.

Our study has several limitations worth mentioning. Most notably, the methodology used cannot establish a causal link for pathogen spread between patients, particularly since our analysis did not include prospectively collected environmental samples because of its retrospective analytic nature. In addition, our study was powered for infection prevention and control-related outcomes (i.e., transmission); therefore, no reliable conclusions can be made regarding clinical outcomes in this cohort, which would require a larger sample size. The retrospective chart review of patient data is prone to misclassification bias and is susceptible to confounding from unmeasured variables; we attempted to minimize this risk by using a pre-piloted data extraction form and identifying consecutive patients over multiple time periods. Our analyses should thus be interpreted with caution and awareness of the context in which they were produced.

In conclusion, this study represents the largest known multi-year, multi-hospital survey of Smc infections. Our work implies that current MALDI-TOF MS-based identification methods in the clinical laboratory insufficiently reflect the granularity of Smc species constituents, which may have important implications for the investigation of (suspected) hospital outbreaks for infection control practitioners. Over our study period, we identified only one pair of Smc isolates from distinct patients with a spatiotemporal link and a significant average nucleotide identity value. The role of environmental niches within the hospital system should continue to be elaborated upon in future research, to help inform infection prevention and control efforts for Smc.

Data availability statement

Genomes and their linked antimicrobial MIC data are publicly available as of the publication date of this manuscript under NCBI BioProject #PRJNA1232785

(https://www.ncbi.nlm.nih.gov/bioproject/). All extracted patient-level chart data and plasmid maps for novel plasmids identified are available at the Open Science Framework under doi:10.17605/OSF.IO/6NHZ7 (35). Analysis code will be made available upon written request to the corresponding author.

Ethics statement

Institutional ethics board approval for this work was obtained from the Hamilton Integrated Research Ethics Board (Project #13873).

Author contributions (ICJME)

A.S.K.: conceptualization, funding acquisition, methodology, investigation, data curation, formal analysis, project administration, writing – original draft, writing – review and editing.

M.G.S.: methodology, investigation, data curation, formal analysis, resources, software, writing – review and editing.

L.R.: investigation, data curation, writing – review and editing.

D.T.: data curation, formal analysis, software, writing – review and editing.

M.G.: investigation, writing – review and editing.

S.S.: software, writing – review and editing.

A.G.M.: conceptualization, methodology, software, supervision, writing – review and editing.

M.S.: conceptualization, methodology, funding acquisition, supervision, writing – review and editing.

D.M.: conceptualization, methodology, funding acquisition, supervision, writing – review and editing.

All authors warrant the findings presented herein, and consent to the manuscript's submission to the journal. This work has been submitted in partial fulfilment of the requirements for the degree of Master of Science by A.S.K.

Conflicts of Interest

A.S.K. received salary support for this work from a Clinician-Investigator Fellowship administered by Government of Ontario's Ministry of Health, as well as an award from the Research Institute of St. Joe's Hamilton; and received operational funding for this work through a research grant awarded by the Department of Pathology and Molecular Medicine at McMaster University. A.S.K. serves as a member of the Serratia working group of the European Society of Clinical Microbiology and Infectious Diseases Study Group for Epidemiological Markers – Antimicrobial Resistance (ESGEM-AMR), outside of the submitted work.

Funding

This work was supported by operational funding to A.S.K. from the Department of Pathology and Molecular Medicine (Award #PMM 2021-01). Salary support for A.S.K. for this work was provided by the Government of Ontario Ministry of Health Clinician-Investigator fellowship, and the Research Institute of St. Joe's Hamilton. A.G.M. holds McMaster University's inaugural David Braley Chair in Computational Biology, generously supported by the family of the late Mr. David Braley. D.M. holds the Michael G. DeGroote Chair in Infectious Diseases. The authors warrant that none of the funders of this work had a role in the design, data analysis, or the decision to submit this work for publication.

Figure 1 – Participant flow chart and exclusion criteria.

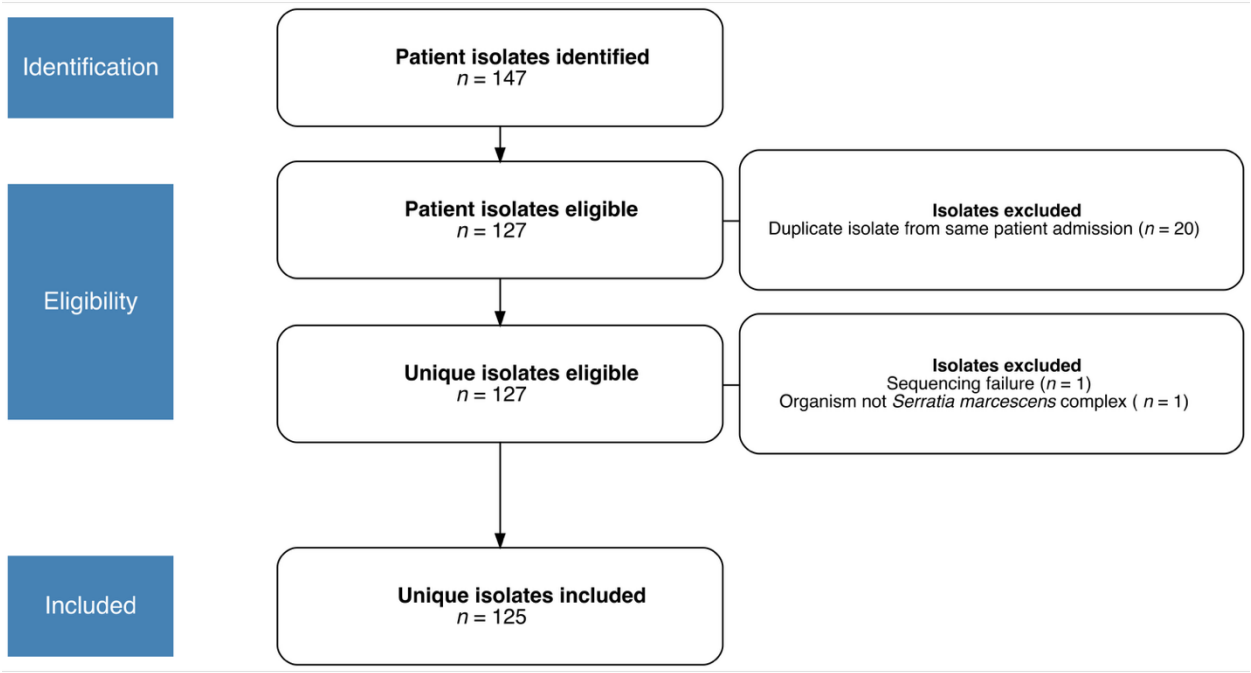
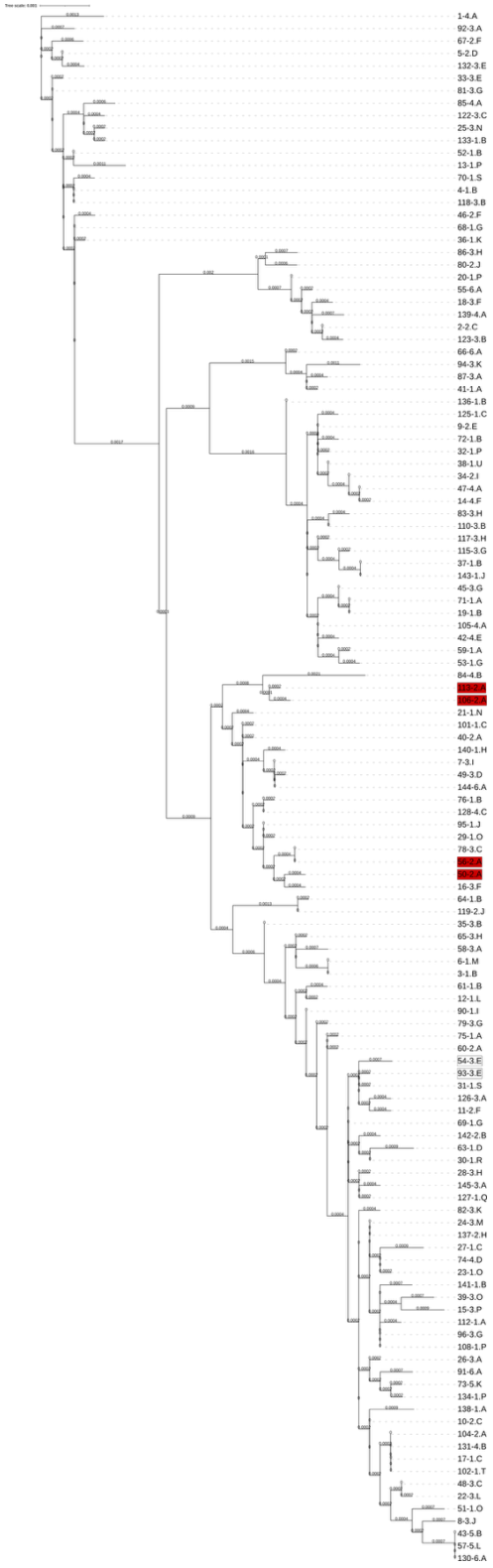


Figure 2 – Phylogenetic tree of *Serratia marcescens* complex clinical isolates



Each node on the phylogenetic tree has been assigned a unique identifier of the format “x-y.z”, where x is the numeric study identification number, y is the anonymized numeric code for the hospital where the specimen originated, and z is the anonymized letter code for the ward where the specimen originated. Adjacent nodes are labelled in red if patients had spatiotemporal overlap. Node lengths are labelled directly on the tree. Tree visualized using iToL, v. 7 (27).

Figure 3 – Distribution of antimicrobial minimum inhibitory concentrations for *Serratia marcescens* complex isolates

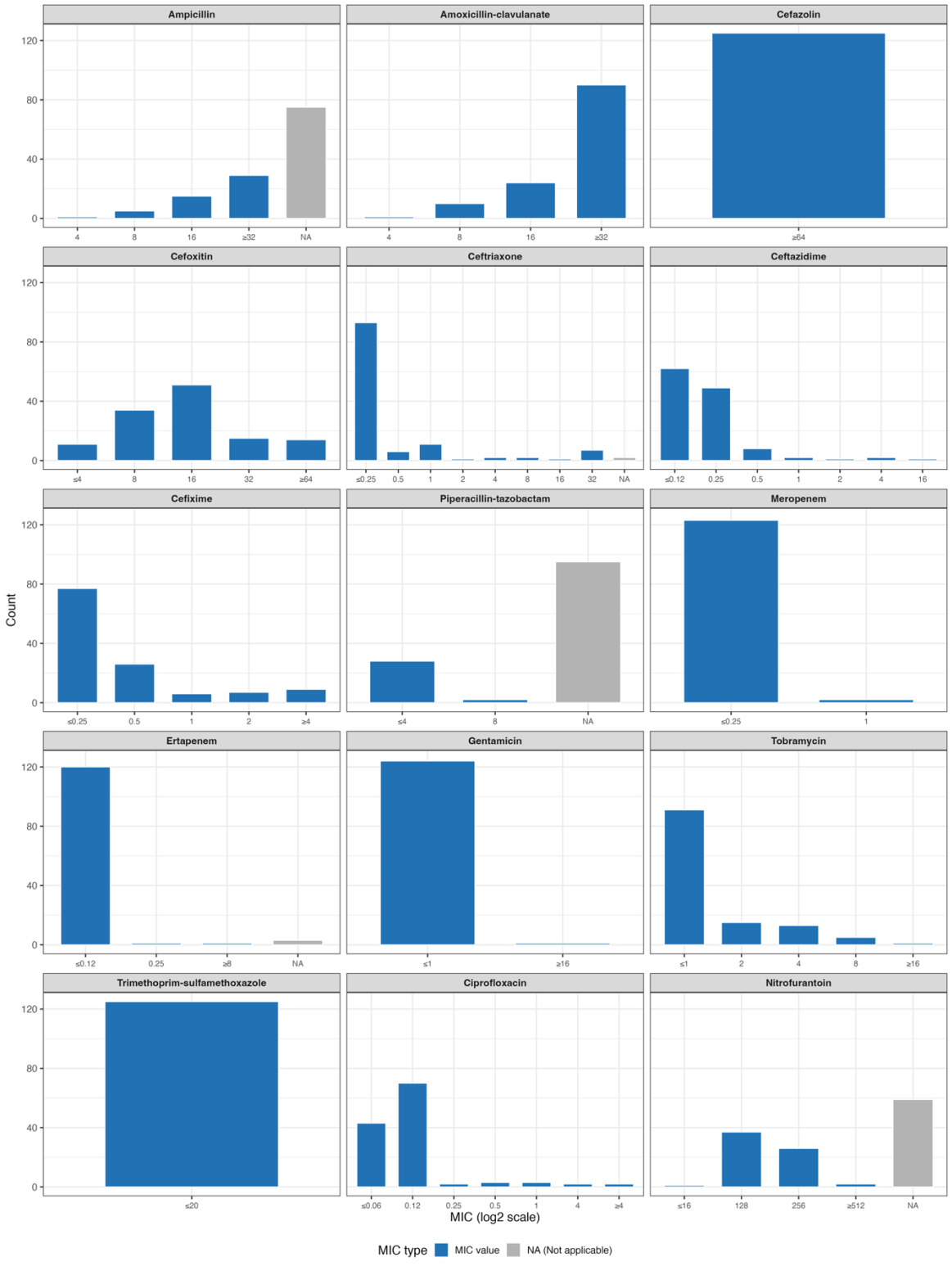
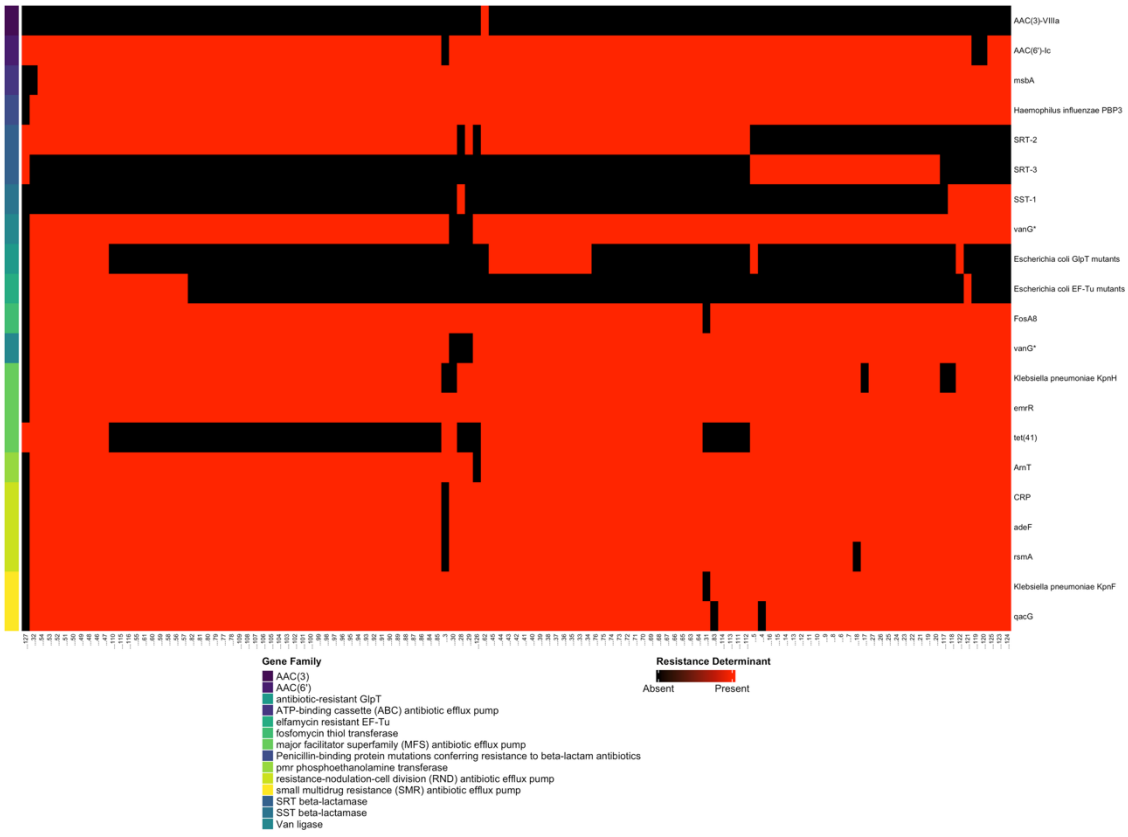


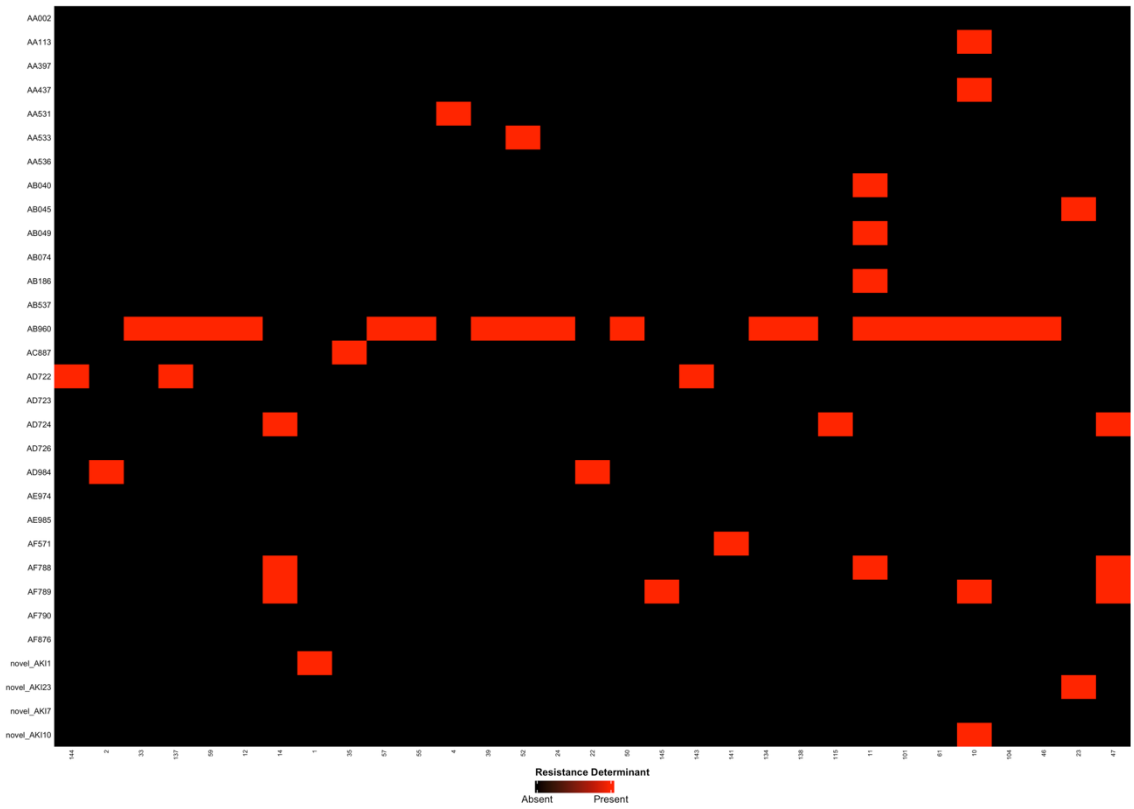
Figure 4 – Distribution of antimicrobial interpretative breakpoints for *Serratia marcescens* complex isolates

Figure 5A – Heat map of Comprehensive Antimicrobial Resistance Database (CARD) Resistance Gene Identifier (RGI) Resistome Prediction



Panel A: Resistance determinants are mapped on the heat map as follows: the lefthand column shows the gene family according to the CARD RGI ontology; the righthand axis labels show the gene names corresponding to the heatmap rows. The binary heat map shows either the presence of a resistance determinant for a given isolate in red, or the absence of said resistance determinant in black. The numerical isolate number is found on the x-axis of the heatmap.

Figure 5B – Heat map of Comprehensive Antimicrobial Resistance Database (CARD) Resistance Gene Identifier (RGI) Resistome Prediction



Panel B: Resistance determinants are mapped on the heat map as follows: the y-axis shows the MOB type of the predicted plasmids; the x-axis labels show the numerical isolate number for those isolates which had predicted plasmids, only. The binary heat map shows either the presence of a resistance determinant for a given isolate in red, or the absence of said resistance determinant in black.

Table 1 – Demographics of patients with *Serratia marcescens* complex infections

Patient characteristic	**All included patients (*n*=125)**	**Patients identified in 2022 (*n*=71)**	**Patients identified in 2024 (*n*=54)**
Median age in years, *n* (IQR)	69.0 (57.0-76.0)	67.0 (59.5-76.0)	69.5 (51.5-78.5)
Female, *n* (%)	37 (29.6)	21 (29.6)	16 (29.6)
Congregate setting, *n* (%)	11 (8.9)	5 (7.0)	6 (11.1)
Median hospitalizations in past year, *n* (IQR)	0 (0-1.0)	1.0 (0-1.0)	0 (0-1.0)
Hospital length of stay at specimen collection, median *n* days (IQR)	1.0 (0-10.0)	1.0 (0-7.0)	2.0 (1.0-20.0)
Receipt of antimicrobials prior to specimen collection, *n* (%)	62 (49.6)	36 (50.7)	26 (48.1)
Receipt of carbapenem as empiric treatment, *n* (%)	5 (4.0)	4 (3.2)	1 (1.8)
Receipt of carbapenem as definitive treatment, *n* (%)	36 (28.8)	17 (23.9)	19 (35.2)
Median days of treatment duration, *n* (IQR)	7.0 (7.0-14.0)	7.0 (7.0-14.0)	7.0 (7.0-10.7)
Patients not treated for positive culture result, *n* (%)	40 (32.0)	16 (22.5)	24 (44.4)
Patient deceased, *n* (%)	17 (13.6)	12 (16.9)	5 (9.3)
Of deceased, *n* (%) with infection as attributable cause of mortality	11 (64.7)	6 (50.0)	5 (100.0)

IQR = interquartile range

Table 2 – Microbiological specimen characteristics for patient isolates with *Serratia marcescens* complex infections

Microbiological characteristic	**All included patients (*n*=125)**
Specimen source, *n* (%)	
Bronchoalveolar lavage fluid	6 (4.8)
Blood	18 (14.4)
Body fluid	8 (6.4)
Sputum	22 (17.6)
Tissue	5 (4.0)
Urine	49 (39.2)
Wound swab	17 (13.6)
Species identification by MALDI, *n* (%)	
Serratia marcescens	125 (100.0)
Discordant species-level identification, MALDI vs. WGS, *n* (%)	64 (51.2)
Species identification by NGS, *n* (%)	
Serratia bockelmannii	19 (15.2)
Serratia marcescens	61 (48.8)
Serratia nematodiphila	4 (3.2)
Serratia nevei	19 (15.2)
Serratia ureilytica	22 (17.6)
Antimicrobial susceptibility profiles, *n* (%)	
Basal AmpC expression[1]	0 (0)
Inducible AmpC expression[2]	109 (87.2)
De-repressed AmpC expression[3]	12 (9.6)
Not applicable[4]	4 (3.2)

MALDI = matrix-assisted laser desorption ionization time-of-flight mass spectrometry; WGS = whole genome sequencing

[1] Defined as testing susceptible to amoxicillin-clavulanate, cefoxitin, and a third-generation cephalosporin using the VITEK® 2 antimicrobial susceptibility testing instrument.

[2] Defined as testing resistant to amoxicillin-clavulanate, cefoxitin, and susceptible a third-generation cephalosporin using the VITEK® 2 antimicrobial susceptibility testing instrument.

[4] In the event that one of the three antimicrobials used to predict AmpC resistance profile was not tested or tested intermediate, a value of “Not applicable” was assigned

References

1. Williams DJ, Grimont PAD, Cazares A, Grimont F, Ageron E, Pettigrew KA, et al. The genus Serratia revisited by genomics. Nat Commun. 2022;13(1):5195.

2. Cristina ML, Sartini M, Spagnolo AM. Serratia marcescens Infections in Neonatal Intensive Care Units (NICUs). Int J Environ Res Public Health. 2019;16(4).

3. Martineau C, Li X, Lalancette C, Perreault T, Fournier E, Tremblay J, et al. Serratia marcescens Outbreak in a Neonatal Intensive Care Unit: New Insights from Next-Generation Sequencing Applications. Journal of Clinical Microbiology. 2018;56(9):10.1128/jcm.00235-18.

4. Fleisch F, Zimmermann-Baer U, Zbinden R, Bischoff G, Arlettaz R, Waldvogel K, et al. Three consecutive outbreaks of Serratia marcescens in a neonatal intensive care unit. Clin Infect Dis. 2002;34(6):767-73.

5. Madani TA, Alsaedi S, James L, Eldeek BS, Jiman-Fatani AA, Alawi MM, et al. Serratia marcescens-contaminated baby shampoo causing an outbreak among newborns at King Abdulaziz University Hospital, Jeddah, Saudi Arabia. J Hosp Infect. 2011;78(1):16-9.

6. Montagnani C, Cocchi P, Lega L, Campana S, Biermann KP, Braggion C, et al. Serratia marcescens outbreak in a neonatal intensive care unit: crucial role of implementing hand hygiene among external consultants. BMC Infect Dis. 2015;15:11.

7. Ono T, Taniguchi I, Nakamura K, Nagano DS, Nishida R, Gotoh Y, et al. Global population structure of the Serratia marcescens complex and identification of hospital-adapted lineages in the complex. Microb Genom. 2022;8(3).

8. Aracil-Gisbert S, Fernández-De-Bobadilla Miguel D, Guerra-Pinto N, Serrano-Calleja S, Pérez-Cobas Ana E, Soriano C, et al. The ICU environment contributes to the endemicity of the “Serratia marcescens complex” in the hospital setting. mBio. 2024;15(5):e03054-23.

9. Tamma PD, Doi Y, Bonomo RA, Johnson JK, Simner PJ. A Primer on AmpC β-Lactamases: Necessary Knowledge for an Increasingly Multidrug-resistant World. Clin Infect Dis. 2019;69(8):1446-55.

10. Jacoby GA. AmpC beta-lactamases. Clin Microbiol Rev. 2009;22(1):161-82, Table of Contents.

11. Philippon A, Arlet G, Labia R, Iorga BI. Class C β-Lactamases: Molecular Characteristics. Clin Microbiol Rev. 2022;35(3):e0015021.

12. Kohlmann R, Bähr T, Gatermann SG. Species-specific mutation rates for ampC derepression in Enterobacterales with chromosomally encoded inducible AmpC β-lactamase. J Antimicrob Chemother. 2018;73(6):1530-6.

13. Vadlamani G, Thomas MD, Patel TR, Donald LJ, Reeve TM, Stetefeld J, et al. The β-lactamase gene regulator AmpR is a tetramer that recognizes and binds the D-Ala-D-Ala motif of its repressor UDP-N-acetylmuramic acid (MurNAc)-pentapeptide. J Biol Chem. 2015;290(5):2630-43.

14. Adler A, Gniadkowski M, Baraniak A, Izdebski R, Fiett J, Hryniewicz W, et al. Transmission dynamics of ESBL-producing Escherichia coli clones in rehabilitation wards at a tertiary care centre. Clinical Microbiology and Infection. 2012;18(12):E497-E505.

15. Hilty M, Betsch BY, Bögli-Stuber K, Heiniger N, Stadler M, Küffer M, et al. Transmission dynamics of extended-spectrum β-lactamase-producing Enterobacteriaceae in the tertiary care hospital and the household setting. Clin Infect Dis. 2012;55(7):967-75.

16. Tschudin-Sutter S, Lucet J-C, Mutters NT, Tacconelli E, Zahar JR, Harbarth S. Contact Precautions for Preventing Nosocomial Transmission of Extended-Spectrum β Lactamase–Producing Escherichia coli: A Point/Counterpoint Review. Clinical Infectious Diseases. 2017;65(2):342-7.

17. Bourdin T, Benoit M, Monnier A, Bédard E, Prévost M, Charron D, et al. Serratia marcescens Colonization in a Neonatal Intensive Care Unit Has Multiple Sources, with Sink Drains as a Major Reservoir. Appl Environ Microbiol. 2023;89(5):e0010523.

18. CLSI. Performance Standards for Antimicrobial Susceptibility Testing. CLSI Supplement M100. 35th. ed. Wayne, PA, USA: Clinical and Laboratory Standards Institute.

19. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20.

20. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol. 2017;13(6):e1005595.

21. Mikheenko A, Saveliev V, Hirsch P, Gurevich A. WebQUAST: online evaluation of genome assemblies. Nucleic Acids Research. 2023;51(W1):W601-W6.

22. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microb Genom. 2021;7(11).

23. Alcock BP, Huynh W, Chalil R, Smith KW, Raphenya AR, Wlodarski MA, et al. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. Nucleic Acids Res. 2023;51(D1):D690-d9.

24. McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, et al. The comprehensive antibiotic resistance database. Antimicrob Agents Chemother. 2013;57(7):3348-57.

25. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics. 2020;36(6):1925-7.

26. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Molecular Biology and Evolution. 2015;32(1):268-74.

27. Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Research. 2024;52(W1):W78-W82.

28. Saratto T, Visuri K, Lehtinen J, Ortega-Sanz I, Steenwyk JL, Sihvonen S. Solu: a cloud platform for real-time genomic pathogen surveillance. BMC Bioinformatics. 2025;26(1):12.

29. Robertson J, Nash JHE. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. Microb Genom. 2018;4(8).

30. Yoon SH, Ha SM, Lim J, Kwon S, Chun J. A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie Van Leeuwenhoek. 2017;110(10):1281-6.

31. Wickham H. ggplot2: Elegant Graphics for Data Analysis: Springer-Verlag New York; 2016.

32. Wickham H, François R, Henry L, Müller K, Vaughan D. dplyr: A Grammar of Data Manipulation. 2025.

33. Ianone R, Roy O. DiagrammeR: Graph/Network Visualization. R package version 1011. 2024.

34. Gu Z. Complex heatmap visualization. Imeta. 2022;1(3):e43.

35. Komorowski AS, Surette MG, McArthur AG, Smieja M, Mertz D. Molecular Epidemiology of Clinical Infections Caused By Serratia marcescens Complex in a Tertiary Care Hospital System: Open Science Framework; 2025 [Available from: doi: 10.17605/OSF.IO/6NHZ7.

36. Harch Susan AJ, Jenkins F, Farhat R, van Hal Sebastiaan J. Complexities in species identification for Serratia marcescens complex for the modern microbiology laboratory. Microbiology Spectrum. 2024;13(1):e01361-24.

37. bioMérieux. VITEK MS. Indications for Use Statement - Form 3881: United States Department of Health and Human Services, Food and Drug Administration; [Available from: https://go.biomerieux.com/us_fda_510_k_organisms_vitek_ms.

38. Tamma PD, Heil EL, Justo JA, Mathers AJ, Satlin MJ, Bonomo RA. Infectious Diseases Society of America 2024 Guidance on the Treatment of Antimicrobial-Resistant Gram-Negative Infections. Clinical Infectious Diseases. 2024:ciae403.