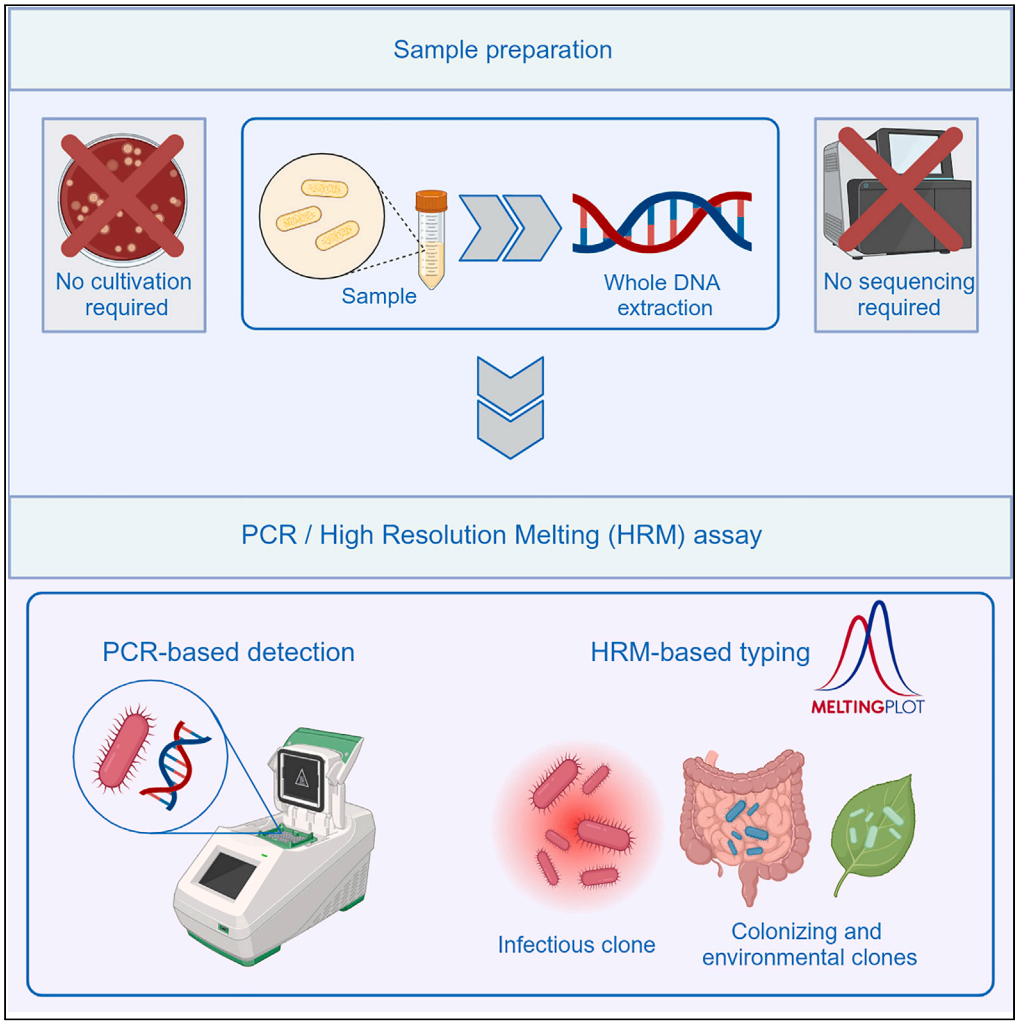
iScience

CellPress
OPEN ACCESS

Article

Cultivation and sequencing-free protocol for *Serratia marcescens* detection and typing



Alessandro Alvaro, Aurora Piazza, Stella Papaleo, ..., Daniela Maria Cirillo, Gianvincenzo Zuccotti, Francesco Comandatore

francesco.comandatore@unimi.it

Highlights

Most *Serratia marcescens* clinical infections are due to a specific clone

We developed a PCR/HRM protocol for fast *Serratia marcescens* detection and typing

The protocol has high sensitivity and sensibility in identify the infectious clone

The protocol is highly reproducible and gives portable results

Alvaro et al., iScience 27, 109402
April 19, 2024
https://doi.org/10.1016/j.isci.2024.109402

iScience

CellPress
OPEN ACCESS

Article

Cultivation and sequencing-free protocol for *Serratia marcescens* detection and typing

Alessandro Alvaro,[1,2] Aurora Piazza,[3] Stella Papaleo,[1] Matteo Perini,[1,4] Ajay Ratan Pasala,[1,5,6] Simona Panelli,[1] Tiago Nardi,[1,7] Riccardo Nodari,[1] Lodovico Sterzi,[1] Cristina Pagani,[8] Cristina Merla,[9] Daniele Castelli,[10] Emanuela Olivieri,[11] Silvia Bracco,[12] Maria Laura Ferrando,[13] Francesca Saluzzo,[13] Sara Giordana Rimoldi,[8] Marta Corbella,[9] Annalisa Cavallero,[10] Paola Prati,[11] Claudio Farina,[12] Daniela Maria Cirillo,[13] Gianvincenzo Zuccotti,[1,14] and Francesco Comandatore[1,15,*]

SUMMARY

***Serratia marcescens* is an opportunistic pathogen that survives in inhospitable environments causing large outbreaks, particularly in neonatal intensive care units (NICUs). Genomic studies revealed that most *S. marcescens* nosocomial infections are caused by a specific clone (here "Infectious clone"). Whole genome sequencing (WGS) is the only portable method able to identify this clone, but it requires days to obtain results. We present a cultivation-free hypervariable-locus melting typing (HLMT) protocol for the fast detection and typing of *S. marcescens*, with 100% detection capability on mixed samples and a limit of detection that can reach the 10 genome copies. The protocol was able to identify the *S. marcescens* infectious clone with 97% specificity and 96% sensitivity when compared to WGS, yielding typing results portable among laboratories. The protocol is a cost and time saving method for *S. marcescens* detection and typing for large environmental/clinical surveillance screenings, also in low-middle income countries.**

INTRODUCTION

Serratia marcescens is a rod-shaped, Gram-negative bacterium that belongs to the order *Enterobacteriales*.[1,2] The bacterium, usually found in soil, water, plants and animals, can also behave as an opportunistic pathogen infecting people with weak immune systems.[3,4] Moreover, it is able to survive in the hospital environment and rapidly spreads among the hospital patients.[2] Indeed, *S. marcescens* causes healthcare-associated infections (HAIs) and outbreaks, in particular in neonatal intensive care units (NICUs), with a mortality rate that ranges from 25% to 58%.[2,5]

Recent genomic studies revealed that most of the *S. marcescens* infections are caused by a specific lineage.[3,6,7] Thus, the rapid identification of this infection-associated *S. marcescens* lineage in the medical practice routine is crucial to limit *S. marcescens* HAIs and to implement effective infection prevention and control (IPC) procedures.[8]

At the state of the art, the most used methods for *S. marcescens* typing are pulsed-field gel electrophoresis (PFGE), matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) and whole genome sequencing (WGS), while no multi-locus sequence typing (MLST) scheme is currently available.[9–11] Unfortunately, each of these methods has limits. PFGE is time-consuming and less discriminatory than sequencing-based typing methods.[12] MALDI-TOF is a fast and largely used method for bacterial species identification in hospital context. However, the application of this technique may be prohibitive in low-income countries hospitals due to the high purchase and maintenance costs of the instrument.[13] Furthermore, there is still limited evidence on the capability of the MALDI-TOF method to distinguish sub-species

[1]Department of Biomedical and Clinical Sciences, Pediatric Clinical Research Center "Romeo and Enrica Invernizzi", University of Milan, 20157 Milan, Italy
[2]Department of Biosciences and Pediatric Clinical Research Center "Romeo Ed Enrica Invernizzi", University of Milan, 20133 Milan, Italy
[3]Unit of Microbiology and Clinical Microbiology, Department of Clinical-Surgical, Diagnostic and Pediatric Sciences, University of Pavia, Pavia 27100, Italy
[4]Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY 10032, USA
[5]Biochemistry, Microbiology and Immunology Department, Faculty of Medicine, University of Ottawa, Ottawa, ON K1H 8M5, Canada
[6]Centre for Innovation, Canadian Blood Services, Ottawa, ON K1G 4J5, Canada
[7]Department of Biology and Biotechnology, University of Pavia, 27100 Pavia, Italy
[8]Laboratorio di Microbiologia Clinica, Virologia e Diagnostica delle Bioemergenze, ASST Fatebenefratelli Sacco, 20157 Milan, Italy
[9]Department of Microbiology & Virology, Fondazione IRCCS Policlinico San Matteo, Viale Camillo Golgi 19, 27100 Pavia, Italy
[10]Microbiology Unit, Fondazione IRCCS San Gerardo, 20900 Monza, Italy
[11]Istituto Zooprofilattico Sperimentale della Lombardia e dell'Emilia Romagna (IZSLER), 27100 Pavia, Italy
[12]Laboratory of Microbiology and Virology, Azienda Socio-Sanitaria Territoriale (ASST) Papa Giovanni XXIII, 24127 Bergamo, Italy
[13]Emerging Bacterial Pathogens Unit, Division of Immunology, Transplantation and Infectious Diseases, IRCCS San Raffaele Scientific Institute, 20132 Milan, Italy
[14]Department of Paediatrics, Children's Hospital "V. Buzzi", 20154 Milano, Italy
[15]Lead contact
*Correspondence: francesco.comandatore@unimi.it
https://doi.org/10.1016/j.isci.2024.109402

Table 1. Information about the PCR/HRM primers designed in the study

Primer pair	Target gene	Primer sequences	Amplicon size (nt)
HLMT_SMA_1	SMDB11_2070	HMLT_SMA_1_F: GAACYTGCGCATGATTTATGCG HMLT_SMA_1_R: CGGWGACGACCTGCAGCTG	118
HLMT_SMA_2	SMDB11_3518	HMLT_SMA_2_F: ATGRCCGGYAAGGCCATCGAT HMLT_SMA_2_R:TTCAGGGCGACCGCGTCG	134

The target genes are from the reference strain *Serratia marcescens* D11 strain (HG326223.1).

lineages.[10] Furthermore, the method has been shown to be less discriminatory than PFGE and MLST in clone assignment.[14] Another recently proposed method, named high-throughput short sequence typing scheme (HiSST), allows for the differentiation of *S. marcescens* clones both in pure culture and in environmental samples.[15,16] However, although this method is effective, it relies on sequencing. Despite the high precision in the discrimination of bacterial clones, sequencing-based typing methods (e.g., HiSST and WGS-based typing) are laborious, time-consuming and expensive.[17] This can limit their applicability for the real-time surveillance of rapidly spreading pathogens, such as *S. marcescens*. High resolution melting (HRM)-based typing is a fast and inexpensive method for the discrimination of bacterial clones on the basis of the melting temperature of PCR amplicons.[18] It can be performed on most real-time PCR instruments and the entire protocol is fast (~5 h) and inexpensive (~5$/sample with three technical replicates), making HRM-based typing suitable for nosocomial infection surveillance.[19] Indeed, HRM analysis has been increasingly used as a research tool, being successfully implemented in fields ranging from forensics to epidemiological investigations.[20–22]

Recently, we proposed hypervariable-locus melting typing (HLMT), a novel approach to HRM-based typing.[19,23–25] This approach is based on the use of hypervariable genetic regions for HRM protocol design and a novel graph-based algorithm to cluster bacterial strains on the basis of their melting temperatures.[23] Recently, we successfully applied this approach for the typing of the nosocomial pathogen *Klebsiella pneumoniae*.[23,24]

During an outbreak, *S. marcescens* is able to rapidly spread among patients and to colonize the hospital environment. Most of the currently available typing methods (e.g., PFGE, WGS, and HRM) are designed to use DNA extracted from pure bacterial isolates that requires a ~24 h cultivation step, limiting the efficiency of the current IPC procedures for *S. marcescens*.

In this work, we present an HRM-based cultivation-free protocol for the detection and typing of *S. marcescens*. This protocol is able to detect and type with high precision *S. marcescens* performing a PCR/HRM amplification experiment on the DNA extracted directly from the samples, without the need for a cultivation step.

RESULTS

Cultivation-free HLMT protocol design

The orthologue analysis performed on the 223 *S. marcescens* genomes identified 8,468 orthologous groups, of which 2,656 present in single copy in all the genomes. The BLASTn[26] search revealed that 40 of these orthologous groups were not, or rarely, present in other bacterial species. The analysis of genetic variability, phylogenetic signal and presence of regions suitable for the designing of HRM/PCR primers led to the selection of two target genes among the 40 suitable orthologous groups, herein labeled HLMT_SMA_1 and HLMT_SMA_2. These two genes were reported on the *S. marcescens* Db11 strain genome (HG326223.1) with the locus tags SMDB11_2070 (CDG12644.1, annotation: "hypothetical protein") and SMDB11_3518 (CDG14081.1, annotation: "alkaline phosphatase"), respectively (see Table 1). The BLASTn[26] search of both the target genes returned 23 common not-*Serratia marcescens* hits: one against a genome of *Serratia bockelmannii*, one *Serratia surfactantfaciens*, two *Serratia nematodiphila*, seven *Serratia* spp., and ten *Serratia ureilytica*. The designed primers are reported in Table 1.

The genomic dataset and WGS-based typing

The Global genomic dataset subjected to WGS-based typing was composed of 516 *S. marcescens* genomes, of which 230 sequenced in this study (BioProject PRJNA957961) and 286 retrieved from public databases. For 369 out of the 516 strains (~72%), information about the isolation material was available: 239/369 (65%) were classified as "Infection," 72 (~19%) as "Colonization," 39 (~11%) as "Environmental," and 19 (~5%) as "Hospital environment" (we will refer to these categories as "sample categories"). Details about the metadata of the isolates and the genome assemblies are reported in Table S1. The core SNP calling procedure produced an alignment of 31,510 core SNPs and the obtained maximum likelihood phylogenetic tree is reported in Figure 1.

The K-means unsupervised clustering grouped the 516 *S. marcescens* strains in five Phylo Clusters (Figure 1C), after that the optimum number of clusters was determined on the basis of the silhouette analysis (Figure 1B). 308 out of the 516 strains (~60%) were classified in Phylo Cluster 1, 38 (~7%) in Phylo Cluster 2, 72 (~14%) in Phylo Cluster 3, 44 (~9%) in Phylo Cluster 4, and 54 (~10%) genomes in Phylo Cluster 5. Among the 230 *S. marcescens* strains sequenced in this study, 109 (~47%) were classified in Phylo Cluster 1, 10 (~4%) in Phylo Cluster 2, 55 (~24%) in Phylo Cluster 3, 13 (~6%) in Phylo Cluster 4, and 43 (~19%) in Phylo Cluster 5.

The analysis of the residuals of the chi squared analysis between Phylo clusters and sample categories in the Global dataset revealed that Phylo Cluster 1 was the only one strongly positively associated to "Infection" and negatively associated both to "Colonization" or "Environmental" (Figure S1A). The Phylo Clusters 2, 3, and 5 were negatively associated with the "Infection" and positively associated with at least one

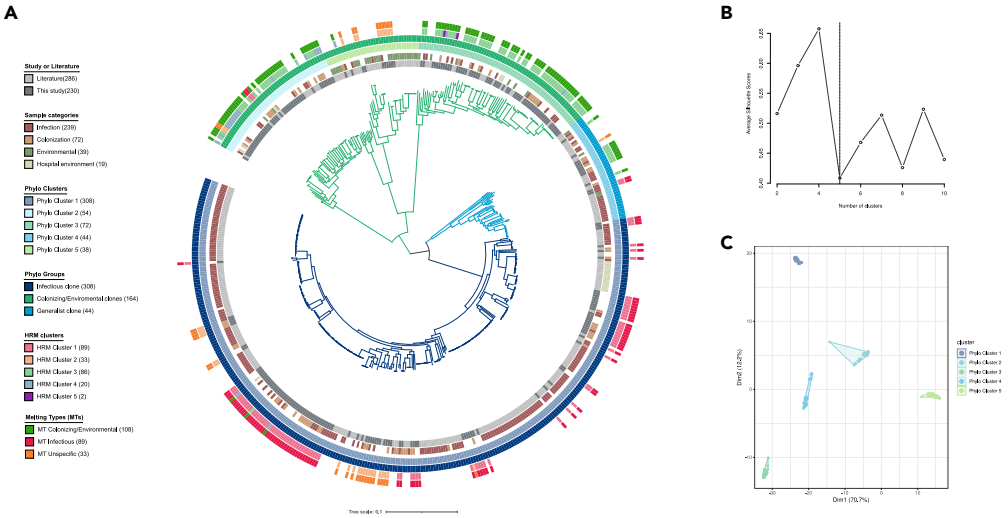


Figure 1. Comparison of WGS-based typing and HRM-based typing

(A) SNP-based maximum likelihood (ML) phylogenetic tree of the 516 *Serratia marcescens* strains of the Global dataset (230 sequenced in this study and 286 from public databases). The tree branches' colors indicate the Phylo Groups: "Infectious clone" is blue, "Colonizing/Environmental clones" in green and "Generalist clone" in azure. The rings around the tree shows metadata of the strains, starting from the inner circle: i) "Study or Literature", divides the strains sequenced in this work ("This Study", dark gray) from those collected from databases ("Literature", light gray); ii) "Sample category" reported the type of samples from which the strains were isolated, "Infection" strains were isolated from usually sterile human body compartments (e.g., blood and urine), "Colonizing" from usually colonized human compartments (e.g., gut), "Environmental" from soil and water, "Hospital environment" from the hospital superficies; iii) "Phylo Clusters" are the clusters identified by K-means clustering on coreSNPs distance (colors reported in legend); iv) "Phylogroups" are the monophyletic macro-lineages identified combining phylogeny and ecological specializations (colors in the legend); v) "HRM clusters" are the clusters identified by the MeltingPlot tool on the basis of melting temperatures; vi) "Melting Type" are the macro-groups of HRM clusters identified on the basis of HRM clustering and melting temperatures. Melting Types are the main results of the HRM protocol presented in this work. (B) Silhouette plot of the number of clusters on the basis of SNP distances. (C) Principal Component Analysis (PCA) performed on SNP distances, with the WGS clusters (identified by K-means algorithm) coloured.

of the other sample categories: Phylo Cluster 2 was particularly skewed toward "Colonization," Phylo Cluster 3 toward both "Colonization" and "Environmental," and Phylo Cluster 5 was strongly associated with "Environmental." Lastly, Phylo Cluster 4 was not associated with any of the sample categories. On the basis of these results, we arranged the Phylo Clusters in three monophyletic Phylo Groups: "Infectious clone" (which includes Phylo Cluster 1 only), "Colonizing/Environmental clones'' (including Phylo Clusters 2, 3, and 5) and "Generalist clone" (which includes Phylo Cluster 4). The Phylo Groups are represented in Figures 1 and 2.

More in detail, the Phylo Group "Infectious clone" included 308/516 (~60%) strains, the "Colonizing/Environmental clones" 164/516 (~32%), while the remaining 44 (~9%) fell into the "Generalist clone." Among the 230 strains of the Study dataset, 109 (47%) were classified as "Infectious clone," 108 (47%) as "Colonizing/Environmental clones," and 13 (6%) fell into the Phylo Group "Generalist clone."

HLMT protocol typing capability

The 230 strains included in the Study dataset were subjected to HRM assay using the primers designed on the target genes HLMT_SMA_1 and HLMT_SMA_2 (see Method details). The resulting melting temperatures (reported in Table S1) were used to group the strains using the graph-based clustering algorithm implemented in the MeltingPlot tool.[25] The analysis grouped the strains in five HRM Clusters (Figure S2A). Chi-Squared analysis showed a significant association (p value $<2.2e^{-16}$) between the HRM Clusters and Phylo Clusters (Figure S1B). Notably, HRM Cluster 1 was positively associated only with Phylo Cluster 1 (which was also associated with Phylo Group "Infection clone," see aforementioned text). The scatterplot of the melting temperatures showed that the five HRM Clusters can be arranged in three Melting Types (MTs) (Figures 2 and S2B): the first one, enriched in strains belonging to the "Infection clone" (Figures S1C and S2B), was called "MT Infectious"; the second one, enriched in strains from the "Colonizing/Environmental clones," was called "MT Colonizing/Environmental"; and the last one, including strains both from "Infection clone" and "Colonizing/Environmental clones," was called "MT Unspecific" (see Figures 1 and 2). Among the 230 Study dataset strains subjected to HMR typing, 89 (~39%) were assigned to the "MT Infectious," 108 (~47%) to the "MT Colonising/Environmental," and 33 (~14%) to the "MT Unspecific." The protocol thus provided a classification (either "MT Infectious" or "MT Colonising/Environmental") for 197/230 strains (~86%). This HRM-based classification correctly distinguished the "Infectious clone" from the other clones in 190/197 (96%) of the strains (Figure S1D), with a specificity of 0.97 and a sensitivity of 0.96.

The phylogeny of the two target genes (HLMT_SMA_1 and HLMT-SMA_2) were compared with the global SNP-based protocol phylogeny to assess whether the incongruences between HRM Clusters and WGS Clusters could be due to horizontal gene transfer (HGT) events. As shown in Figure S3, in case of congruent HRM and WGS clustering (Figures S3A and S3C), the phylogeny of the HLMT_SMA_1 and HLMT_SMA_2 target genes are mostly coherent with the global SNP-based phylogeny. While, as shown in Figures S3B and S3D, for most

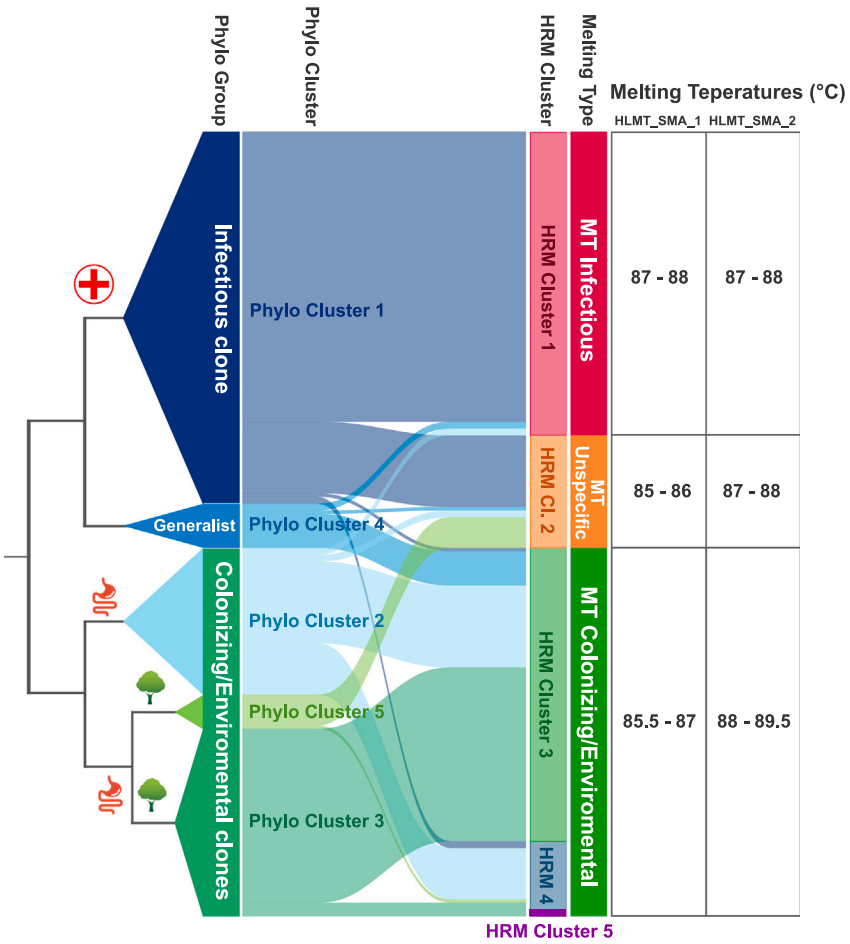


Figure 2. WGS-based typing vs. HRM_based typing

In this figure we collected the most relevant information regarding the typing capabilities of the HRM protocol presented in this work. On the left, the SNP-based phylogenetic analysis of the 230 strains sequenced in this work and subjected to HRM typing. Colored triangles represent the five Phylo Clusters identified using the unsupervised K-means clustering algorithm. The icons on the nodes highlight the Phylo Cluster's specializations: Phylo Cluster 1 is enriched in strains isolated from infections, Phylo Cluster 2 from colonizations, Phylo Cluster 3 from colonization and environment, Phylo Cluster 4 has no specialization, lastly Phylo Cluster 5 from environment. In the middle, an alluvial plot connects the strains on the tree to the relative HRM Clusters and Melting Types. The melting temperatures specific for each Melting Type are reported on the right.

of the misclassified strains (HRM clustering not congruent to WGS clustering), the position on one or both the target gene trees is evidently not coherent with the species tree, suggesting HGT events. In several cases, these phylogenetic incongruences involve strains classified by HRM-based typing as "MT Unspecific" (Figures S3B and S3D).

HLMT protocol reproducibility and portability

The reproducibility and portability of the HLMT protocol was evaluated on 16 representative strains. Four out of the 16 test dataset isolates (40%) were selected among the 7/197 strains wrongly classified by HLMT typing (see aforementioned text). As shown in Figure S4, the obtained melting temperatures were significantly correlated (linear regression, p value <0.01 for both the primers) and the obtained melting temperature varied less than 0.5°C in 30 out of 32 (~94%) HRM experiments (two primers for each of the 16 isolates). Lastly, the MTs assigned by the two independent laboratories were perfectly congruent for 14 out of the 16 (~88%) tested isolates (Table S2). The differences regarded isolates classified as "MT Unspecific" or "Undetermined." The coherence between the HLMT classifications obtained by the two laboratories is observed both for isolates with HLMT wrong and correct typing.

HLMT protocol specificity for *S. marcescens* species

The protocol specificity for *S. marcescens* has been evaluated by *in silico* and *in vitro* analyses. The *in silico* Primer-BLAST[26] search of both HLMT_SMA_1 and HLMT_SMA_2 primer pairs, returned a total of 158 shared hits: 146 (~92%) on *S. marcescens* genomic sequences, five *Serratia* sp., and seven *S. ureilytica*.

The *in vitro* tests were carried out by PCR experiments (using both the primer pairs) on: (1) *S. marcescens*-free fecal samples; (2) isolates of five *Serratia* species; (3) representative *S. marcescens* isolates; (4) mixed samples containing DNA from 16 *Serratia marcescens*-free fecal samples and *S. marcescens* isolates; (5) DNA extracted from mixed samples obtained by spiking an *S. marcescens*-free soil sample with different amounts of *S. marcescens* cells from overnight cultures. The images of the electrophoresis gels obtained from the PCR experiments are reported in Image S1 (for point 1) and Image S2 (for points 2, 3, and 4) and Image S3 (for point 5).

As shown in Image S1, PCR experiments on the *S. marcescens*-free fecal samples did not produce observable bands relative to the expected amplicon size, even if some non-specific bands are observed for HLMT_SMA_2 primer set. Furthermore, both primer pairs (HLMT_SMA_1 and HLMT_SMA_2) successfully amplified *S. marcescens* DNA (Images S2A and S2B), while no amplification was obtained for *S. rubidaea*, *S. ficaria*, *S. liquefaciens*, and *S. odorifera*. Moreover, no significant non-specific bands were observed in the agarose gel after electrophoresis. As shown in Images S2C–S2F and S3, both primer pairs (HLMT_SMA_1 and HLMT_SMA_2) were able to specifically amplify *S. marcescens* DNA in different types of mixed samples, respectively containing DNA from *S. marcescens* isolates and *S. marcescens*-free fecal samples (Image S2), and DNA extracted from an *S. marcescens*-free soil sample spiked with different amounts of *S. marcescens* intact cells (Image S3).

HLMT protocol limit of detection

The limit of detection for the two primer pairs (HLMT_SMA_1 and HLMT_SMA_2) was evaluated both via PCR and real-time qPCR. In the PCR experiments the primers successfully amplified *S. marcescens* DNA both at 0.1 ng/μL and 0.01 ng/μL *S. marcescens* DNA concentrations, equivalent to 10^4 and 10^3 genome copies, respectively (Image S2). During qPCR experiments, DNA of a fecal sample was mixed with *S. marcescens* DNA at 10 ng/μL, 1 ng/μL, 0.1 ng/μL, 0.01 ng/μL, 0.001 ng/μL, and 0.0001 ng/μL concentration, corresponding to the order of 10^6, 10^5, 10^4, 10^3, 10^2, and 10 genome copies. *S. marcescens* DNA was detected down to 0.0001 ng/μL concentration, equivalent to 10 genome copies. Similar experiments were repeated on mixed samples obtained by spiking an *S. marcescens*-free soil with different amounts (from 10^6 to 10^3) of S. *marcescens* intact cells from overnight cultures. In the PCR experiments, the primers successfully amplified the DNA extracted from the soil sample spiked with 10^5 cells (10^4 for some *S. marcescens* clusters). Concerning qPCR, the HRM analysis was successfully conducted till the soil sample spiked with 10^4 cells.

DISCUSSION

Serratia marcescens is one of the most important nosocomial pathogens world-wide. The bacterium is able to survive on various surfaces, including soap, and to rapidly spread among hospital patients, causing large nosocomial outbreaks, particularly in NICUs.[16,27–32] *S. marcescens* mainly colonize the environment (soil, plants, and animals) and recent genomic studies revealed that the nosocomial *S. marcescens* infections are mostly caused by a specific lineage of the bacterium (here called "Infectious clone").[3,6,7] Considering the lethality of the bacterium, its fast-spreading capability and the existence of a strictly infection-associated lineage, a fast and precise detecting/typing method is required to establish effective infection control strategies in hospitals to prevent and to limit outbreaks.

In this work we present a novel HLMT protocol for the cultivation-free detection and typing of *S. marcescens*. The rationale of this cultivation-free HLMT protocol is to use HRM/PCR primer pairs on target genes specifically present in *S. marcescens* and designed to amplify highly variable genetic regions. The use of these primer pairs in HRM assays allows distinguishing *S. marcescens* strains of the infectious clone on the basis of their melting temperatures. Moreover, the designing of primer pairs on target genes specifically present in *S. marcescens* allows the use of this protocol for the detection and typing of *S. marcescens* starting from the DNA extracted directly from the samples, avoiding the cultivation step and thus greatly decreasing the time needed for the results. The target genes were identified by analyzing more than 200 *S. marcescens* genome assemblies present in public databases. Following the HLMT approach, the primers were designed to amplify highly variable genetic regions with the aim to obtain a highly discriminatory HRM protocol.

PCR and real-time qPCR experiments on mixed DNA (*S. marcescens*-free fecal samples + *S. marcescens* strain) showed that the protocol is able to detect *S. marcescens* DNA in mixed samples as low as ten genomic copies. The same approach was applied on a complex matrix (*S. marcescens*-free soil sample) spiked with various amounts of intact *S. marcescens* cells. Following DNA extraction, PCR showed that the protocol detects up to 10^4 (up to 10^3 in some cases) *S. marcescens* cells added to a freshly sampled soil. Experiments on both types of mixed sample evidence that the protocol can thus be considered also for the direct detection of the bacterium using DNA extracted from environmental or clinical samples, avoiding the cultivation step.

We also evaluated the typing capability of the HLMT protocol on 230 *S. marcescens* strains (collected from seven hospitals and one community health institution) by comparing the results of WGS-based and HLMT-based typing. The WGS-based typing (SNP-based phylogenetic analysis), distinguished three main groups (Phylo Groups) that we called "Infectious clone," "Colonizing/Environmental clones," and "Generalist clone," with the first one being strongly associated with infections in clinical context. On the other hand, HLMT typing revealed the existence of three MTs, groups of strains identified by their specific melting temperatures. One MT was strongly associated with the Phylo Group "Infectious clone" (we called this MT "Infectious"), one to the "Colonizing/Environmental clones" (but also to "Generalist clone") (MT "Colonizing/Environmental"), and one was not associated with any Phylo group (MT "Unspecific"). The co-occurrence analysis revealed that these MTs can be successfully used to distinguish the strains belonging to the Phylo group "Infectious clone" from the others. Indeed, the HRM assignment to "MT Infectious" or "MT Colonizing/Environmental" predicts the relative Phylo Group with an exceptionally high sensitivity and specificity.

We also evaluated the reproducibility of the HLMT protocol and the portability of typing results among different laboratories on a representative subset of eight strains: a different operator repeated the same HRM/PCR experiments in a different laboratory using the same model of real-time PCR instrument (BioRad CFX real-time PCR System). The melting temperatures obtained from the two independent experiments are highly correlated: they vary for less than 0.5°C, that is the sensitivity of the instrument used for measuring the melting temperature. Consequently, the assignment of the MT resulted perfectly coherent. This result clearly shows that this HLMT protocol and the MT assignment is reproducible and portable, thus the results obtained from different laboratories using the same protocol and the BioRad CFX real-time PCR instrument are comparable.

To test the possible cross-reaction with other bacterial species we tested the primers on DNA extracted from 16 *S. marcescens*-free fecal samples previously subjected to V3-V4 16S rRNA metagenomics by D'Auria et al. 2021. These fecal samples contained several bacterial species, including *Enterobacteriales* (the same order of *S. marcescens*) but not *S. marcescens*. The absence of bands with the expected size on the PCR gels clearly indicates the absence of cross-reaction with the numerous bacterial species present in the samples. Moreover, PCR amplification experiments on DNA extracted from strains of *Serratia rubidaea*, *Serratia ficaria*, *Serratia liquefaciens*, and *Serratia odorifera*, show the absence of cross-reactions with these other *Serratia* species. It has to be noted that *in silico* analysis of the designed primers suggested a possible cross-reaction with *Serratia ureilytica*, a species rarely isolated in hospital context but also able to cause nosocomial outbreaks.[33] Unfortunately, no sample for this species was available in our collection to be tested *in vitro*.

The HLMT protocol presented in this work can be used to detect and type *S. marcescens* from DNA directly extracted from samples, without the need of an isolation step. Indeed, the experiment is run on a qPCR real-time instrument, that can quantify the bacterium in the sample (as low as 10-10^4 units) and classify it as "MT Infectious," "MT Colonizing/Environmental," or "MT Unspecific." The main advantages of this cultivation-free approach are that: (1) it can give a rapid response about the presence of the bacterium in the sample, saving precious time e.g., during a nosocomial outbreak; and (2) it can be applied in large screening (e.g., hospital environmental screening) saving the resources and time needed to perform the cultivation step.

The identification of strains belonging to the Infectious clone by WGS approach requires a cultivation step, DNA sequencing and complex bioinformatic analyses. Furthermore, WGS-based typing is also time and cost consuming, requiring days and 50$–100$ per sample. Using the HLMT protocol proposed here, it is possible to identify Infectious clone strains just performing a real-time PCR/HRM assay, an experiment that requires a few hours and ~5$ per sample. This makes the protocol suitable for hospital surveillance programs and large epidemiological screenings, in particular in lower- or middle-income countries.

Limitations of the study

The presented cultivation-free HLMT protocol employs two HRM/PCR primer pairs on *S. marcescens*-specific genes. Although both *in silico* and *in vitro* analyses and experiments demonstrate the absence of cross-reaction of these primers with four other *Serratia* species, a possible cross-reaction with *Serratia ureilytica* was noticed by *in silico* analysis. This species causes nosocomial outbreaks, but it's seldom isolated in hospitals. However, we could not test *in vitro* this *Serratia* species to check for this possible cross-reaction.

Some isolates were classified as "MT Unspecific" or "Undetermined" and, among the "MT Unspecific" strains, several showed signals of possible HGT for one or both the target genes. These results suggest that HGT events do not reduce the detecting capability of the protocol, but they could affect the typing results. The protocol presented in this work has been tested on the BioRad CFX Connect real-time PCR instrument, which has an HRM temperature sensitivity of 0.5°C. Although this is one of the most diffused/cheap PCR/HRM instruments, others with higher sensitivity are available. As shown in Pasala et al., 2020,[19] the model of the PCR/HRM instrument affects the measured HRM temperatures, while the operator and laboratory have non-significative effects. As a consequence, the discriminatory power and portability of the presented HRM protocol refers only to the use of BioRad CFX Connect real-time PCR instrument. Other studies are necessary to test the discriminatory power and portability using other instruments.

It is to be noticed that the qPCR on soil mixed samples evidenced some technical problems, likely associated with the efficiency of DNA extraction, and/or with the presence of strong inhibitors of DNA amplification in this specific sample. For this reason, when dealing with complex and "difficult" matrices, we suggest our protocol to be used for detection and just preliminary typing. In these cases, a more precise typing result would be obtained on the isolated strains.

WGS remains the gold standard for the typing of the *S. marcescens* infection-associated lineage, and other high throughput sequencing methods such as HiSST proven to be successful in typing *S. marcescens*.[15,16] However, the time needed to carry out the sequencing may represent an important drawback. Indeed, this species is able to spread very rapidly among patients and can exploit different sources to grow in the hospital environment. Therefore, the high specificity and sensitivity of our HLMT protocol, combined with its speed of realization and cost-effectiveness, make it a valuable alternative to WGS-based methods, especially in low-middle income countries and in contexts where rapid and/or very large screenings are needed.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact

- Materials availability
- Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Cultivation-free HLMT protocol design
  - Whole Genome Sequencing (WGS) of the Study dataset
  - Global genomic dataset setup and WGS-based typing
  - HLMT protocol typing capability
  - HLMT protocol portability
  - HLMT protocol specificity for *S. marcescens* detection
  - HLMT protocol limit of detection
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Primer design
  - HRM Clusters grouping

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.109402.

ACKNOWLEDGMENTS

We want to thank Fondazione Romeo ed Enrica Invernizzi and the PLINIANA S.P.A. for supporting the project. The project was partially funded also by the University of Milan with Line 2 Action A grant. Lastly, we want to thank reviewers for their suggestions.

AUTHOR CONTRIBUTIONS

Experimental design: F.C., A.P., and S.Panelli; bioinformatics analyses: F.C., T.N., M.P., R.N., and L.S.; molecular biology experiments: A.A., S.Papaleo, and A.R.P.; paper writing: F.C., A.A.; sample provisioning: C.P., C.M., D.C., E.O., S.B., M.L.F., F.S., S.G.R., M.C., A.C., P.P., C.F., D.M.C.; funding and supervision: F.C. and G.Z.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 25, 2023
Revised: January 8, 2024
Accepted: February 28, 2024
Published: March 1, 2024

REFERENCES

1. Adeolu, M., Alnajar, S., Naushad, S., and S Gupta, R. (2016). Genome-based phylogeny and taxonomy of the "Enterobacteriales": proposal for Enterobacterales ord. nov. divided into the families Enterobacteriaceae, Erwiniaceae fam. nov., Pectobacteriaceae fam. nov., Yersiniaceae fam. nov., Hafniaceae fam. nov., Morganellaceae fam. nov., and Budviciaceae fam. nov. Int. J. Syst. Evol. Microbiol. *66*, 5575–5599.
2. Khanna, A., Khanna, M., and Aggarwal, A. (2013). Serratia marcescens- a rare opportunistic nosocomial pathogen and measures to limit its spread in hospitalized patients. J. Clin. Diagn. Res. *7*, 243–246.
3. Abreo, E., and Altier, N. (2019). Pangenome of Serratia marcescens strains from nosocomial and environmental origins reveals different populations and the links between them. Sci. Rep. *9*, 46.
4. Bush, L.M., and Vazquez-Pertejo, M.T.. Klebsiella, Enterobacter, and Serratia Infections. MSD Manual Professional Edition. https://www.merckmanuals.com/professional/infectious-diseases/gram-negative-bacilli/klebsiella,-enterobacter,-and-serratia-infections.
5. Kim, S.B., Jeon, Y.D., Kim, J.H., Kim, J.K., Ann, H.W., Choi, H., Kim, M.H., Song, J.E., Ahn, J.Y., Jeong, S.J., et al. (2015). Risk factors for mortality in patients with Serratia marcescens bacteremia. Yonsei Med. J. *56*, 348–354.
6. Saralegui, C., Ponce-Alonso, M., Pérez-Viso, B., Moles Alegre, L., Escribano, E., Lázaro-Perona, F., Lanza, V.F., de Pipaón, M.S., Rodríguez, J.M., Baquero, F., and Del Campo, R. (2020). Genomics of Serratia marcescens Isolates Causing Outbreaks in the Same Pediatric Unit 47 Years Apart: Position in an Updated Phylogeny of the Species. Front. Microbiol. *11*, 451.
7. Sterzi, L., Nodari, R., Di Marco, F., Ferrando, M.L., Saluzzo, F., Spitaleri, A., Allahverdi, H., Papaleo, S., Panelli, S., Rimoldi, S.G., et al. (2023). Genetic barriers more than ecological adaptations shaped Serratia marcescens diversity. Preprint at bioRxiv. https://doi.org/10.1101/2023.07.14.548978.
8. Rajapaksha, P., Elbourne, A., Gangadoo, S., Brown, R., Cozzolino, D., and Chapman, J. (2019). A review of methods for the detection of pathogenic microorganisms. Analyst *144*, 396–411.
9. Shi, Z.Y., Liu, P.Y., Lau, Y.J., Lin, Y.H., and Hu, B.S. (1997). Use of pulsed-field gel electrophoresis to investigate an outbreak of Serratia marcescens. J. Clin. Microbiol. *35*, 325–327.
10. Rödel, J., Mellmann, A., Stein, C., Alexi, M., Kipp, F., Edel, B., Dawczynski, K., Brandt, C., Seidel, L., Pfister, W., et al. (2019). Use of MALDI-TOF mass spectrometry to detect nosocomial outbreaks of Serratia marcescens and Citrobacter freundii. Eur. J. Clin. Microbiol. Infect. Dis. *38*, 581–591.
11. Phan, H.T.T., Stoesser, N., Maciuca, I.E., Toma, F., Szekely, E., Flonta, M., Hubbard, A.T.M., Pankhurst, L., Do, T., Peto, T.E.A., et al. (2018). Illumina short-read and MinION long-read WGS to characterize the molecular epidemiology of an NDM-1 Serratia marcescens outbreak in Romania. J. Antimicrob. Chemother. *73*, 672–679.
12. Gona, F., Comandatore, F., Battaglia, S., Piazza, A., Trovato, A., Lorenzin, G., Cichero, P., Biancardi, A., Nizzero, P., Moro, M., and Cirillo, D.M. (2020). Comparison of

core-genome MLST, coreSNP and PFGE methods for Klebsiella pneumoniae cluster analysis. Microb. Genom. *6*, e000347. https://doi.org/10.1099/mgen.0.000347.

13. Tran, A., Alby, K., Kerr, A., Jones, M., and Gilligan, P.H. (2015). Cost Savings Realized by Implementation of Routine Microbiological Identification by Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry. J. Clin. Microbiol. *53*, 2473–2479.
14. Jiang, F., Kong, Z., Cheng, C., Kang, H., Gu, B., and Ma, P. (2019). Overestimated discriminatory power of MALDI-TOF mass spectrometry for typing of carbapenem-resistant clones. Epidemiol. Infect. *147*, e324.
15. Bourdin, T., Monnier, A., Benoit, M.È., Bédard, E., Prévost, M., Quach, C., Déziel, E., and Constant, P. (2021). A High-Throughput Short Sequence Typing Scheme for Serratia marcescens Pure Culture and Environmental DNA. Appl. Environ. Microbiol. *87*, e0139921.
16. Bourdin, T., Benoit, M.È., Monnier, A., Bédard, E., Prévost, M., Charron, D., Audy, N., Gravel, S., Sicard, M., Quach, C., et al. (2023). Serratia marcescens Colonization in a Neonatal Intensive Care Unit Has Multiple Sources, with Sink Drains as a Major Reservoir. Appl. Environ. Microbiol. *89*, e0010523.
17. Sabat, A.J., Budimir, A., Nashev, D., Sá-Leão, R., van Dijl, J.M., Laurent, F., Grundmann, H., and Friedrich, A.W.; ESCMID Study Group of Epidemiological Markers ESGEM (2013). Overview of molecular typing methods for outbreak detection and epidemiological surveillance. Euro Surveill. *18*, 20380.
18. Słomka, M., Sobalska-Kwapis, M., Wachulec, M., Bartosz, G., and Strapagiel, D. (2017). High Resolution Melting (HRM) for High-Throughput Genotyping-Limitations and Caveats in Practical Case Studies. Int. J. Mol. Sci. *18*, 2316. https://doi.org/10.3390/ijms18112316.
19. Pasala, A.R., Perini, M., Piazza, A., Panelli, S., Di Carlo, D., Loretelli, C., Cafiso, A., Inglese, S., Gona, F., Cirillo, D.M., et al. (2020). Repeatability and reproducibility of the wzi high resolution melting-based clustering analysis for Klebsiella pneumoniae typing. Amb. Express *10*, 217.
20. Grando, D., Said, M.M., Mayall, B.C., and Gurtler, V. (2012). High resolution melt analysis to track infections due to ribotype 027 Clostridium difficile. J. Microbiol. Methods *89*, 87–94.
21. Ngui, R., Lim, Y.A.L., and Chua, K.H. (2012). Rapid detection and identification of human hookworm infections through high resolution melting (HRM) analysis. PLoS One *7*, e41996.
22. Tamburro, M., and Ripabelli, G. (2017). High Resolution Melting as a rapid, reliable, accurate and cost-effective emerging tool for genotyping pathogenic bacteria and enhancing molecular epidemiological surveillance: a comprehensive review of the literature. Ann Ig *29*, 293–316. https://doi.org/10.7416/ai.2017.2153.
23. Perini, M., Piazza, A., Panelli, S., Papaleo, S., Alvaro, A., Vailati, F., Corbella, M., Saluzzo, F., Gona, F., Castelli, D., et al. (2022). Hypervariable-Locus Melting Typing: a Novel Approach for More Effective High-Resolution Melting-Based Typing, Suitable for Large Microbiological Surveillance Programs. Microbiol. Spectr. *10*, e0100922.
24. Perini, M., Piazza, A., Panelli, S., Di Carlo, D., Corbella, M., Gona, F., Vailati, F., Marone, P., Cirillo, D.M., Farina, C., et al. (2020). EasyPrimer: user-friendly tool for pan-PCR/HRM primers design. Development of an HRM protocol on wzi gene for fast Klebsiella pneumoniae typing. Sci. Rep. *10*, 1307.
25. Perini, M., Batisti Biffignandi, G., Di Carlo, D., Pasala, A.R., Piazza, A., Panelli, S., Zuccotti, G.V., and Comandatore, F. (2021). MeltingPlot, a user-friendly online tool for epidemiological investigation using High Resolution Melting data. BMC Bioinf. *22*, 76.
26. Madden, T. (2003). The BLAST Sequence Analysis Tool (National Center for Biotechnology Information (US)).
27. Buffet-Bataillon, S., Rabier, V., Bétrémieux, P., Beuchée, A., Bauer, M., Pladys, P., Le Gall, E., Cormier, M., and Jolivet-Gougeon, A. (2009). Outbreak of Serratia marcescens in a neonatal intensive care unit: contaminated unmedicated liquid soap and risk factors. J. Hosp. Infect. *72*, 17–22.
28. Lompo, P., Agbobli, E., Heroes, A.-S., Van den Poel, B., Kühne, V., Kpossou, C.M.G., Zida, A., Tinto, H., Affolabi, D., and Jacobs, J. (2023). Bacterial contamination of antiseptics, disinfectants, and hand hygiene products used in healthcare settings in low- and middle-income countries—A systematic review. Hygiene (Basel) *3*, 93–124.
29. Johnson, J., and Quach, C. (2017). Outbreaks in the neonatal ICU: a review of the literature. Curr. Opin. Infect. Dis. *30*, 395–403.
30. Piccirilli, A., Cherubini, S., Brisdelli, F., Fazii, P., Stanziale, A., Di Valerio, S., Chiavaroli, V., Principe, L., and Perilli, M. (2022). Molecular Characterization by Whole-Genome Sequencing of Clinical and Environmental Strains Isolated during an Outbreak in a Neonatal Intensive Care Unit (NICU). Diagnostics *12*, 2180. https://doi.org/10.3390/diagnostics12092180.
31. Cristina, M.L., Sartini, M., and Spagnolo, A.M. (2019). Infections in Neonatal Intensive Care Units (NICUs). Int. J. Environ. Res. Public Health *16*, 610. https://doi.org/10.3390/ijerph16040610.
32. Guel-Gomez, M., Angulo-Zamudio, U.A., Leon-Sicairos, N., Flores-Villaseñor, H., Mendívil-Zavala, E., Plata-Guzmán, A., Martinez-Garcia, J.J., Angulo-Rocha, J., Ochoa-Espinoza, R., Crespo-Palazuelos, P., et al. (2023). Outbreak of in the Neonatal Intensive Care Unit of a Tertiary Care Hospital in Mexico. Adv. Met. Med. *2023*, 3281910.
33. Dahdouh, E., Lázaro-Perona, F., Ruiz-Carrascoso, G., Sánchez García, L., Saenz de Pipaón, M., and Mingorance, J. (2021). Intestinal Dominance by Serratia marcescens and Serratia ureilytica among Neonates in the Setting of an Outbreak. Microorganisms *9*, 2271. https://doi.org/10.3390/microorganisms9112271.
34. Hyatt, D., Chen, G.-L., Locascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinf. *11*, 119.
35. Li, L., Stoeckert, C.J., Jr., and Roos, D.S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. *13*, 2178–2189.
36. Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797.
37. Wall, L., Christiansen, T., and Orwant, J. (2000). Programming Perl (O'Reilly Media, Incorporated).
38. R Core Team (2018). R: A language and environment for statistical computing (Vienna, Austria: R Foundation for Statistical Computing). https://www.R-project.org/.
39. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312–1313.
40. Darriba, D., Posada, D., Kozlov, A.M., Stamatakis, A., Morel, B., and Flouri, T. (2020). ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. Mol. Biol. Evol. *37*, 291–294.
41. Darling, A.C.E., Mau, B., Blattner, F.R., and Perna, N.T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. *14*, 1394–1403.
42. Babraham bioinformatics - FastQC A quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
43. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. *19*, 455–477.
44. Olson, R.D., Assaf, R., Brettin, T., Conrad, N., Cucinell, C., Davis, J.J., Dempsey, D.M., Dickerman, A., Dietrich, E.M., Kenyon, R.W., et al. (2023). Introducing the Bacterial and Viral Bioinformatics Resource Center (BV-BRC): a resource combining PATRIC, IRD and ViPR. Nucleic Acids Res. *51*, D678–D689. https://doi.org/10.1093/nar/gkac1003.
45. Revell, L.J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. *3*, 217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x.
46. Ripley, B.D. (2001). The R project in statistical computing. MSOR Connect *1*, 23–25.
47. Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. *49*, W293–W296.
48. Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. J. Stat. Softw. *28*, 1–26.
49. D'Auria, E., Panelli, S., Lunardon, L., Pajoro, M., Paradiso, L., Beretta, S., Loretelli, C., Tosi, D., Perini, M., Bedogni, G., et al. (2021). Rice flour fermented with Lactobacillus paracasei CBA L74 in the treatment of atopic dermatitis in infants: A randomized, double- blind, placebo- controlled trial. Pharmacol. Res. *163*, 105284.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
SsoAdvanced Universal SYBR® Green Supermix	Bio-Rad	1725274
PCR amplification-grade water	Promega	DW0991
Roche Diagnostics KAPA HiFi HotStart ReadyMix	Roche Diagnostics	07958927001 | KK2601
Critical commercial assays		
Qiagen QIAcube Connect automated extractor	Qiagen	9002864
Qiagen DNeasy Blood & Tissue reagents	Qiagen	69504
Illumina NextSeq platform	Illumina	N/A
Oligonucleotides		
Primers for *S.marcescens* detection and typing	This manuscript	Ordered from Eurofins genomics
Software and algorithms		
BV-BRC database	NIAID Bioinformatics Resource Centers	https://www.bv-brc.org/
Prodigal	Hyatt et al.[34]	Version 2.6.3
OrthoMCL	Li et al.[35]	Version 2.0.9
Muscle	Edgar[36]	Version 5
Perl	Wall et al.[37]	Version 5.38.0
R	Ripley[38]	Version 4.3.1
RAxML8	Stamatakis[39]	Version 8
ModelTest-NG	Darriba et al.[40]	Version 0.1.7
Mauve	Darling et al.[41]	Version 2.4.0
Purple	Gona et al.[12]	Version 1.23
EasyPrimer	Perini et al.[24]	https://doi.org/10.1038/s41598-020-57742-z
MeltingPlot	Perini et al.[25]	https://doi.org/10.1186/s12859-021-04020-y
FastQC	Babraham bioinformatics[42]	Version 0.12.1
BLAST	NCBI[26]	https://blast.ncbi.nlm.nih.gov/Blast.cgi#
SPAdes	Bankevich[43]	Version 3.15.4
Bio-Rad CFX Manager™ Software	Bio-Rad	1845000
Deposited data		
Serratia marcescens genome assemblies	This manuscript	NCBI - PRJNA957961
Other		
BioRad CFX Connect real-time PCR System	Bio-Rad	1855201
Bio-Rad T100 Thermal cycler	Bio-Rad	1861096

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr. Francesco Comandatore (francesco.comandatore@unimi.it).

Materials availability

Primer sequences generated in this study are available in this manuscript and by request by the lead contact, Dr. Francesco Comandatore (francesco.comandatore@unimi.it).

Data and code availability

- All the data published in this paper will be available from the lead contact upon request. Genome assemblies have been deposited on the NCBI database and are publicly available as of the date of publication. The accession number is listed in the key resources table.
- The EasyPrimer and the MeltingPlot tools codes have been deposited on github (https://github.com/MatteoPS/EasyPrimer; https://github.com/MatteoPS) and are publicly available as of the date of publication. The corresponding DOI are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

The specificity and typing capability of the protocol was evaluated on a dataset including 230 *Serratia marcescens* strains (from here "Study dataset"), one *Serratia rubidaea*, one *Serratia ficaria*, one *Serratia liquefaciens* and one *Serratia odorifera*. The 230 *S. marcescens* isolates of the Study dataset were collected from seven hospitals and one community health institution: 71 isolates from ASST Papa Giovanni XXIII Hospital in Bergamo, 66 from IRCCS San Raffaele Hospital (HSR) in Milan, 36 from IRCCS San Matteo Hospital in Pavia, 28 from ASST Fatebenefratelli Sacco Hospital in Milan, 15 from ASST San Gerardo Hospital in Monza, 11 from ASST Fatebenefratelli Buzzi Hospital in Milan, three from the Instituto Zooprofilattico di Pavia. The metadata of the strains are reported in Table S1. The *S. rubidaea*, *S. ficaria*, *S. liquefaciens* and *S. odorifera* were retrieved from the collection of the IRCCS San Matteo Hospital in Pavia.

The *S. marcescens* isolates were seeded in McConkey agar medium and incubated overnight at 37°C. The following day, single colonies of the isolates were dissolved in 5 ml of LB medium and incubated for 22 hours. DNA extraction was then carried out using a Qiagen QIAcube Connect automated extractor (Qiagen, Hilden, Germany) following the "bacterial pellet" protocol which employs Qiagen DNeasy Blood & Tissue reagents.

METHOD DETAILS

Cultivation-free HLMT protocol design

To design a cultivation-free High Resolution Melting (HRM)-based typing protocol using the Hypervariable-Locus Melting Typing (HLTM) approach, we performed an *ad hoc* genomic analysis for the identification of target genes that are: i) present in all the *S. marcescens* strains; ii) rarely present (or absent) in other species; iii) characterized by a high level of genetic variability; iv) have a phylogenetic signal coherent with the *S. marcescens* species tree; v) contain DNA regions suitable for HRM/PCR primers design (hyper-variable regions flanked by conserved regions).

The gene selection started from the *S. marcescens* genome assembly collection of the Bacterial and Viral Bioinformatics Resource Center database (BV-BRC https://www.bv-brc.org/, formerly PATRIC database)[44] retrieved on 23th May 2018. More in detail, the 223 genome assemblies (categories "WGS" or "Complete") associated with a published paper and classified by the database as quality "Good" were selected for the further analyses. Each of the selected genomes was then subjected to the Prodigal tool for Open Reading Frame (ORF) calling.[34] The obtained ORFs were then grouped in ortholog clusters using the OrthoMCL tool[35] and the core clusters (those present in all the 223 assemblies in single copy) were selected. For each core cluster, one representative nucleotide sequence was used to perform a BLASTn[26] search against the Nucleotide database (nt) (e-value threshold $< 10^{-5}$ and % identity > 90), excluding any match with *S. marcescens* sequences. The sequences with less than 10 BLASTn[26] hits were then manually assessed to select *S. marcescens* sequences rarely present in other species.

For each selected core cluster, the genetic variability, the co-cladogenesis with the *S. marcescens* species tree and the possibility to design HRM primers were considered for the further selection step. To evaluate the genetic variability, the sequences of each selected core cluster were aligned using Muscle, then the alignment length and average nucleotide distance were computed using Perl and R, respectively.[35–37] The co-cladogenesis was evaluated using the Cophylo R library.[45] More in detail, the Maximum Likelihood phylogenetic analysis for each selected cluster was carried out using RAxML8 (with 100 pseudo bootstraps) after best evolutionary model selection using the ModelTest-NG tool.[39,40] The *S. marcescens* species phylogenetic tree was obtained using the SNP-based approach: the 223 genome assemblies (see above) were aligned against the reference genome of *S. marcescens* Db11 (GCF_000513215.1) using Mauve tool,[41] core SNPs were called using Purple tool[12] and phylogenetic analysis was carried out using RAxML8 as described above. Lastly, for each selected core cluster the possibility to design HRM/PCR primers on the alignments was manually evaluated using the EasyPrimer tool.[24] The obtained results were then manually evaluated to select the target clusters and HRM/PCR primers were then manually designed using the outputs of EasyPrimer tool.[24]

Whole Genome Sequencing (WGS) of the Study dataset

Libraries were prepared and 2 × 150 bp paired-end run sequencing was carried out on the Illumina NextSeq platform. The obtained reads were quality checked using FastQC tool and assembled using SPAdes.[42,43]

Global genomic dataset setup and WGS-based typing

The 230 *S. marcescens* strains of the Study dataset were added to a larger Global dataset (including a total of 516 genomes) and classified using a WGS-based phylogenetic approach as follows. The Global genomes dataset included: i) the 230 *S. marcescens* isolate genomes of the

Study dataset; ii) the 223 *S. marcescens* genome assemblies used for the protocol design; iii) the 63 genomes of the dataset previously included in the studies by Abreo and colleagues[3] or by Saralegui and colleagues that are not included in the genome dataset previously retrieved from the BV-BRC database (see above); iv) the genome of *Serratia odorifera* FDAARGOS_353 (GCF_002206385.2), used as outgroup. For the dataset analyzed by Abreo and colleagues (2019), when only the reads were available, the reads files were downloaded and assembled using the SPAdes tool.[43] See Table S1 for details.

The genome assemblies were aligned against the genome of the reference strain *S. marcescens* Db11 (GCF_000513215.1) using Mauve tool[41] and core SNPs were called using Purple tool.[12,41] The core SNPs alignment was then subjected to Maximum Likelihood phylogenetic analysis, with 100 pseudo bootstraps, using RAxML8 setting the TVM+G evolutionary model, as selected by the ModelTest-NG tool.[39,40] The obtained tree was then rooted on the outgroup strain (*S. odorifera*), which was then removed from the tree. Furthermore, the *S. marcescens* strains were grouped on the basis of core SNP distance using the unsupervised clustering method K-means, after silhouette analysis for the determination of the optimal number of clusters. The analyses were performed using R.[46]

The strains for which the isolation material was reported, were classified as "Infection", "Colonization", "Environmental", "Hospital environment" (see Table S1). More in detail, among the strains isolated from humans, those from body compartments usually sterile (e.g. blood or urine) were classified as "Infection" while those from human compartments usually colonized by bacteria (e.g. gut) were classified as "Colonization". The strains isolated from animals, plants, soil or water were classified as "Environmental", while those isolated from the hospital environment were classified as "Hospital environment".

HLMT protocol typing capability

HRM assays were performed on the genomic DNA extracted from each of the 230 *S. marcescens* isolates of the Study dataset. Each 15 μL reaction mix contained 7.5 μL of 2x SsoAdvanced Universal SYBR® Green Supermix ((BioRad, Hercules, California), 0.4 μL of each primer (see Table 1) (10 μM), 5,7 μL of PCR amplification-grade water (Promega Corporation, Wisconsin, USA) and 1 μL of *S. marcescens* DNA. Each *S. marcescens* sample was qPCR-amplified in three technical replicates. The qPCR assays were performed on a BioRad CFX Connect real-time PCR System (BioRad, Hercules, California). The thermal profile for both primer pairs was as follows: 98°C for 2 min and 40 cycles of [95°C for 7 s, 63°C for 7 s, and 72°C for 15 s]. Both profiles were characterized by HRM with an initial denaturation (95°C for 2 min) followed by 0.5°C increments with fluorescence data acquisition from 70°C to 95°C.

The resulting melting temperatures were then used to group the 230 *S. marcescens* strains in HLMT clusters using the MeltingPlot tool and to determine the temperature ranges to be used to define the Melting Types.[25] The instrument used to perform the HRM experiments (BioRad CFX Connect real-time PCR instrument, BioRad, Hercules, California) has a temperature sensitivity of 0.5°C, thus it is not able to distinguish two strains for which the melting temperatures differ less than 0.5°C for all the HRM primers used in the typing protocol. Using MeltingPlot[25] we built a network including all the 230 Study dataset strains connecting the indistinguishable pairs. The strains were then grouped in HRM Clusters by the MeltingPlot tool on the basis of the network topology (the clustering algorithm is described in detail in Perini et al.[25] The HRM Clusters were then mapped on the coreSNP-based phylogenetic tree (see above) using iTol[47] and the HRM Cluster(s) associated with the clinical lineage were identified. These results were then used to identify the melting temperatures associated with specific lineages and to define the Melting Types (MTs). Lastly, the sensitivity and specificity of the HLMT protocol were determined using the caret R library.[48]

HLMT protocol portability

All the HRM assays described so far have been performed by several operators from the University of Milan using a CFX Connect real-time PCR instrument (BioRad, Hercules, California) placed in a laboratory of this institution. The reproducibility and portability of the protocol was evaluated by repeating the experiments on a subset of 16 strains, by a different operator in a different laboratory (at the Hospital San Raffaele), using the same real-time PCR instrument model and the same protocol described above. The 16 strains were selected to be representative of the different Melting Types (see above). The melting temperatures obtained from the two laboratories were then compared by linear regression analysis and passed to the MeltingPlot tool for Melting Type assignment.

HLMT protocol specificity for *S. marcescens* detection

The specificity of the designed primers for *S. marcescens* were tested *in silico* and *in vitro*. The *in silico* test was carried out using Primer-BLAST,[26] selecting matches with amplicon size that differ less than 20% from the expected length (see Table 1).

The *in vitro* tests were carried out on four samples groups: i) 16 *S. marcescens*-free fecal samples; ii) isolates of five *Serratia* species (including *S. marcescens*); iii) five *S. marcescens* isolates representative of the 230 Study dataset; iv) five *S. marcescens*-free fecal samples added with decreasing concentrations of *S. marcescens* (from here, mixed samples); v) DNA extracted from mixed samples obtained by spiking a *S. marcescens-free* soil sample with different amounts of *S. marcescens* cells from overnight cultures. More in detail, the extracted DNA of 16 *S. marcescens*-free fecal samples were retrieved exploiting the D'Auria et al. 2021 collection.[49] The D'Auria et al. study was aimed to investigate the effects of a probiotic to atopic dermatitis in infants (58 individuals aged between 6 and 36 months), also by V3-V4 16S rRNA metagenomics. All the analyzed samples were *S. marcescens*-free, thus we decided to exploit this DNA collection (the same tubes of extracted DNA) to test the specificity of the designed primers. More in detail, the 16 DNA samples were selected as follows: eight samples with the highest relative abundance of *Enterobacteriaceae* and eight with the lowest. As stated in D'Auria et al.[49] DNA was extracted from 250mg of fecal material using the QIAamp PowerFecal Pro DNA Kit (QIAGEN). The total DNAs were then diluted 1:50 in Milli-q water

and amplified through PCR using the *S. marcescens* primers and the protocol described in this study. A *S. marcescens* genomic DNA was employed as positive control and the PCR mix was composed of 10 μL of Roche Diagnostics KAPA HiFi HotStart ReadyMix (Roche Diagnostics, Risch-Rotkreuz, Switzerland), 1 μL of each primer (10 μM), 7 μL of Promega PCR amplification-grade water (Promega Corporation, Madison, Wisconsin, USA) and 1 μL of *Serratia* DNA, corresponding to about 5-50 ng of DNA (per sample). The PCR mix of fecal samples included 10 μL of Roche Diagnostics KAPA HiFi HotStart ReadyMix, 1 μL of each primer (10 μM), 6 μL of Promega PCR amplification-grade water and 2 μL of the fecal sample DNA.

Five isolates of five *Serratia* species (*S. marcescens, S. rubidaea, S. ficaria, S. liquefaciens* and *S. odorifera*) were retrieved from the microbiology collection of IRCCS San Matteo Hospital and the genomic DNA was extracted as described above. DNA concentration was assayed with Qubit Fluorometer (Invitrogen, Waltham, Massachusetts, USA) and PCR experiments were carried out on DNA at 1 ng/μL concentration after dilutions with Milli-Q water.

Five out of the 230 *S. marcescens* Study dataset strains were manually selected on the basis of SNP-based phylogenetic tree (see above) to span the known genetic variability of the species. The genomic DNA of the five selected strains were subjected to PCR experiments as described above for the other *Serratia* species strains. The PCR experiments were performed with the same mix of the *S. marcescens* positive control described above.

The specificity of the protocol was also tested by PCR experiments on two types of mixed samples: i) fecal DNA + *S. marcescens* DNA; ii) fresh soil + intact *S. marcescens* cells. For the former, a mixed DNA sample was prepared for each of the five *S. marcescens* strains selected above, adding the bacterium DNA to an *Enterobacteriaceae*-rich fecal sample. More in detail, for each isolate, the PCR test was composed by the *S. marcescens* DNA at 1 ng/μL concentration (equivalent to the order of 10^5 genome copies) as a positive control, the fecal sample added with the *S. marcescens* sample at 0.1 ng/μL and 0.01 ng/μL concentration (equivalent to the order of 10^4 and 10^3 genome copies, respectively), and the fecal sample as a negative control. For mixed samples, the PCR mix was composed of 10 μL of Roche Diagnostics KAPA HiFi HotStart ReadyMix, 1 μL of each primer (10 μM), 5 μL of Promega PCR amplification-grade water, 2 μL of fecal sample DNA (5-50 ng) and 1 μL of *S. marcescens* DNA. Negative control and positive control PCR mixes were the same described for *S.marcescens*-free samples and *Serratia* species strains, respectively. For the second type of mixed samples, fresh soil was collected in Milan. 250 mg-aliquots of this soil were spiked with decreasing amounts (from 10^6 to 10^3) intact *S. marcescens* cells cultured overnight in LB broth as described above. A soil aliquot, not spiked with *S. marcescens* cells, was preserved for a double experimental verification of the absence of *S. marcescens* from this soil, carried out both by means of the two PCR protocols developed in this paper and by a V3-V4 16S rRNA amplicon metagenomics analysis. After DNA extraction using the DNeasy Powersoil kit (Qiagen, Hilden, Germany), mixed soil samples were PCR amplified as described above for the mixed fecal samples. The "native" soil sample, not spiked with *S. marcescens* cells, was used as a negative control. The positive control was the same used for the mixed fecal samples (see above).

For each of the above-mentioned PCR experiments, the thermal profile and agarose gel preparation were as follows. The PCR thermal profile for both sets of primers was the following: 95°C for 2 min, 35 cycles of [95°C for 30 s, 63°C for 45 s, 72°C for 30 s] and 72°C for 5 mins. The PCR reaction was performed using a Bio-Rad T100 Thermal cycler (BioRad, Hercules, California, USA). Finally, the results of the PCR amplification were analyzed through electrophoresis on 2.5% agarose gel.

HLMT protocol limit of detection

To evaluate the limit of detection of the HLMT primers, real-time PCR (qPCR) experiments were performed on both types of mixed samples. Mixed fecal samples, containing DNA of a *S. marcescens*-free fecal sample and DNA of a *S. marcescens* strain at 10 ng/μL, 1 ng/μL, 0.1 ng/μL, 0.01 ng/μL, 0.001 ng/μL and 0.0001 ng/μL concentrations, corresponding to the order of 10^6, 10^5, 10^4 , 10^3 , 10^2, and 10 genome copies. The qPCR experiment was carried out adding 1 μL of each *S. marcescens* dilution to 1 μL of fecal DNA. Each reaction mix was prepared as described above in a final volume of 15 μL. The undiluted *S. marcescens* isolate and a fecal DNA sample were used as positive and negative control, respectively. Two μL DNA extracted from each mixed soil sample were also subjected to qPCR. The reaction mix is the same as above as well as the positive control. The negative control here is the unspiked soil sample.

The qPCR was performed on a BioRad CFX Connect real-time PCR System, with the thermal profile described above for the HLMT protocol. The results of the experiment were visualized and retrieved from the Bio-Rad CFX Manager™ software and the statistical analyses were performed using R.[28]

QUANTIFICATION AND STATISTICAL ANALYSIS

Primer design

The possibility to design HRM/PCR primers on the clusters alignments was manually evaluated using the EasyPrimer tool. The acquired results were subsequently assessed manually to identify the desired clusters, and HRM/PCR primers were manually designed based on the findings from the EasyPrimer tool outputs.

HRM Clusters grouping

The melting temperatures resulting from the HRM assay were grouped with the graph-based clustering algorithm implemented in the MeltingPlot tool. Chi-Squared analysis showed a significant association (p-value < 2.2e-16) between the HRM Clusters and Phylo Clusters (residuals are shown in Figure S1B).