



# *Serratia marcescens* Outbreak in a Neonatal Intensive Care Unit: New Insights from Next-Generation Sequencing Applications

Christine Martineau,<sup>a,b,\*</sup> Xuejing Li,<sup>c</sup> Cindy Lalancette,<sup>a</sup> Thérèse Perreault,<sup>d</sup> Eric Fournier,<sup>a</sup> Julien Tremblay,<sup>e</sup> Milagros Gonzales,<sup>f</sup> Étienne Yergeau,<sup>g</sup> Caroline Quach<sup>b,f,h</sup>

<sup>a</sup>Laboratoire de Santé Publique du Québec, Ste-Anne-de-Bellevue, Québec, Canada

<sup>b</sup>Department of Microbiology, Infectious Diseases, and Immunology, Université de Montréal, Montréal, Québec, Canada

<sup>c</sup>Institute of Parasitology, McGill University, Montréal, Québec, Canada

<sup>d</sup>Department of Paediatrics, McGill University Health Centre, Montréal, Québec, Canada

<sup>e</sup>National Research Council of Canada, Energy, Mining, and Environment, Montréal, Québec, Canada

<sup>f</sup>Research Institute, McGill University Health Centre, Montréal, Québec, Canada

<sup>g</sup>Institut Armand Frappier, Institut National de la Recherche Scientifique, Laval, Québec, Canada

<sup>h</sup>Infection Prevention and Control, CHU Sainte-Justine, Montréal, Québec, Canada

**ABSTRACT** *Serratia marcescens* is an environmental bacterium that is commonly associated with outbreaks in neonatal intensive care units (NICUs). Investigations of *S. marcescens* outbreaks require efficient recovery and typing of clinical and environmental isolates. In this study, we investigated how the use of next-generation sequencing applications, such as bacterial whole-genome sequencing (WGS) and bacterial community profiling, could improve *S. marcescens* outbreak investigations. Phylogenomic links and potential antibiotic resistance genes and plasmids in *S. marcescens* isolates were investigated using WGS, while bacterial communities and relative abundances of *Serratia* in environmental samples were assessed using sequencing of bacterial phylogenetic marker genes (16S rRNA and *gyrB* genes). Typing results obtained using WGS for the 10 *S. marcescens* isolates recovered during a NICU outbreak investigation were highly consistent with those obtained using pulsed-field gel electrophoresis (PFGE), the current standard typing method for this bacterium. WGS also allowed the identification of genes associated with antibiotic resistance in all isolates, while no plasmids were detected. Sequencing of the 16S rRNA and *gyrB* genes both showed greater relative abundances of *Serratia* at environmental sampling sites that were in close contact with infected babies. Much lower relative abundances of *Serratia* were observed following disinfection of a room, indicating that the protocol used was efficient. Variations in the bacterial community composition and structure following room disinfection and among sampling sites were also identified through 16S rRNA gene sequencing. Together, results from this study highlight the potential for next-generation sequencing tools to improve and to facilitate outbreak investigations.

**KEYWORDS** amplicon deep sequencing, bacterial community profiling, NICU, next-generation sequencing, outbreak investigation, *Serratia marcescens*, whole-genome sequencing

Newborns admitted to neonatal intensive care units (NICUs) are at high risk of contracting health care-associated infections (HAIs) due to the immaturity of their immune systems and the intensity of medical interventions required. *Serratia marcescens* is a Gram-negative bacterium that is commonly associated with HAI outbreaks in NICUs. In a recent review of the literature on NICU outbreaks, *S. marcescens* was

Received 12 February 2018 Returned for modification 5 April 2018 Accepted 12 June 2018

Accepted manuscript posted online 13 June 2018

**Citation** Martineau C, Li X, Lalancette C, Perreault T, Fournier E, Tremblay J, Gonzales M, Yergeau É, Quach C. 2018. *Serratia marcescens* outbreak in a neonatal intensive care unit: new insights from next-generation sequencing applications. J Clin Microbiol 56:e00235-18. <https://doi.org/10.1128/JCM.00235-18>.

**Editor** Alexander Mellmann, University Hospital Münster

**Copyright** © 2018 American Society for Microbiology. All Rights Reserved.

Address correspondence to Caroline Quach, [c.quach@umontreal.ca](mailto:c.quach@umontreal.ca).

\* Present address: Christine Martineau, Canadian Forest Service, Natural Resources Canada, Québec, Québec, Canada.

associated with 5 of 39 outbreaks (1), with a broad clinical spectrum, including septicemia, conjunctivitis, pneumonia, urinary tract infections, and meningitis.

*S. marcescens* is ubiquitous in the environment, and previously identified sources of contamination included the hands of health care workers, medical devices, soap, sinks, and milk (2–4). Although environmental sampling occasionally led to the isolation of *S. marcescens*, the source of contamination remained unidentified in most outbreaks (5, 6). In NICUs, *S. marcescens* gastrointestinal carriage was identified as a potential reservoir (5, 7) and was associated with greater prematurity, use of antibiotics, and mechanical ventilation (7).

Previous investigations showed that generally more than one clone can be identified during *S. marcescens* outbreak investigations (2, 5, 8, 9). Typing of bacterial isolates can be particularly challenging and relies on the use of molecular biology tools such as pulsed-field gel electrophoresis (PFGE), multilocus sequence typing (MLST), and ribotyping. PFGE is the more commonly used typing technique and was applied previously to the investigations of *S. marcescens* outbreaks in NICUs (2, 10, 11). Although PFGE is still the gold standard for most bacterial species, whole-genome sequencing (WGS) is a promising and increasingly used tool for strain typing, often providing higher resolution than other tools (12–14).

Other recently developed next-generation sequencing (NGS) tools, such as targeted amplicon sequencing of bacterial phylogenetic marker genes, could improve outbreak investigations, allowing investigators to profile the microbial community composition in environmental samples and improving our understanding of microbial spread and colonization of the hospital environment (15–17). In this study, we describe how the use of advanced molecular biology tools, such as bacterial WGS and sequencing of amplicons from environmental DNA, is challenging but can provide crucial information in the context of a NICU *S. marcescens* outbreak.

## MATERIALS AND METHODS

**Study population and sample collection.** The McGill University Health Centre (MUHC) has a 52-bed level III NICU, with approximately 900 admissions annually from both in-born and referred populations, as it serves as a reference center for other hospitals in the province of Quebec. During the outbreak, specimens from symptomatic newborns were collected and cultured; screening of all babies in the NICU (using nasal and rectal swabs, as well as tracheal samples for intubated or tracheotomized babies) was implemented following the detection of the first 3 cases. Measures for infection prevention and control included contact precautions (patients were all in single rooms) with eye protection during aspiration, discarding of all material and equipment upon discharge of the patient, enhancement of environmental cleaning, and reinforcement of hand hygiene compliance among parents. A disinfection protocol using hydrogen peroxide was implemented. This protocol was applied daily to all rooms but was repeated twice for in-depth cleaning following discharge of a patient.

Environmental sampling was performed upon detection of *S. marcescens* in clinical samples, first on day 27 of the outbreak investigation from 3 rooms with cases (positive rooms) and 1 control room that had been decontaminated prior to sampling following discharge of a case. Environmental samples were collected from different sites in each room (around the bed [outside the isolette] and from the drain, table cart, ventilator/infusion pump, wipe bottle/disinfectant pump, bench, and computer keyboard [outside the room]). Two swabs were collected from each sampling site, one for bacterial culture and the other for DNA extraction. Both swabs were moistened with saline and streaked on 25-cm<sup>2</sup> surfaces (or the entire surface, for <25-cm<sup>2</sup> surfaces). Swabs for cultures were placed in Amies medium, while swabs for DNA extraction were placed in universal transport medium (Copan Diagnostics, Murrieta, CA, USA), and all swabs were frozen at –80°C until tested. The same 3 positive rooms were sampled again on day 59 of the outbreak investigation, along with 2 additional rooms hosting newly infected babies; a single swab per site was collected for bacterial culture.

**Strain isolation and identification.** Samples collected from patients and the environment were inoculated on culture media (Columbia blood agar and MacConkey agar) and incubated at 37°C for 24 h. Isolates were identified with the Vitek 2 system (bioMérieux, Marcy l’Etoile, France) by the hospital laboratory and submitted to the provincial public health laboratory (Laboratoire de Santé Publique du Québec [LSPQ]) for typing.

**Strain typing by PFGE.** Colonies from 24-h growth on blood agar plates were harvested in cellular suspension buffer (100 mM Tris [pH 8.0], 100 mM EDTA [pH 8.0]) with proteinase K (0.1 mg/ml). SeaKem Gold agarose (Lonza, Basel, Switzerland) was added, and the final mixture was transferred into a well of the plug mold. The solidified plug was transferred in lysis buffer (50 mM Tris [pH 8.0], 50 mM EDTA [pH 8.0], 1% Sarkosyl) with proteinase K and incubated at 55°C for 1 h, with agitation. The plug was rinsed and sliced, and the DNA was digested with 150  $\mu$ l of 40 U of *SpeI* solution (Roche, Canada) and incubated at 50°C for a minimum of 2 h. The digested slice was loaded onto a 1% SeaKem Gold PFGE agarose gel.

Electrophoresis was performed in Tris-borate-EDTA (TBE) buffer at 14°C for 19.5 h at 6 V/cm, using switch times of 5.3 to 42 s. Gels were stained in GelRed (Cedarlane, Canada). The Spel patterns were interpreted using the standards described by Tenover et al. (18). Band position tolerances of 0.8% and optimization values of 1% were used for all analyses. Similarity coefficients were obtained within BioNumerics v6.5 by calculating Dice coefficients. Cluster analysis was performed with the unweighted pair group method with arithmetic averages (UPGMA), using 80% similarity to assign a cluster. PFGE patterns were considered similar visually according to the criteria described by Tenover et al. (18). Letters identified the different PFGE patterns.

**Whole-genome sequencing.** Genomic DNA for WGS of bacterial isolates was extracted using the MagAttract DNA Mini M48 kit on the BioRobot M48 automated platform (Qiagen, Toronto, ON, Canada) and was quantified using the Qubit double-stranded DNA (dsDNA) high-sensitivity assay (Thermo Fisher Scientific, Mississauga, ON, Canada). Sequencing libraries were prepared with the Nextera XT library preparation kit (Illumina, Inc., San Diego, CA, USA), following the manufacturer's instructions, and sequenced on the Illumina MiSeq platform, using MiSeq v3 reagent kits (600 cycles, paired-end reads). The quality of the raw sequence reads was verified using FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>).

**Strain typing by hqSNV.** Phylogenomic links between isolates were investigated using a core genome high-quality single-nucleotide variant (hqSNV) analysis pipeline (19) implemented in a Galaxy workflow (20). Briefly, reads were mapped to a reference genome (GenBank accession number CP018923.1) using Smalt (<http://www.sanger.ac.uk/science/tools/smalt-0>). hqSNV sites were identified with FreeBayes (<https://github.com/ekg/freebayes>) and SAMtools/BCFtools (21, 22). A filtering step was then applied to mask regions with single-nucleotide variants (SNVs) occurring at densities of 2 or more within a moving window of 20 bp (default settings) or 500 bp (higher stringency) (19). The remaining hqSNV sites were concatenated to produce a pseudo-multiple-sequence alignment. This alignment was finally used as the input to infer a maximum likelihood tree with PhyML (23) and to generate other output files needed to construct a minimum spanning tree with PhyloViz (24).

**Genome assembly and annotation.** SPAdes v3.7.0 (25) was used with default settings to assemble the reads into contigs. The resulting assemblies were filtered with an in-house Python script in which contigs with coverage of  $>5\times$  and lengths of  $>500$  bp were kept. Genome annotation was added by the NCBI prokaryotic annotation pipeline (v4.2 [May 2017]) following sequence deposition in GenBank. Antibiotic resistance genes were further investigated using the Resistance Gene Identifier (RGI) of the Comprehensive Antibiotic Resistance Database (CARD) (26). The presence of plasmids was detected using PlasmidFinder, with the threshold set at 90% identity (27).

**Sequencing of bacterial 16S rRNA and *gyrB* genes.** Genomic DNA extraction of environmental swab samples was performed using the NucliSENS EasyMAG automated platform, with input and elution volumes of 500  $\mu$ l and 100  $\mu$ l, respectively. A negative control (diethyl pyrocarbonate [DEPC]-treated water) was included in each extraction run and processed with the samples. A PCR negative control (i.e., DEPC-treated water added to the PCR mixture) was also included to PCR runs. Primers 343F (5'-TACG GRAGGCAGCAG-3') and 803R (5'-CTACCAGGTATCTAATCC-3') were used for PCR amplification of the V3 to V4 regions of the 16S rRNA genes of bacteria (not archaea). The use of these primers generated ~460-bp amplicons that allowed the identification of bacteria at the genus level. A second set of primers, *gyrB*\_aF64 (5'-MGNCCNGSNATGTAYATHGG-3') and *gyrB*\_aR353 (5'-ACNCCRTGNARDCCDCNGA-3') (28), was used to amplify an ~290-bp portion of the bacterial *gyrB* gene, which was previously described as a genetic marker that allowed the identification at the species level of bacteria of the *Enterobacteriaceae* family, such as *Serratia marcescens* (28–30). PCR amplification with the HotStarTaq DNA polymerase (Qiagen) was carried out for 32 cycles (16S) or 40 cycles (*gyrB*), at annealing temperatures of 55°C (16S) or 60°C (*gyrB*). A mock bacterial community composed of a mixture of genomic DNA from 20 bacterial strains containing different rRNA operon counts (1,000 to 1,000,000 operons per organism per  $\mu$ l) (BEI783D; BEI Resources, Manassas, VA, USA) was used as a positive control. Library preparation for Illumina sequencing was performed according to the manufacturer's instructions (31). Sequencing was performed on the Illumina MiSeq platform using MiSeq v3 reagent kits (600 cycles, paired-end reads).

**16S rRNA and *gyrB* gene sequence processing.** The 16S rRNA gene sequences were analyzed through a custom rRNA short amplicon analysis pipeline, as described previously (32). Briefly, reads were filtered, merged with their overlapping paired ends, and clustered at 97% identity. Taxonomy was then assigned to each cluster based on the SILVA taxonomy (33), and an operational taxonomic unit (OTU) table was generated, filtered to exclude eukaryotes and chloroplasts, and normalized as described previously (34). This normalized OTU table was used for downstream analysis and for computing of beta diversity metrics. Taxonomy tables were also generated for both the *gyrB* and 16S rRNA gene sequences using MiniKraken (35), a tool that was initially designed for the taxonomic classification of shotgun metagenomic sequence reads that can be applied to amplicon sequencing data (36, 37). In this case, relative abundances were calculated by dividing the number of sequences belonging to a taxonomic group by the total number of bacterial sequences within each sample.

**Statistical analyses.** All statistical analyses were performed in R v3.4.2 (The R Foundation for Statistical Computing). Data normality was tested prior to analysis of variance (ANOVA) using the Shapiro.test function. Data were transformed (transformTukey function, with the rcompanion library) when necessary to meet parametric ANOVA assumptions. ANOVA and the *post hoc* Tukey honestly significant difference (HSD) test were then carried out using the aov and TukeyHSD functions, respectively. Spearman linear correlation analysis was performed using the cor.test function. Principal-coordinate analysis (PCoA) was carried out on weighted UniFrac matrices (38) using the pcoa function of

**TABLE 1** Characteristics of the *S. marcescens* isolates recovered in this study

Isolate no.	Patient or environment no.	Outbreak investigation day	Collection site	Room no.	PFGE pattern
ID148137	Patient 1	1	Blood	2	A
NA <sup>a</sup>	Patient 4	1	Eye	3	NA
ID148138	Patient 2	17	Blood	4 (control room) <sup>b</sup>	A
ID147729	Patient 3	21	Tracheal sputum	1	B
ID147728	Patient 4	24	Endotracheal aspirate	3	A
ID148299	Environment 1	27	Around bed	2	A
NA	Environment 2	27	Drain	3	NA
ID147991	Patient 5	31	Eye	5	C
ID148587	Patient 6	52	Surgical wound (abdomen)	6	D
ID148696	Environment 3	59	Drain	5	E
ID149856	Patient 7	101	Pus (rectal swab)	7	E
ID149855	Patient 7	101	Nasal swab	7	E

<sup>a</sup>NA, not applicable; the isolate was not sent to the public health laboratory for further characterization.

<sup>b</sup>Room 4 was decontaminated twice with H<sub>2</sub>O<sub>2</sub> following discharge of the patient and was used as a control room during the first environmental sampling.

the ape package (39). Effects of the sampling site on the community structure were also tested using permutational multivariate ANOVA (PERMANOVA) (Adonis function, with the vegan library).

**Accession number(s).** Raw sequence data produced from *S. marcescens* isolates were deposited in NCBI under BioProject accession number [PRJNA430250](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA430250) and Sequence Read Archive (SRA) accession number [SRP129857](https://www.ncbi.nlm.nih.gov/sra/SRP129857). Annotated genomes are available under GenBank accession numbers [PQOB000000000](https://www.ncbi.nlm.nih.gov/genbank/PQOB000000000) to [PQOK000000000](https://www.ncbi.nlm.nih.gov/genbank/PQOK000000000). Raw sequence data produced from environmental swab samples were deposited in NCBI under BioProject accession number [PRJNA432186](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA432186) and SRA accession number [SRP131754](https://www.ncbi.nlm.nih.gov/sra/SRP131754).

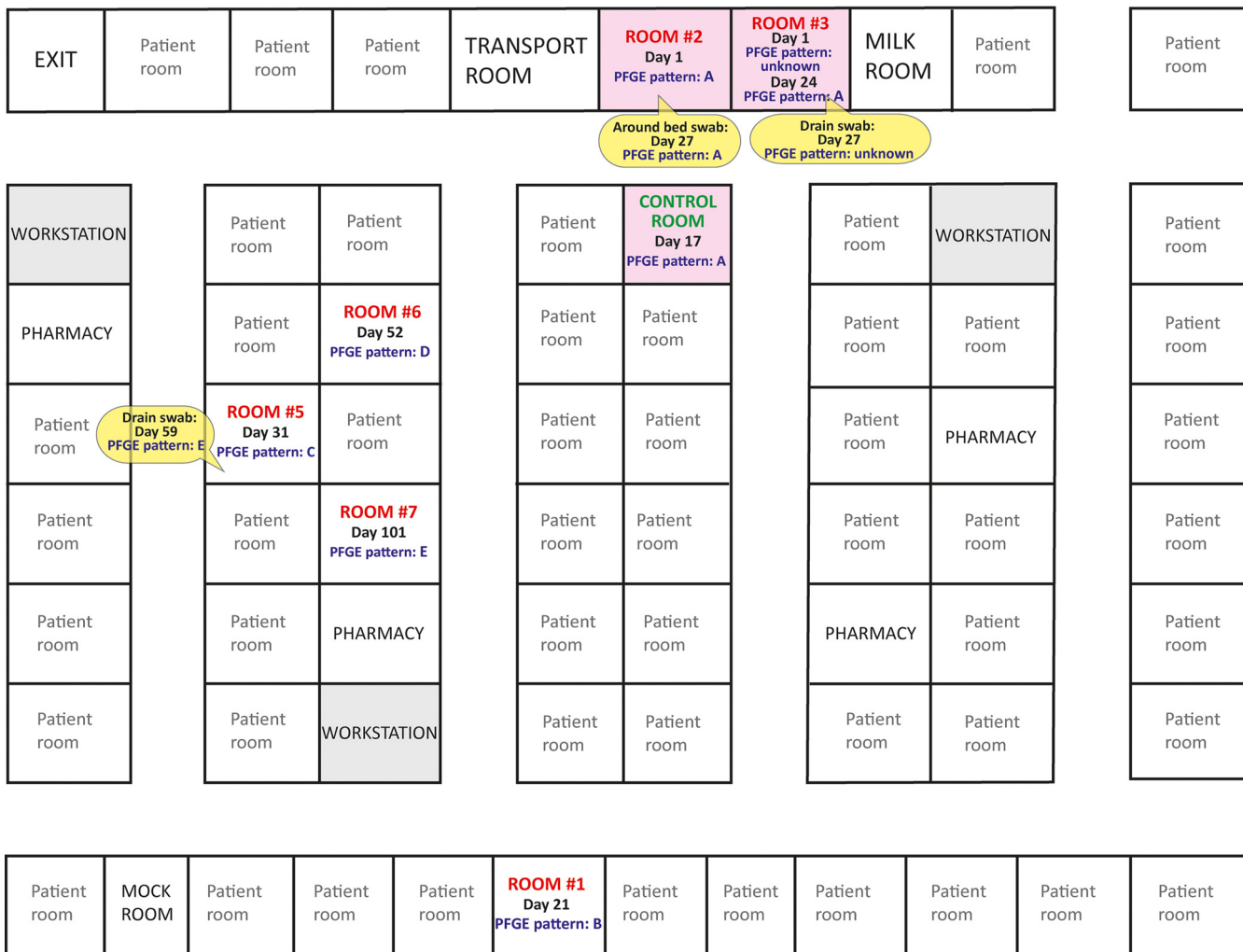
## RESULTS

**Patient and strain characteristics.** Twelve isolates were collected during the study period, which lasted 101 days in 2016, with 9 isolates being collected from 7 patients and 3 isolates being collected from the environment (2 on day 27 and 1 on day 59) (Table 1). Isolates were recovered from various types of samples, including sputum, eye, blood, surgical wound, rectal swab, nasal swab, and endotracheal aspirate samples. All patients stayed in different rooms (Fig. 1). Two patients were female. The median chronological age at the onset of first infection was 45 days (interquartile range [IQR], 14 to 74 days); corrected gestational age and corrected age data can be found in Table S1 in the supplemental material. One patient had a birth weight equal to 750 g, 4 had birth weights between 751 and 1,000 g, and 2 had birth weights of more than 2,500 g. The median length of NICU stay was 117 days (IQR, 114 to 279 days). One patient died at 10 days of life; death was considered to be directly related to the infection.

**PFGE results.** Two isolates (1 from patient 4 and 1 from the room 3 drain) were not sent to the reference laboratory and could not be included in further analyses. PFGE analysis of the 10 remaining isolates identified 5 different banding patterns (patterns A to E) (Table 1). Four isolates shared the same PFGE pattern (pattern A); 3 isolates were from 3 patients and 1 was from the environment (around the bed) in the room of a patient with banding pattern A. All isolates with banding pattern A were recovered from closely located rooms (Fig. 1). PFGE banding pattern E was identified for 2 isolates collected from the same patient and 1 isolate collected from a nearby room (drain). The remaining 3 PFGE banding patterns were represented by single patient isolates.

**hqSNV results.** Analysis of the 43,628 or 55 hqSNV positions obtained using density filtering cutoff values of 2 SNVs/20 bp and 2 SNVs/500 bp, respectively, showed that the 4 isolates associated with PFGE pattern A had 0 hqSNVs between them (Fig. 2). The 2 isolates from the same patient (PFGE pattern E) also had 0 hqSNVs, and both isolates were separated by 5 hqSNVs (cutoff value of 2 SNVs/20 bp) or 0 hqSNVs (cutoff value of 2 SNVs/500 bp) from the environmental isolate sharing the same PFGE banding pattern. All of the other isolates were more distantly related to each other, with the closest isolates at 16,550 to 18,155 hqSNV positions with a density filtering cutoff value of 2 SNVs/20 bp and 13 to 16 hqSNV positions with a density filtering cutoff value of 2 SNVs/500 bp.

**Plasmid and resistance genes.** Following sequence assembly with SPAdes, the number of contigs obtained for each isolate varied between 34 and 84, with average

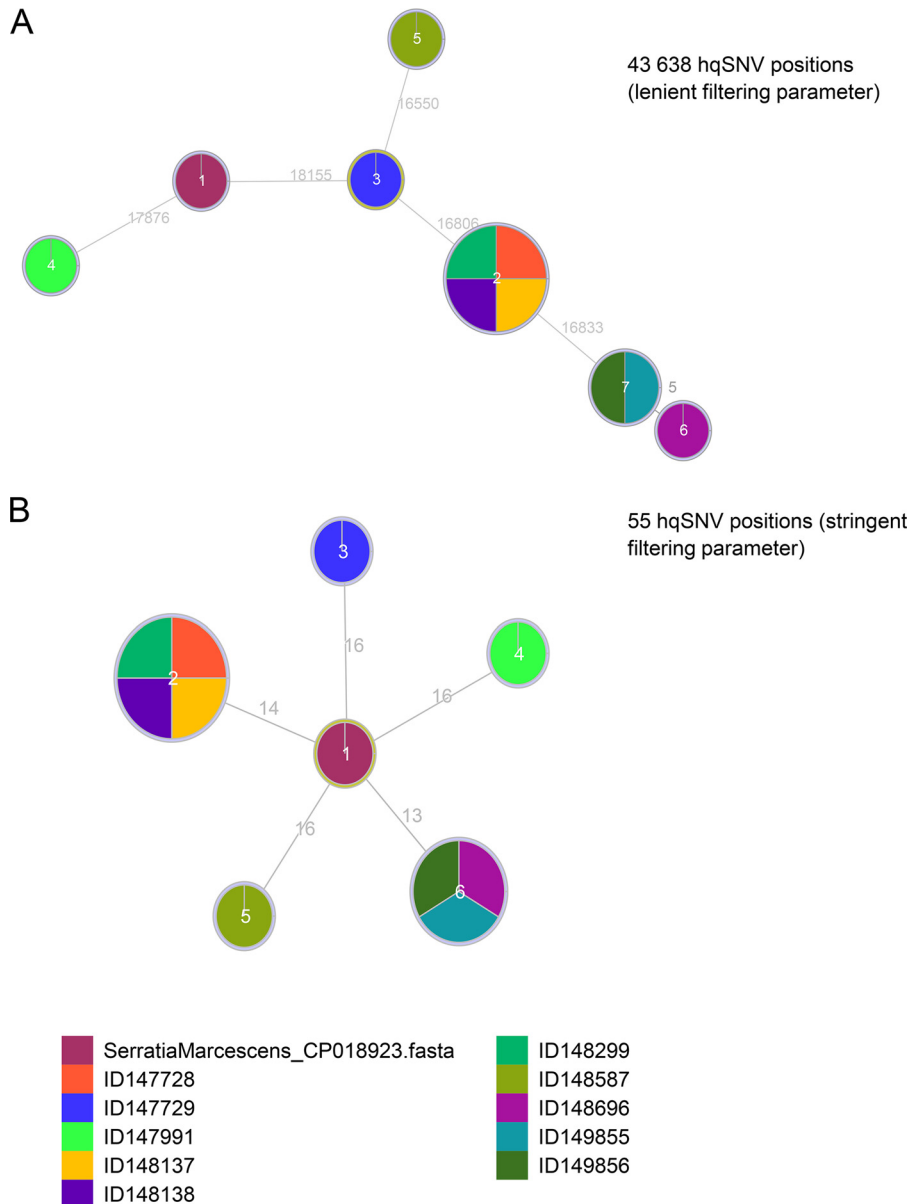


**FIG 1** Map of the NICU. Rooms in which *S. marcescens* isolates were recovered from the environment and/or from patients are identified in red font, along with the date of infection onset and the PFGE results. The control room (identified in green font) was decontaminated twice with hydrogen peroxide prior to sampling, following discharge of a case. Pink shading identifies rooms in which the dominant clone of the outbreak was identified.

coverage of 38× to 52×. Contigs of the 10 bacterial isolates were screened for plasmids and antibiotic resistance genes (Table 2). Four genes associated with antibiotic resistance were identified in all isolates, with percentages of identity to sequences in the CARD database varying between 90.95% and 100%. No plasmids were identified.

**Relative abundance of *Serratia* in environmental samples.** The relative abundance of *Serratia* in environmental samples varied between 0.0% and 25.6%, depending on the target gene (16S or *gyrB*), the method used for taxonomic assignment (custom pipeline or Kraken), and the sampling site (Fig. 3). The relative abundance of *Serratia* in the extraction and PCR negative controls varied between 0.0% and 0.25% (Fig. 3). Correlation analyses of *Serratia* (for 16S) or *S. marcescens* (for *gyrB*) relative abundances showed overall good agreement between the methods, with significant correlations between all data sets being identified (*gyrB* Kraken versus 16S Kraken, Spearman  $\rho = 0.68, P < 0.001$ ; *gyrB* Kraken versus 16S custom pipeline, Spearman  $\rho = 0.64, P < 0.001$ ; 16S Kraken versus 16S custom pipeline, Spearman  $\rho = 0.94, P < 0.001$ ).

The relative abundance of *Serratia* was higher in rooms with a patient positive for *Serratia* than in the control room, although the statistical significance of these differences could not be tested because of the lack of a replicate for the control room. Within positive rooms, however, the sampling site had a significant effect on the relative abundance of *Serratia* sequences (*gyrB* Kraken,  $P = 0.006$ ; 16S Kraken,  $P = 0.048$ ; 16S



**FIG 2** Graphic representation of the minimum spanning trees resulting from WGS typing of the 10 isolates recovered during the outbreak investigation, using core genome hqSNV analysis with density cutoff values of 2 SNVs/20 bp (A) or 2 SNVs/500 bp (B). The numbers on the branches denote the numbers of hqSNV differences. Isolates ID149855 and ID149856 are from the same patient.

custom pipeline,  $P = 0.0046$ ), with higher values around beds and on ventilator/infusion pumps and lower values in drains and on computer keyboards located outside the rooms (Fig. 3).

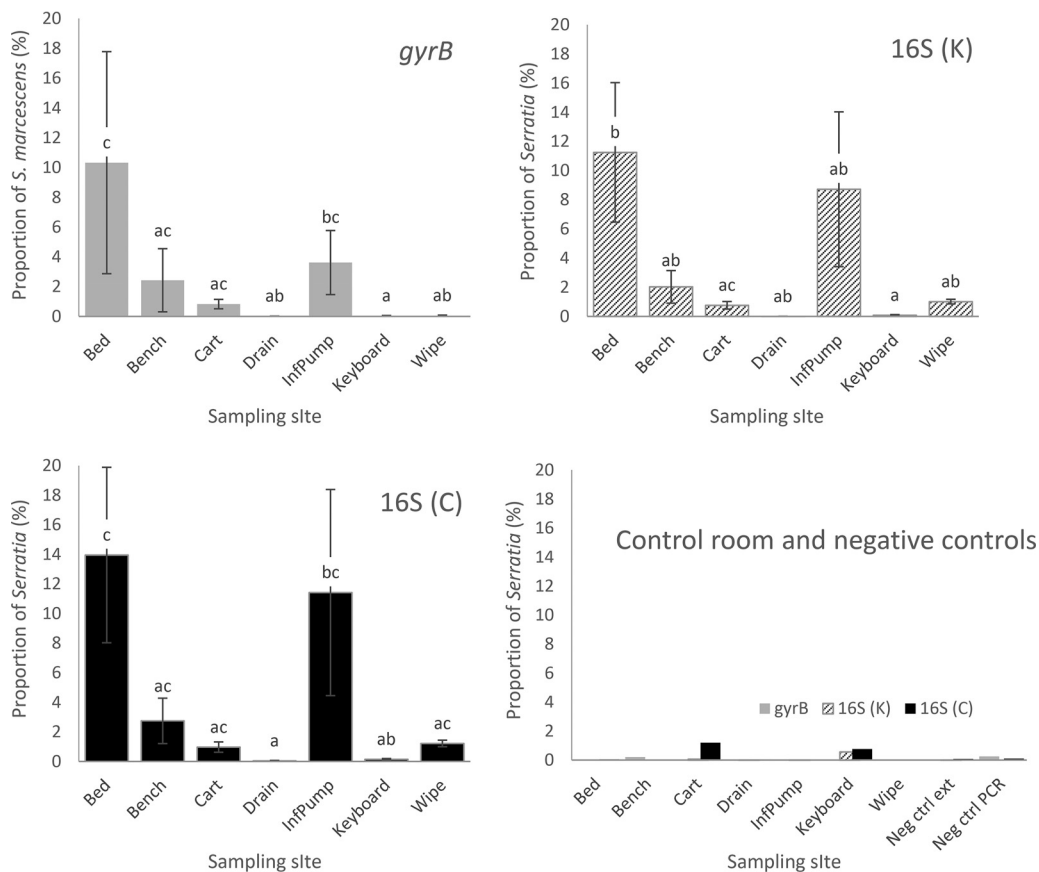
**Bacterial community composition and structure.** The accuracy of the three bacterial amplicon sequencing methods (i.e., *gyrB* analyzed with Kraken, 16S analyzed with Kraken, and 16S analyzed with the custom pipeline) for taxonomic assignment at the genus level was assessed by comparing the results obtained with each method to theoretical results for a mock community DNA mixture composed of 20 bacterial species from 17 genera (see Fig. S1 and S2 in the supplemental material). The 16S rRNA gene sequencing combined with the custom pipeline was identified as the most accurate method and was selected for further bacterial community analyses (Fig. S1 and S2).

**TABLE 2** Genes associated with antibiotic resistance identified in genomes of *S. marcescens* isolates

Gene product	ARO accession no. <sup>a</sup>	Description	Sequence identity (%)															
			ID148137	ID148138	ID147729	ID147728	ID148299	ID147991	ID148587	ID148696	ID149856	ID149855						
AAC(6')-Ic	3002549	Aminoglycoside resistance	98.63	98.63	97.95	98.63	98.63	98.63	99.32	100.0	98.63	98.63	98.63	98.63	98.63	98.63		
CRP <sup>b</sup>	3000518	Indirect inhibitor of antimicrobial secondary metabolites	99.05	99.05	99.05	99.05	99.05	99.05	99.05	99.05	99.05	99.05	99.05	99.05	99.05	99.05		
<i>E. coli</i> CpxR	3004055	Stress response regulator	90.95	90.95	90.95	90.95	90.95	90.95	90.95	90.95	90.95	90.95	90.95	90.95	90.95	90.95		
SRT-2	3002494, 3002493	AmpC $\beta$ -lactamase	98.69	98.69	99.21	98.69	98.69	99.47	99.74	98.69	99.74	99.74	99.74	99.74	99.74	99.74		

<sup>a</sup>From CARD (26).

<sup>b</sup>CRP, cAMP receptor protein.



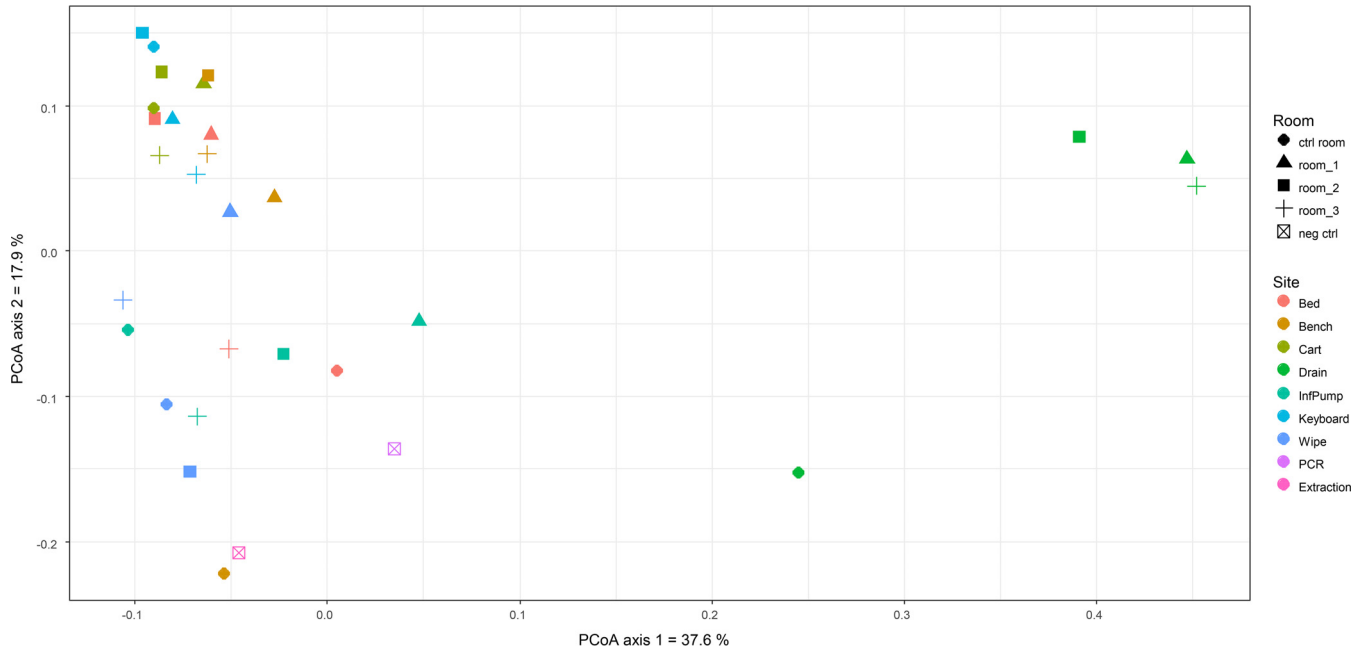
**FIG 3** Relative abundance of *Serratia* in environmental samples from positive rooms of a NICU *S. marcescens* outbreak, as determined by different targeted amplicon sequencing protocols. *gyrB*, *gyrB* gene analyzed with Kraken; 16S (K), 16S rRNA gene analyzed with Kraken; 16S (C), 16S analyzed with the custom pipeline. Error bars represent the standard error of the mean. Different letters within a series indicate significant differences according to Tukey's test ( $P < 0.05$ ). InfPump, infusion pump; Neg ctrl ext, negative control for DNA extraction; Neg ctrl PCR, negative control for PCR.

The bacterial community composition and structure of the control room were clearly distinct from those of the positive rooms for several sampling sites (around the bed, the table cart, and the drain) (Fig. 4 and 5), although these observations could not be statistically confirmed. In positive rooms, PERMANOVA showed that the bacterial community structures were significantly different among environmental sampling sites ( $P < 0.001$ ). Drain community structures were highly distinct, compared to other sampling sites, as shown in the PCoA ordination plot (Fig. 4). These samples were characterized by large proportions of sequences related to bacteria of the genera *Methylobacterium* and *Elizabethkingia* (Fig. 5). In other samples, *Streptococcus* and *Staphylococcus* were generally dominant (Fig. 5). Together with *Serratia*, other bacteria commonly associated with outbreaks in NICUs, such as *Enterobacter*, *Neisseria*, and *Pseudomonas*, were among the top 20 most abundant bacterial genera detected throughout the samples (Fig. 5). Environmental bacteria with no known clinical impact (e.g., *Tumebacillus*, *Acidovorax*, and *Aquabacterium*) were also identified at high relative abundances in some samples, as well as in the extraction negative control, indicating that these sequences might be from contaminants. The PCR negative control had a more distinct bacterial community profile (Fig. 4 and 5), although some bacterial taxa detected (*Haemophilus* and *Acinetobacter*) were also found in swab samples.

## DISCUSSION

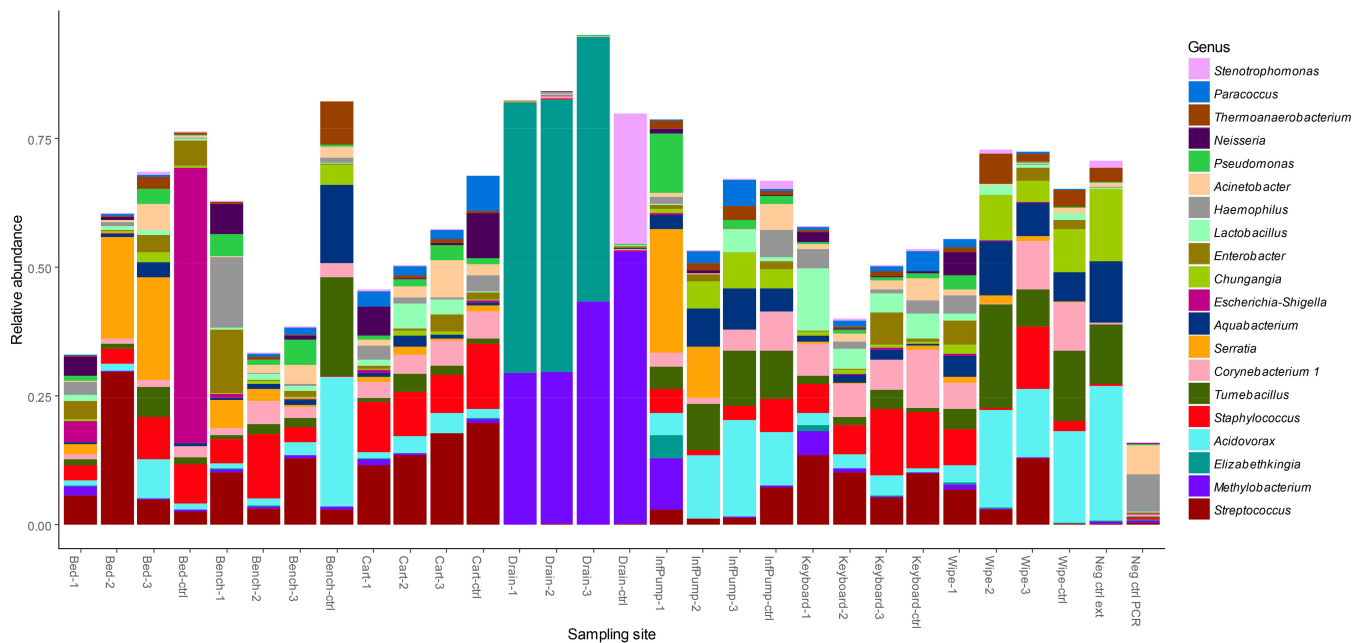
NGS tools have great potential for use in clinical microbiology and outbreak investigations (40, 41). Here, WGS of bacteria and amplicon sequencing of environmental samples were applied to the characterization of a NICU *S. marcescens* outbreak.





**FIG 4** Results from the PCoA of weighted UniFrac distances between swab samples, calculated from 16S rRNA gene sequencing data. Each point represents a single swab sample, and the distances between points illustrate how the bacterial community differed from one sample to another. Colors correspond to the sampling sites and shapes represent the different rooms, as indicated. InfPump, infusion pump; neg ctrl, negative controls.

WGS was mainly used for strain typing, using a core genome hqSNV bioinformatics approach (19). hqSNV was used previously for typing of other members of the *Enterobacteriaceae* family (12, 42, 43), as well as other groups of bacteria (13, 14, 44), but this is the first report, to our knowledge, of its use for the investigation of a *S. marcescens* outbreak. Good correlation was found between PFGE and hqSNV results, with between 0 and 5 SNV positions being identified for isolates sharing the same PFGE profile; higher



**FIG 5** Relative abundances of the 20 most abundant bacterial genera throughout the samples, as assessed by 16S rRNA gene sequencing. Numbers following the name of the sampling site indicate the room number. ctrl, control room; InfPump, infusion pump; Neg ctrl ext, negative control for extraction; Neg ctrl PCR, negative control for PCR.

numbers of SNVs were identified for isolates with distinct PFGE patterns. Previous studies found that the numbers of SNVs among isolates from the same outbreak can vary depending of the bacterial taxon analyzed. Using the same bioinformatics approach, the numbers of SNVs identified among isolates from the same outbreak varied from 0 to 4 for *Salmonella enterica* serovar Heidelberg (12), from 0 to 3 for *Salmonella enterica* serovar Enteritidis (45), from 0 to 5 for *Escherichia coli* O157 (46), and from 0 to 2 for *Vibrio cholerae* (47). As reported here, the use of a SNV density filtering step at different cutoff values has a strong effect on the number of hqSNVs identified among clusters (19). Although the supporting epidemiological information and the number of isolates tested were not sufficient to define the SNV density and the numbers of core hqSNVs that should be used as cutoff values to identify isolates belonging to an outbreak, the results showed the clear potential for using WGS for *S. marcescens* typing.

Both PFGE and core hqSNV identified multiple clones during this NICU *S. marcescens* outbreak investigation, a phenomenon reported previously (2, 5, 8). Strains isolated from patients belonged to 5 different PFGE banding patterns and hqSNV clusters, but only 1 of those was identified in more than 1 patient (PFGE pattern A). Similarly, implementation of *S. marcescens* screening procedures in the NICU of a German hospital led to the identification of isolates corresponding to 8 different PFGE patterns over a 3-month period, with 6 of the patterns occurring in the same unit and 1 dominant strain leading to multiple cases (8). These results indicate that multiple sources could be involved in *S. marcescens* contamination of hospitalized neonates and that screening procedures implemented during outbreak investigations might have detected *S. marcescens* strains that would have gone undetected under normal conditions.

Environmental sampling led to the isolation of 2 strains from drains and 1 strain from around a bed. Previous studies identified *S. marcescens* in drains (48–51), although in some cases there was no link between the isolates and the outbreak under investigation (48, 49). Drains might have been one of the contamination sources here; 1 drain isolate had a matching PFGE banding pattern and a distance of 0 to 5 hqSNVs in comparison with 2 isolates from the same patient. The other drain isolate was collected from the room of a patient infected with the dominant clone of this outbreak, but unfortunately the isolate was lost prior to typing. It is impossible to know which event came first, i.e., the drain microbiota colonizing the patient or the patient microbiota colonizing the environment.

While allowing for strain typing, the use of WGS in outbreak investigations can provide extended information about outbreak isolates, such as the presence of antibiotic resistance genes and plasmids. In this case, no plasmid was detected but 4 chromosomal genes associated with antibiotic resistance were identified in all strains, 2 of which are known to be linked to aminoglycoside and  $\beta$ -lactam resistance in *S. marcescens* (52, 53). Although determining the presence of antibiotic resistance genes was not clinically relevant and could not be related to phenotypes, due to the absence of antibiotic susceptibility testing data, these results showed the potential of WGS for the investigation of resistant strains of *S. marcescens*, as reported for other bacteria (reviewed in reference 54).

Metagenomic sequencing will become essential for outbreak investigations as culture-based methods are increasingly replaced by molecular assays for diagnostic purposes in clinical microbiology laboratories (55). Here, we used a high-throughput targeted amplicon sequencing approach in order to characterize the bacterial communities and to evaluate the relative abundance of *Serratia* in swabs collected from different sampling sites in positive and negative rooms during a NICU *S. marcescens* outbreak. In contrast to shotgun metagenomics, in which the total DNA contained in a sample is sequenced, the amplicon metagenomic approach includes a step involving PCR amplification of a phylogenetic marker gene, usually the 16S rRNA gene for bacteria. Although it is less informative than shotgun metagenomics and provides taxonomic information only, marker gene amplicon sequencing might be better suited as a diagnostic tool because it is faster, cost-effective, and applicable to low-biomass

samples such as the environmental swabs collected in this study. However, results have to be interpreted with caution since, as reported previously and exemplified here, low-biomass samples are prone to contamination with exogenous DNA from various sources, including reagents and human skin (56). Although strategies to filter contaminants have been proposed (57), parsing sequences of contaminants from sequences of microbes truly belonging to the sample remains a challenge that needs to be addressed in order to facilitate the use of amplicon sequencing in clinical settings. Selection of reagents with less contaminating DNA could help to circumvent the problem. Although the *gyrB* sequencing protocol allowed the identification of *Serratia marcescens* at the species level in environmental samples, further optimization will be required for accurate bacterial community profiling, since the method failed to detect several bacterial genera in a mock community. This finding might be linked to the primers that were used, which were highly degenerate and could fail to amplify certain groups of bacteria.

Whereas culture only allowed the detection of *S. marcescens* in 3 environmental samples, metagenomic sequencing of the 16S and *gyrB* genes identified sequences related to *Serratia* in most of the samples tested, with greater proportions being observed around beds and in ventilator/infusion pumps, indicating that these sites are potential reservoirs for this bacterium. These sampling sites were directly in contact with babies infected or colonized by *S. marcescens*, in contrast to the other surfaces tested. A previous NICU study based on 16S rRNA gene amplicon sequencing also reported greater relative abundances of outbreak-related bacteria in samples collected from neonate-associated surfaces, compared with other environmental samples (17). In this outbreak, 1 *S. marcescens* isolate was recovered from around 1 bed, and the particular isolate was of the same genotype as 3 isolates from newborns, supporting the metagenomic findings. The presence of *S. marcescens* around beds might be linked to the presence of this bacterium in the gastrointestinal tract of newborns, which represents an important reservoir for cross-contamination (5–7).

The relative abundances of *Serratia* were low in all samples collected in the control room, indicating that the disinfection protocol used to eliminate potential sources of contamination following discharge of a case was efficient. Interestingly, the community composition and structure at several sampling sites in the control room were also clearly distinct from those of the positive rooms, illustrating the strong effects of disinfection on bacterial communities, as described previously (17). Further testing would be required to confirm these observations, since a single disinfected control room was available for this study.

Low relative abundances of *Serratia* sequences were detected in the drains of positive rooms, while 2 of 3 environmental isolates were recovered from those sampling sites. This phenomenon illustrates a potential limitation of the use of relative abundances in gene-targeted metagenomic studies; although the relative abundance of *Serratia* in drains was low, the total numbers of *Serratia* cells might still be higher than on dry surfaces, due to the massive amounts of biomass colonizing drains, which could have facilitated the recovery of *S. marcescens* isolates from those sites. This issue could be circumvented in further studies by using quantitative PCR to estimate the bacterial load in each sample (58) or by spiking exogenous bacteria prior to sample processing (59).

While providing sensitive and informative data regarding the environmental spread of a causative agent during an outbreak, as exemplified here in the context of a *S. marcescens* outbreak, sequencing of bacterial taxonomic marker genes facilitates better overall understanding of bacterial communities in hospitals. Projects such as the Hospital Microbiome Project specifically aim to identify factors that modulate microbial population development in health care environments through extended surface, air, staff, and patient sampling (15). One trend observed in a recent publication from that project was that surfaces in patients' rooms, especially bedrails, harbored microbial communities resembling those of the skin of the current patient (60). Although we could not confirm such a trend, as no such samples were collected from patients,

bacterial genera previously identified as dominant in the skin microbiome of newborns (61), such as *Streptococcus* and *Staphylococcus*, were found at high relative abundances in NICU surface samples. However, these findings have to be interpreted with caution, since DNA from skin bacteria tend to be found everywhere and are often detected as contaminants, as discussed above. Interestingly, *Serratia* was detected at low relative abundances in the negative controls and was not identified as a dominant member of bacterial communities on surfaces in previous NICU metagenomic studies (16, 17), indicating that the high relative abundances observed here at newborn-affected sites might be directly linked to the *S. marcescens* outbreak. However, those studies reported the presence of several bacterial genera commonly linked to NICU outbreaks, as was the case in our study.

In conclusion, outbreak investigations of hospital-acquired infections are challenging but necessary in order to control and to limit the burden of this increasing threat. In this study, we described how the use of high-throughput sequencing applications, such as bacterial WGS and targeted amplicon sequencing, can provide new tools that have the potential to contribute to outbreak investigations. Links between *S. marcescens* isolates and potential antibiotic resistance genes and plasmids were assessed by WGS, while sequencing of bacterial phylogenetic marker genes (16S rRNA and *gyrB* genes) was used to map the distribution of *Serratia* strains and to describe bacterial communities in hospital rooms. Although we applied these tools to a specific NICU outbreak, they could easily be used for investigations of outbreaks caused by other bacteria, with the potential to improve our knowledge of bacterial ecology and transmission pathways in the hospital environment, which in turn could support and provide evidence for the development and evaluation of strategies for infection prevention and control.

#### SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/JCM.00235-18>.

**SUPPLEMENTAL FILE 1**, PDF file, 0.1 MB.

#### ACKNOWLEDGMENTS

We thank the bacteriology and molecular biology technicians at LSPQ for excellent technical support for this study, with special thanks to technical coordinators Lyne Désautels, Dominique Paquette, and Isabelle Robillard. We also thank the personnel from the MUHC Vaccine Study Centre for administrative support and Shauna O'Donnell for her help.

We acknowledge Compute Canada for access to the University of Waterloo High Performance Computing infrastructure (Graham system). This work was funded through internal funds.

#### REFERENCES

- Johnson J, Quach C. 2017. Outbreaks in the neonatal ICU: a review of the literature. *Curr Opin Infect Dis* 30:395–403. <https://doi.org/10.1097/QCO.0000000000000383>.
- Fleisch F, Zimmermann-Baer U, Zbinden R, Bischoff G, Arlettaz R, Waldvogel K, Nadal D, Christian R. 2002. Three consecutive outbreaks of *Serratia marcescens* in a neonatal intensive care unit. *Clin Infect Dis* 34:767–773. <https://doi.org/10.1086/339046>.
- Almuneef MA, Baltimore RS, Farrel PA, Reagan-Cirincione P, Dembry LM. 2001. Molecular typing demonstrating transmission of Gram-negative rods in a neonatal intensive care unit in the absence of a recognized epidemic. *Clin Infect Dis* 32:220–227. <https://doi.org/10.1086/318477>.
- Madani TA, Alsaedi S, James L, Eldeek BS, Jiman-Fatani AA, Alawi MM, Marwan D, Cudal M, Macapagal M, Bahlas R, Farouq M. 2011. *Serratia marcescens*-contaminated baby shampoo causing an outbreak among newborns at King Abdulaziz University Hospital, Jeddah, Saudi Arabia. *J Hosp Infect* 78:16–19. <https://doi.org/10.1016/j.jhin.2010.12.017>.
- Montagnani C, Cocchi P, Lega L, Campana S, Biermann KP, Braggion C, Pecile P, Chiappini E, de Martino M, Galli L. 2015. *Serratia marcescens* outbreak in a neonatal intensive care unit: crucial role of implementing hand hygiene among external consultants. *BMC Infect Dis* 15:11. <https://doi.org/10.1186/s12879-014-0734-6>.
- Gastmeier P. 2014. *Serratia marcescens*: an outbreak experience. *Front Microbiol* 5:81. <https://doi.org/10.3389/fmicb.2014.00081>.
- Moles L, Gómez M, Heilig H, Bustos G, Fuentes S, de Vos W, Fernández L, Rodríguez JM, Jiménez E. 2013. Bacterial diversity in meconium of preterm neonates and evolution of their fecal microbiota during the first month of life. *PLoS One* 8:e66986. <https://doi.org/10.1371/journal.pone.0066986>.
- Dawczynski K, Proquitté H, Roedel J, Edel B, Pfeifer Y, Hoyer H, Dobermann H, Hagel S, Pletz MW. 2016. Intensified colonisation screening according to the recommendations of the German Commission for Hospital Hygiene and Infectious Diseases Prevention (KRINKO): identification and containment of a *Serratia marcescens* outbreak in the neo-

- natal intensive care unit, Jena, Germany, 2013–2014. *Infection* 44: 739–746. <https://doi.org/10.1007/s15010-016-0922-y>.
9. David MD, Weller TMA, Lambert P, Fraise AP. 2006. An outbreak of *Serratia marcescens* on the neonatal unit: a tale of two clones. *J Hosp Infect* 63:27–33. <https://doi.org/10.1016/j.jhin.2005.11.006>.
  10. Miranda G, Kelly C, Solorzano F, Leanos B, Coria R, Patterson JE. 1996. Use of pulsed-field gel electrophoresis typing to study an outbreak of infection due to *Serratia marcescens* in a neonatal intensive care unit. *J Clin Microbiol* 34:3138–3141.
  11. Lai KK, Baker SP, Fontecchio SA. 2004. 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* 25:730–734. <https://doi.org/10.1086/502468>.
  12. Bekal S, Berry C, Reimer AR, Van Domselaar G, Beaudry G, Fournier E, Doualla-Bell F, Levac E, Gaulin C, Ramsay D, Huot C, Walker M, Sieffert C, Tremblay C. 2016. Usefulness of high-quality core genome single-nucleotide variant analysis for subtyping the highly clonal and the most prevalent *Salmonella enterica* serovar Heidelberg clone in the context of outbreak investigations. *J Clin Microbiol* 54:289–295. <https://doi.org/10.1128/JCM.02200-15>.
  13. Salipante SJ, SenGupta DJ, Cummings LA, Land TA, Hoogstraal DR, Cookson BT. 2015. Application of whole-genome sequencing for bacterial strain typing in molecular epidemiology. *J Clin Microbiol* 53: 1072–1079. <https://doi.org/10.1128/JCM.03385-14>.
  14. Fitzpatrick MA, Ozer EA, Hauser AR. 2016. Utility of whole-genome sequencing in characterizing *Acinetobacter* epidemiology and analyzing hospital outbreaks. *J Clin Microbiol* 54:593–612. <https://doi.org/10.1128/JCM.01818-15>.
  15. Shogan BD, Smith DP, Packman AI, Kelley ST, Landon EM, Bhangar S, Vora GJ, Jones RM, Keegan K, Stephens B, Ramos T, Kirkup BC, Levin H, Rosenthal M, Foxman B, Chang EB, Siegel J, Cobey S, An G, Alverdy JC, Olszewski PJ, Martin MO, Marrs R, Hernandez M, Christley S, Morowitz M, Weber S, Gilbert J. 2013. The Hospital Microbiome Project: meeting report for the 2nd Hospital Microbiome Project, Chicago, USA, January 15th, 2013. *Stand Genomic Sci* 8:571–579. <https://doi.org/10.4056/sigs.4187859>.
  16. Hewitt KM, Mannino FL, Gonzalez A, Chase JH, Caporaso JG, Knight R, Kelley ST. 2013. Bacterial diversity in two neonatal intensive care units (NICUs). *PLoS One* 8:e54703. <https://doi.org/10.1371/journal.pone.0054703>.
  17. Bokulich NA, Mills DA, Underwood MA. 2013. Surface microbes in the neonatal intensive care unit: changes with routine cleaning and over time. *J Clin Microbiol* 51:2617–2624. <https://doi.org/10.1128/JCM.00898-13>.
  18. Tenover FC, Arbeit RD, Goering RV, Mickelsen PA, Murray BE, Persing DH, Sminathan B. 1995. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. *J Clin Microbiol* 33:2233–2239.
  19. Petkau A, Mabon P, Sieffert C, Knox NC, Cabral J, Iskander M, Iskander M, Weedmark K, Zaheer R, Katz LS, Nadon C, Reimer A, Taboada E, Beiko RG, Hsiao W, Brinkman F, Graham M, Van Domselaar G. 2017. SNVPhyl: a single nucleotide variant phylogenomics pipeline for microbial genomic epidemiology. *Microb Genom* 3:e000116.
  20. Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, Zhang Y, Blankenberg D, Albert I, Taylor J, Miller W, Kent WJ, Nekrutenko A. 2005. Galaxy: a platform for interactive large-scale genome analysis. *Genome Res* 15:1451–1455. <https://doi.org/10.1101/gr.4086505>.
  21. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
  22. Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27:2987–2993. <https://doi.org/10.1093/bioinformatics/btr509>.
  23. Lefort V, Longueville J-E, Gascuel O. 2017. SMS: smart model selection in PhyML. *Mol Biol Evol* 34:2422–2424. <https://doi.org/10.1093/molbev/msx149>.
  24. Ribeiro-Gonçalves B, Francisco AP, Vaz C, Ramirez M, Carriço JA. 2016. PHYLOViZ Online: web-based tool for visualization, phylogenetic inference, analysis and sharing of minimum spanning trees. *Nucleic Acids Res* 44:W246–W251. <https://doi.org/10.1093/nar/gkw359>.
  25. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
  26. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, Lago BA, Dave BM, Pereira S, Sharma AN, Doshi S, Courtot M, Lo R, Williams LE, Frye JG, Elsayegh T, Sardar D, Westman EL, Pawlowski AC, Johnson TA, Brinkman FSL, Wright GD, McArthur AG. 2017. CARD 2017: expansion and model-centric curation of the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res* 45:D566–D573. <https://doi.org/10.1093/nar/gkw1004>.
  27. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, Møller Aarestrup F, Hasman H. 2014. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother* 58:3895–3903. <https://doi.org/10.1128/AAC.02412-14>.
  28. Barret M, Briand M, Bonneau S, Préveaux A, Valière S, Bouchez O, Hunault G, Simoneau P, Jacques M-A. 2015. Emergence shapes the structure of the seed microbiota. *Appl Environ Microbiol* 81:1257–1266. <https://doi.org/10.1128/AEM.03722-14>.
  29. Dauga C. 2002. Evolution of the *gyrB* gene and the molecular phylogeny of Enterobacteriaceae: a model molecule for molecular systematic studies. *Int J Syst Evol Microbiol* 52:531–547. <https://doi.org/10.1099/00207713-52-2-531>.
  30. Delmas J, Breyse F, Devulder G, Flandrois J-P, Chomarat M. 2006. Rapid identification of Enterobacteriaceae by sequencing DNA gyrase subunit B encoding gene. *Diagn Microbiol Infect Dis* 55:263–268. <https://doi.org/10.1016/j.diagmicrobio.2006.02.003>.
  31. Illumina. 2013. 16S metagenomic sequencing library preparation: preparing 16S ribosomal RNA gene amplicons for the Illumina MiSeq system. Illumina, San Diego, CA. [https://support.illumina.com/documents/documentation/chemistry\\_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf](https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf).
  32. Tremblay J, Singh K, Fern A, Kirton E, He S, Woyke T, Lee J, Chen F, Dangl J, Tringe S. 2015. Primer and platform effects on 16S rRNA tag sequencing. *Front Microbiol* 6:771. <https://doi.org/10.3389/fmicb.2015.00771>.
  33. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41: D590–D596. <https://doi.org/10.1093/nar/gks1219>.
  34. McMurdie PJ, Holmes S. 2014. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol* 10:e1003531. <https://doi.org/10.1371/journal.pcbi.1003531>.
  35. Wood DE, Salzberg SL. 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 15:R46. <https://doi.org/10.1186/gb-2014-15-3-r46>.
  36. Valenzuela-González F, Martínez-Porchas M, Villalpando-Canchola E, Vargas-Albores F. 2016. Studying long 16S rDNA sequences with ultrafast-metagenomic sequence classification using exact alignments (Kraken). *J Microbiol Methods* 122:38–42. <https://doi.org/10.1016/j.mimet.2016.01.011>.
  37. Gao X, Lin H, Revanna K, Dong Q. 2017. A Bayesian taxonomic classification method for 16S rRNA gene sequences with improved species-level accuracy. *BMC Bioinformatics* 18:247. <https://doi.org/10.1186/s12859-017-1670-4>.
  38. Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. 2011. UniFrac: an effective distance metric for microbial community comparison. *ISME J* 5:169–172. <https://doi.org/10.1038/ismej.2010.133>.
  39. Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20:289–290. <https://doi.org/10.1093/bioinformatics/btg412>.
  40. Deurenberg RH, Bathoorn E, Chlebowicz MA, Couto N, Ferdous M, García-Cobos S, Kooistra-Smid AMD, Raangs EC, Rosema S, Veloo ACM, Zhou K, Friedrich AW, Rossen JWA. 2017. Application of next generation sequencing in clinical microbiology and infection prevention. *J Biotechnol* 243:16–24. <https://doi.org/10.1016/j.jbiotec.2016.12.022>.
  41. Dunne WM, Westblade LF, Ford B. 2012. Next-generation and whole-genome sequencing in the diagnostic clinical microbiology laboratory. *Eur J Clin Microbiol Infect Dis* 31:1719–1726. <https://doi.org/10.1007/s10096-012-1641-7>.
  42. Mellmann A, Harmsen D, Cummings CA, Zentz EB, Leopold SR, Rico A, Prior K, Szczepanowski R, Ji Y, Zhang W, McLaughlin SF, Henkhaus JK, Leopold B, Bielaszewska M, Prager R, Brzoska PM, Moore RL, Guenther S, Rothberg JM, Karch H. 2011. Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid

- next generation sequencing technology. PLoS One 6:e22751. <https://doi.org/10.1371/journal.pone.0022751>.
43. Lindsey RL, Pouseele H, Chen JC, Strockbine NA, Carleton HA. 2016. Implementation of whole genome sequencing (WGS) for identification and characterization of Shiga toxin-producing *Escherichia coli* (STEC) in the United States. *Front Microbiol* 7:766. <https://doi.org/10.3389/fmicb.2016.00766>.
  44. Lewis T, Loman NJ, Bingle L, Jumaa P, Weinstock GM, Mortibooy D, Pallen MJ. 2010. High-throughput whole-genome sequencing to dissect the epidemiology of *Acinetobacter baumannii* isolates from a hospital outbreak. *J Hosp Infect* 75:37–41. <https://doi.org/10.1016/j.jhin.2010.01.012>.
  45. Taylor AJ, Lappi V, Wolfgang WJ, Lapierre P, Palumbo MJ, Medus C, Boxrud D. 2015. Characterization of foodborne outbreaks of *Salmonella enterica* serovar Enteritidis with whole-genome sequencing single nucleotide polymorphism-based analysis for surveillance and outbreak detection. *J Clin Microbiol* 53:3334–3340. <https://doi.org/10.1128/JCM.01280-15>.
  46. Dallman TJ, Byrne L, Ashton PM, Cowley LA, Perry NT, Adak G, Petrovska L, Ellis RJ, Elson R, Underwood A, Green J, Hanage WP, Jenkins C, Grant K, Wain J. 2015. Whole-genome sequencing for national surveillance of Shiga toxin-producing *Escherichia coli* O157. *Clin Infect Dis* 61:305–312. <https://doi.org/10.1093/cid/civ318>.
  47. Reimer AR, Van Domselaar G, Stroika S, Walker M, Kent H, Tarr C, Talkington D, Rowe L, Olsen-Rasmussen M, Frace M, Sammons S, Dahourou GA, Boncy J, Smith AM, Mabon P, Petkau A, Graham M, Gilmour MW, Gerner-Smidt P. 2011. Comparative genomics of *Vibrio cholerae* from Haiti, Asia, and Africa. *Emerg Infect Dis* 17:2113–2121. <https://doi.org/10.3201/eid1711.110794>.
  48. Milisavljevic V, Wu F, Larson E, Rubenstein D, Ross B, Drusin LM, Della-Latta P, Saiman L. 2004. Molecular epidemiology of *Serratia marcescens* outbreaks in two neonatal intensive care units. *Infect Control Hosp Epidemiol* 25:719–722. <https://doi.org/10.1086/502466>.
  49. Maragakis LL, Winkler A, Tucker MG, Cosgrove SE, Ross T, Lawson E, Carroll KC, Perl TM. 2008. Outbreak of multidrug-resistant *Serratia marcescens* infection in a neonatal intensive care unit. *Infect Control Hosp Epidemiol* 29:418–423. <https://doi.org/10.1086/587969>.
  50. Maltezou HC, Tryfinopoulou K, Katerelos P, Ftika L, Pappa O, Tseroni M, Kostis E, Kostalos C, Prifti H, Tzanetou K, Vatopoulos A. 2012. Consecutive *Serratia marcescens* multiclonal outbreaks in a neonatal intensive care unit. *Am J Infect Control* 40:637–642. <https://doi.org/10.1016/j.ajic.2011.08.019>.
  51. McGeer A, Low DE, Penner J, Ng J, Goldman C, Simor AE. 1990. Use of molecular typing to study the epidemiology of *Serratia marcescens*. *J Clin Microbiol* 28:55–58.
  52. Shaw KJ, Rather PN, Sabatelli FJ, Mann P, Munayyer H, Mierzwa R, Petrikos GL, Hare RS, Miller GH, Bennett P, Downey P. 1992. Characterization of the chromosomal *aac(6′-Ic)* gene from *Serratia marcescens*. *Antimicrob Agents Chemother* 36:1447–1455.
  53. Wu L-T, Tsou M-F, Wu H-J, Chen H-E, Chuang Y-C, Yu W-L. 2004. Survey of CTX-M-3 extended-spectrum  $\beta$ -lactamase (ESBL) among cefotaxime-resistant *Serratia marcescens* at a medical center in middle Taiwan. *Diagn Microbiol Infect Dis* 49:125–129. <https://doi.org/10.1016/j.diagmicrobio.2004.02.004>.
  54. Köser CU, Ellington MJ, Peacock SJ. 2014. Whole-genome sequencing to control antimicrobial resistance. *Trends Genet* 30:401–407. <https://doi.org/10.1016/j.tig.2014.07.003>.
  55. Forbes JD, Knox NC, Ronholm J, Pagotto F, Reimer A. 2017. Metagenomics: the next culture-independent game changer. *Front Microbiol* 8:1069. <https://doi.org/10.3389/fmicb.2017.01069>.
  56. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman NJ, Walker AW. 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* 12:87. <https://doi.org/10.1186/s12915-014-0087-z>.
  57. Nguyen NH, Smith D, Peay K, Kennedy P. 2015. Parsing ecological signal from noise in next generation amplicon sequencing. *New Phytol* 205:1389–1393. <https://doi.org/10.1111/nph.12923>.
  58. Nadkarni MA, Martin FE, Jacques NA, Hunter N. 2002. Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. *Microbiology* 148:257–266. <https://doi.org/10.1099/00221287-148-1-257>.
  59. Stämmler F, Gläsner J, Hiergeist A, Holler E, Weber D, Oefner PJ, Gessner A, Spang R. 2016. Adjusting microbiome profiles for differences in microbial load by spike-in bacteria. *Microbiome* 4:28. <https://doi.org/10.1186/s40168-016-0175-0>.
  60. Lax S, Smith D, Sangwan N, Handley K, Larsen P, Richardson M, Taylor S, Landon E, Alverdy J, Siegel J, Stephens B, Knight R, Gilbert JA. 2017. Colonization and succession of hospital-associated microbiota. *Sci Transl Med* 9:eaah6500. <https://doi.org/10.1126/scitranslmed.aah6500>.
  61. Capone KA, Dowd SE, Stamatias GN, Nikolovski J. 2011. Diversity of the human skin microbiome early in life. *J Invest Dermatol* 131:2026–2032. <https://doi.org/10.1038/jid.2011.168>.