Genome Evolution and Plasticity of *Serratia marcescens*, an Important Multidrug-Resistant Nosocomial Pathogen

Atsushi Iguchi1,13, Yutaka Nagaya2,14, Elizabeth Pradel3,4,5,15, Tadasuke Ooka6, Yoshitoshi Ogura6,7, Keisuke Katsura7, Ken Kurokawa8, Kenshiro Oshima9, Masahira Hattori9, Julian Parkhill10, Mohamed Sebaihia10, Sarah J. Coulthurst11, Naomasa Gotoh3, Nicholas R. Thomson10,12,*, Jonathan J. Ewbank3,4,5, and Tetsuya Hayashi6,7,*

1Interdisciplinary Research Organization, University of Miyazaki, Japan
2Department of Microbiology and Infection Control Science, Kyoto Pharmaceutical University, Japan
3Centre d’Immunologie de Marseille-Luminy, UM2 Aix-Marseille Université, Marseille, France
4INSERM, U1104, Marseille, France
5CNRS, UMR7280, Marseille, France
6Department of Infectious Diseases, Faculty of Medicine, University of Miyazaki, Japan
7Department of Genomics and Bioenvironmental Science, Frontier Science Research Center, University of Miyazaki, Japan
8Earth-Life Science Institute, Tokyo Institute of Technology, Kanagawa, Japan
9Graduate School of Frontier Sciences, University of Tokyo, Chiba, Japan
10Pathogen Genomics, The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridge, United Kingdom
11Division of Molecular Microbiology, College of Life Sciences, University of Dundee, United Kingdom
12Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom
13Present address: Department of Animal and Grassland Sciences, Faculty of Agriculture, University of Miyazaki, Japan
14Present address: Kashima ONC QC, Oncology DCU, Eisai Demand Chain Systems, Eisai Co., Ltd., Ibaraki, Japan
15Present address: CIL-Inserm U1019, Institut Pasteur de Lille, Lille, France
*Corresponding author: E-mail: nrt@sanger.ac.uk, thayash@med.miyazaki-u.ac.jp.

Accepted: July 21, 2014

Data deposition: The sequence data for this study have been deposited at DDBJ/EMBL/GenBank under the accession AP013063, AP013064, AP013065, and HG326223.

Abstract

*Serratia marcescens* is an important nosocomial pathogen that can cause an array of infections, most notably of the urinary tract and bloodstream. Naturally, it is found in many environmental niches, and is capable of infecting plants and animals. The emergence and spread of multidrug-resistant strains producing extended-spectrum or metallo beta-lactamases now pose a threat to public health worldwide. Here we report the complete genome sequences of two carefully selected *S. marcescens* strains, a multidrug-resistant clinical isolate (strain SM39) and an insect isolate (strain Db11). Our comparative analyses reveal the core genome of *S. marcescens* and define the potential metabolic capacity, virulence, and multidrug resistance of this species. We show a remarkable intraspecies genetic diversity, both at the sequence level and with regards genome flexibility, which may reflect the diversity of niches inhabited by members of this species. A broader analysis with other *Serratia* species identifies a set of approximately 3,000 genes that characterize the genus. Within this apparent genetic diversity, we identified many genes implicated in the high virulence potential and antibiotic resistance of SM39, including the metallo beta-lactamase and multiple other drug resistance determinants carried on plasmid pSMC1. We further show that pSMC1 is most closely related to plasmids circulating in *Pseudomonas* species. Our data will provide a valuable basis for future studies on *S. marcescens* and new insights into the genetic mechanisms that underlie the emergence of pathogens highly resistant to multiple antimicrobial agents.

Key words: *Serratia marcescens*, genome plasticity, virulence, multidrug resistance.
**Materials and Methods**

**S. marcescens Strains, Culture Media, and Growth Conditions**

Strain SM39 was isolated from a septicemic patient in Japan in 1999 (Nakamura et al. 2002). Strain Db11, a kind gift from Dominique Ferrandon, is a spontaneous streptomycin-resistant derivative of strain Db10 which was isolated from a moribund *Drosophila melangaster* in Sweden (Flyg et al. 1980). It is available, together with Db10, from the Caenorhabditis Genetics Center (http://www.cbs.umn.edu/CGC, last accessed August 6, 2014). For routine bacterial cultivation, the strains were aerobically grown at 37°C in Luria–Bertani (LB) broth with shaking or on LB agar plates.

**Genome Sequencing, Gene Prediction, and Annotation**

Genomic DNA was prepared from overnight cultures of the two strains using the Genomic-tip 100G and Genomic DNA buffer set (Qiagen, Inc.) according to the manufacturer’s instructions. The genome of SM39 was shotgun sequenced to approximately 10-fold coverage from 102,218 end sequences of two genomic shotgun libraries based on pUC118 and pCC1BAC with average insert sizes of 3 and 10 kb, respectively. The genome of Db11 was shotgun sequenced to approximately 10-fold coverage from 90,142 end sequences from multiple shotgun libraries: pMAQ1Sac_BstXI (with insert sizes of 5.5–6, 9–10, and 10–12 kb), pUC19 (with insert sizes of 1.4–2 and 2–2.8 kb), pOTTW2 (with insert sizes of 3–3.3 and 2–2.8 kb), and M13mp18 (with insert sizes of 1–1.4, 0.4–0.8, 1.1–1.4, and 0.5–1 kb). This was supplemented by approximately 0.2-fold coverage from 1,713 end sequences derived from large insert fosmid libraries: pBACE3.6_BamHI with an insert size 18–23 kb. All sequencing was performed using big-dye terminator chemistry on ABI3730 or ABI3700 capillary automated sequencers (Applied Biosystems). All assemblies were generated using Phrap. All repeat regions, gaps and low quality regions were bridged using large insert fosmid libraries, read-pairs, or end-sequenced polymerase chain reaction (PCR) products. The sequences were manipulated to the community standard of “Finished” (Chain et al. 2009).

Nucleotide sequence position 1 of each chromosome was assigned according to that of published *Escherichia coli* genomes. Gene prediction and annotation were performed using Microbial Genome Annotation Pipeline (Sugawara et al. 2009), followed by manual curation on the basis of the results of BLASTP homology search against the public nonredundant protein database (http://www.ncbi.nlm.nih.gov/, last accessed August 6, 2014). The annotated genome sequences of SM39 and Db11 have been deposited to the DDBJ/EMBL/GenBank database under the accession numbers AP013063 for the SM39 chromosome, AP013064 for the
pSMC1 plasmid of SM39, AP013065 for pSMC2 of SM39, and HG326223 for the Db11 chromosome.

Genome-Wide Comparative Analysis

Intra- and interstrain clustering analyses of protein coding sequences (CDSs) were performed based on the results of all-to-all BLASTP analysis with a threshold of ≥90% amino acid sequence identity over ≥90% aligned length of the query sequence (supplementary fig. S1, Supplementary Material online, for details). In Silico Molecular Cloning (InSilico Biology, Yokohama, Japan) was used for genome sequence comparison. The average nucleotide identity (ANI) was calculated using the JSpecies software (Richter and Rossello-Mora 2009). Multiple alignments of sequences were constructed using the ClustalW program (Thompson et al. 1994) and a neighbor-joining tree was generated using the Tamura–Nei model in MEGAS (Tamura et al. 2007). Pathway analysis was performed using the KEGG (Kyoto Encyclopedia of Genes and Genomes) online database (Kanehisa and Goto 2000). Signal peptide-containing proteins were identified using the SignalP 4.1 server (http://www.cbs.dtu.dk/services/SignalP/, last accessed August 6, 2014) (Petersen et al. 2011). For the interspecies comparison, following four genomes were used: S. proteamaculans strain 568 (GenBank accession number CP000826) and S. plymuthica strains AS9 (CP002773) (Neupane et al. 2012) and 4xR13 (CP006250) (Weise et al. 2014), which were all isolated from plants, and S. odorifera strain DSM4582 (from human sputum) (ADBY00000000).

Semiautomated Screen for Mutants with an Attenuated Virulence Using Caenorhabditis elegans

A library of individual S. marcescens Db10 mini-Tn5Sm mutant clones was generated using standard methods. Using a COPAS Biosort (Union Biometica, Boston), 30 L4 stage wild type (N2) worms were sorted into the wells of 96-well plates as previously described (Garvis et al. 2009). Then, 30 μl of an overnight culture (30°C in 200 μl LB with 100 μg/ml streptomycin in 96-well plates) of individual mutant clones was transferred to the worm-containing wells. Assay plates containing worms and bacteria were incubated at 25°C with agitation, whereas overnight culture plates were kept at 4°C. A duplicate assay was performed the following day, again with 30 L4 worms per well, using the same bacterial cultures. For this, bacterial cultures were warmed up from 4°C to room temperature for at least an hour before use. For each assay and its duplicate, the number of worms alive in each well was scored each day for 5 days by examination with a dissecting microscope. Wells were scored as − (<5 worms alive), + (5–9 alive), ++ (10–14 alive), and +++ (at least 15 alive). Mutants were considered attenuated when they scored either ++ or +++ (at least 15 alive) after 3 days, and were subsequently assayed for worm killing on solid medium as previously described (Kurz et al. 2003). The mini-Tn5-Sm insertion site in attenuated mutants was determined by direct genomic DNA sequencing (MilleGen, Labège, France) with primers JEP131 and JEP132 (5′-CGGCCGACTTTGTTATAA-3′ and 5′-CTAGCGGGCCAGATCTGATCAG-3′, respectively).

Construction of KS3, a pSMC1-Cured Derivative of SM39

Because the pSMC1 plasmid of SM39 is highly stable, we could not obtain SM39-derivatives that have spontaneously lost pSMC1. To obtain a pSMC1-cured derivative of SM39, we first deleted the parABC genes (pSMC1_53-55) of pSMC1 using a method described by Masuda et al. (2000). The upstream and downstream parABC-flanking regions were amplified by PCR using primer pairs parAUF4/parAUR2 (5′-GGATCCAGCTGCTTCTCAGATGACCAGA-3′ and 5′-CAGAGAAACAAAGATAGATTCTCGTACAT-3′) and parADF2/parADR3 (5′-GCTAAAACCTATCTGTTGTTGTCCTGATTACCTCCT-3′ and 5′-CTGAGCGTCTGTGCACATCAGTG-3′), respectively. The PCR products were concatenated by fusion PCR with primers parAUF4 and parADR3. The concatenated DNA fragment was inserted into pLOI2223 to yield pPAR001, and then a NotI fragment of pMT5071, which encodes the Mob cassette, was inserted into pPAR001 to yield pPAR002. pPAR002 was mobilized conjugaly from E. coli strain S17-1 to S. marcescens SM39. Transconjugants, in which pPAR002 was inserted into pSMC1 (the first recombination), were selected on BM2 plates containing 150 μg/ml of chloramphenicol (Cm) and sucrose-sensitive clones were further screened using sucrose (10%)-containing and sucrose-free LB plates. Insertion of pPAR002 into pSMC1 was confirmed by PCR, and then pPAR002 excision (with or without the parABC genes) was selected on sucrose LB plates. Cm-susceptible clones were screened by PCR for the loss of the parABC genes. SM39-derivatives cured of pSMC1 were sensitive to HgCl2 (16 μg/ml in LB plates). The loss of pSMC1 was finally confirmed by plasmid profiling, and one of these derivatives (designated as KS3) was used for further analysis. See supplementary figure S2, Supplementary Material online, for the entire process and the composition of BM2 medium.

Antimicrobial Susceptibility Testing

Minimum inhibitory concentrations (MICs) of strains SM39, KS3, and Db11 against 36 antimicrobial agents were determined by the usual 2-fold agar dilution technique with Mueller-Hinton II agar (Becton Dickinson Microbiology Systems, Cockeysville, MD) with an inoculum size of 10^4 cells as described previously (Masuda et al. 2000).

Results and Discussion

Comparative Genomics of Human Clinical and Insect Pathogenic S. marcescens Strains SM39 and Db11

The chromosomes of SM39 and Db11 are similar in size, 5,225,577 and 5,113,802 bp, respectively, and gene
number (table 1 and fig. 1A). SM39 carries two plasmids, pSMC1 (41,517 bp) and pSMC2 (58,929 bp). No extrachromosomal elements were found in Db11. Overall, the chromosomes of the two strains are highly conserved and essentially collinear, except for a large inversion of the oriC region flanked by rRNA operons (fig. 1B). Surprisingly for members of the same species, the ANI between the conserved genomic regions of the two strains (4,391 kb) is only 95.1%. This can be compared with a figure of 97.0% for the ANI between two distantly related members of *E. coli*, strains K-12 and E2348/69. Despite this, the 16S rRNA genes of the two strains show 99.4% sequence identity, consistent with DB11 and SM39 being members of the same species (Kim et al. 2014; see also supplementary fig. S3, Supplementary Material online). Further with a criterion of ≤90% nucleotide sequence identity and ≥5 kb length, we identified 44 genomic regions specific to SM39, totaling 628 kb (12% of the genome), and 39 specific regions (447 kb or 8.7% of the genome) for Db11 (fig. 1A; see also supplementary tables S1 and S2, Supplementary Material online). The relatively low ANI value for the two *S. marcescens* strains and the considerable amount of isolate-specific sequences may reflect the diversity of niches that this species can occupy.

To extend the analysis of unique and shared genes between the two strains, we performed a clustering analysis of the SM39 and Db11 CDSs using an all-against-all BLASTP approach (outlined in supplementary fig. S1, Supplementary Material online). This showed that 3,970 genes (or gene families) were conserved in both strains with additional 860 SM39-specific and 728 Db11-specific genes (fig. 2A). Functional classification of these genes based on the Cluster of Orthologous Groups (COG) categories indicated that genes belonging to “category L (replication, recombination, and repair)” and “category V (defense mechanisms)” are more

### Table 1

Genomic Features of *Serratia marcescens* Strains SM39 and Db11

<table>
<thead>
<tr>
<th>Strains</th>
<th>SM39</th>
<th>Db11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromosome</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size (bp)</td>
<td>5,225,577</td>
<td>5,113,802</td>
</tr>
<tr>
<td>GC content</td>
<td>59.8</td>
<td>59.5</td>
</tr>
<tr>
<td>CDSs*</td>
<td>4,866 (19)</td>
<td>4,722 (15)</td>
</tr>
<tr>
<td>rRNA operons</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>tRNAs</td>
<td>87</td>
<td>87</td>
</tr>
<tr>
<td>Prophages</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Integrative elements</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>IS elements</td>
<td>17</td>
<td>5</td>
</tr>
<tr>
<td>Plasmid s</td>
<td>pSMC1 (41,517)</td>
<td>pSMC2 (58,929)</td>
</tr>
<tr>
<td>Size (bp)</td>
<td>41,517</td>
<td>58,929</td>
</tr>
<tr>
<td>GC content</td>
<td>61.5</td>
<td>51.9</td>
</tr>
<tr>
<td>CDSs*</td>
<td>55 (4)</td>
<td>72 (0)</td>
</tr>
<tr>
<td>IS elements</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

*The number of pseudogenes is indicated in parentheses.
highly represented in SM39 than in Db11 (fig. 2B). These differences are largely attributable to the presence of more genes associated with lateral gene transfer in SM39 such as insertion sequence (IS) transposases and integrases (category L), and genes for restriction-modification systems and multidrug transport systems (category V). For example, SM39 carries a higher variety and number of IS elements: 18 compared with the 5 carried by Db11 (table 1). These IS elements were classified into 11 distinct types, nine of which represent novel IS elements. Eight types were found in SM39 and four in Db11, with only ISSe5 being shared by both isolates (supplementary table S3, Supplementary Material online). More prophages were also found in SM39 (seven) than in Db11 (only two), with none of the prophages and other integrative elements being shared by the two strains (table 1 and supplementary table S4, Supplementary Material online).

We further compared the S. marcescens gene sets with those of four sequenced strains from other Serratia species: S. proteamaculans strain 568 and S. plymuthica strains AS9 (Neupane et al. 2012) and 4Rx13 (Weise et al. 2014), which were all isolated from plants, and S. odorifera strain DSM4582 (from human sputum). This showed that only 23% and 38% of the genes that we denoted as SM39- or Db11-specific had homologs in these isolates representing other Serratia species (fig. 2C). In contrast, 71% (2,852 genes) of the 3,970 core genes shared by the two S. marcescens strains were conserved in all the four isolates representing other Serratia species. This suggests that, despite the genetic diversity of this genus (see supplementary fig. S3, Supplementary Material online, for the phylogenetic relationship of these species in genus Serratia), there is a relatively large shared or core genome of 2,852 CDSs. This analysis also identified 358 genes that are conserved in the two S. marcescens strains but absent in other Serratia species. This group includes a number of genes related to pathogenicity in other bacteria (data not shown).

Common and Strain-Specific Metabolic Capabilities in S. marcescens SM39 and Db11

Central metabolic pathways are well conserved between the two strains and characteristic of the species as a whole: Positive fermentation for sucrose and D-sorbitol, and negative for D- and L-arabinose, L-rhamnose, D-xylose and cellubiose. Not uncommonly, both isolates lack the genes for prodigiosin biosynthesis, the characteristic red pigment often associated with S. marcescens (Harris et al. 2004). The differences between the two isolates include, for carbohydrate metabolism, alternative pathways for L-ascorbate utilization; the ula-type encoded pathway is carried by SM39 (SM39_4088-4095) and the sgb-type by Db11 (SMDB11_3334-3337). Major differences in nitrogen metabolism are also observed between
the two isolates: Unlike Db11, SM39 carries the nrtABC operon (SM39_1878-1880) encoding an ATP-binding cassette (ABC) transporter for nitrate/nitrite, the nasAC operon (SM39_1885 and SM39_1886) encoding a nitrate reductase, and the nasB gene (SM39_1884) encoding a nitrite reductase (all in the SM39_E18 region; supplementary table S1, Supplementary Material online). Thus SM39 appears to have a potential capacity to utilize nitrate, which is present in urine, as a nitrogen source. Related to this, both SM39 and Db11 carry an operon necessary for the metabolism of allantoin but lack the allB gene encoding allantoinase (Cusa et al. 1999). Allantoin is the principal nitrogen source in urine in most mammals. Humans and nonhuman primates, however, carry a genetic lesion, thus rather than allantoin, they accumulate urate that can be catabolized despite the loss of allB. Together, these genes could be associated with the potential capacity of S. marcescens strains to grow in urine and cause urinary tract infections in humans (Mahlen 2011).

Diversity and Differential Evolution of the Potential Virulence of S. marcescens SM39 and Db11

Serratia marcescens is well known for its ability to secrete numerous exoenzymes and other proteins. We identified many known or predicted secreted or surface-exposed proteins potentially related to virulence, most of which (30/40) are conserved in both strains (supplementary table S5, Supplementary Material online).

Surface Structures and Polysaccharides

Here again, the genomic analysis revealed conservation and specialization of the outer surface of the two sequenced strains. Serratia marcescens is characteristically motile, and both strains carry a complete E. coli-like gene set for the biosynthesis of flagella and for chemotaxis (supplementary table S6, Supplementary Material online). These gene sets are highly conserved between SM39 and Db11, the exception being the fliC flagellin-encoding gene, which shows a remarkable sequence diversity (only 68.3% identity in the amino acid sequence). Variation in flagellin structure is linked with antigenic variation and differences in the helicity of the flagellum that in turn are associated with niche adaptation.

Both Db11 and SM39 possess multiple operons for the biosynthesis of chaperone–usher fimbriae associated with biofilm formation and attachment to biotic and abiotic surfaces (Labbate et al. 2007; Shanks et al. 2007). Among these, six are conserved in the two strains (two are significantly divergent in sequence), but three and four are specific to SM39 and Db11, respectively (fig. 3). In addition, although both strains contain a complete set of type IV fimbriae-related genes (Sauvonnet et al. 2000; Kulkarni et al. 2009; Xicohtencatl-Cortes et al. 2009) (supplementary table S7, Supplementary Material online), only SM39 contains a homolog (SM39_0944) of yagZ/ecpA/matB, the gene for E. coli common pili (also known as Mat fimbriae) (Pouttu et al. 2001; Rendon et al. 2007).

Although the SM39 and Db11 LPS core polysaccharide biosynthesis gene clusters (the waa genes) are broadly similar,
there is significant sequence divergence (25–71% amino acid sequence identity for the corresponding predicted proteins) in four orthologous genes (SM39_4551-4554 and SMDB11_4052-4055) in this cluster. These four genes encode two glycosyltransferases, a polymer ligase (WaaL), and an UDP-galactose-4-epimerase. Those of Db11 showed high similarity to those of S. marcescens strain N28b serovar O4 (Coderch et al. 2004). These data suggest that at least two LPS core types exist in S. marcescens.

The gene cluster for O antigen biosynthesis was identified between the S-layer biosynthesis gene cluster and the his operon in both strains (fig. 4). As expected, the gene contents of the loci significantly differ between the two strains; Db11 was serotyped conventionally as O28:K7, whereas SM39 was untypeable. There is also good evidence from the genome sequences that both strains carry a complete gene cluster for group 1 capsule polysaccharide (CPS) biosynthesis, equivalent to the colanic acid biosynthesis genes in E. coli. Differences in gene content in this region suggest, however, that the two strains produce different types of group 1 CPS (fig. 4). This will need to be addressed experimentally in future studies.

Secretion Systems and Secreted Proteins

One unexpected difference between Db11 and SM39 is the presence of a type II secretion system (T2SS) in SM39 but not in Db11 (supplementary fig. S4, Supplementary Material online). The T2SS is a multiprotein secretion complex, present in a wide variety of organisms and frequently implicated in virulence (Korotkov et al. 2012). Although the substrates of the SM39 T2SS are unknown, it could well contribute to the virulence of SM39.

Multiple type V secretion systems (T5SSs) are present in SM39 and Db11. T5SSs include both autotransporters and two-partner systems (Grijpstra et al. 2013; Jacob-Dubuisson et al. 2013). An archetypal example of a two-partner TSSS is ShlBA; the hemolysin ShlA is one of the known major virulence factors in S. marcescens (Goluszko and Nowacki 1989). The shlBA operon is conserved in both strains. In contrast, three additional gene clusters, encoding two-partner systems with partial similarity to ShlA and other haemagglutinin-like proteins, were identified only in SM39 (fig. 5). Two of these, SM39_2080-2077 and SM39_3145-3141, appear to encode contact-dependent inhibition (Cdi) systems. In such systems, CdiB is a translocator protein that assembles the large, hemagglutinin domain-containing CdiA passenger protein on the cell surface (Aoki et al. 2010). Variable toxin domains are found at the C-terminus of CdiA (CdiACt), which mediate contact-dependent growth inhibition of competitor bacteria. Additionally, immunity proteins (CdiI) cognate to the CdiACt are encoded downstream of CdiA. Both loci contain a typical cdiBAI arrangement. The SM39_3144 locus also includes an “orphan” CdiACt- CdiI pair (SM39_3142-3141). Orphan CdiACt-CdiI pairs contain truncated CdiA C-termini thought to represent remnants or reservoirs of alternative toxin domains, which can be exchanged to provide new antibacterial capability (Poole et al. 2011).

Both SM39 and Db11 contain Lip (lipase) and Has (hemophore) type I secretion systems (Kanonenberg et al. 2013), and both possess a single type VI secretion system (T6SS) conserved between the two strains (supplementary fig. S5, Supplementary Material online). T6SS can be used by bacteria either to target eukaryotic cells as a direct virulence
mechanism or to kill rival bacterial cells as a competitive fitness mechanism (Coulthurst 2013). The \textit{S. marcescens} T6SS has been shown to have potent bacterial killing activity (Murdoch et al. 2011). Furthermore, six effectors (antibacterial toxins) secreted by this system have recently been identified in Db11, together with cognate immunity proteins that protect against self-toxicity (English et al. 2012; Fritsch et al. 2013). Of these, the Sp1 and Sp2 effectors, Tae4-family peptidoglycan amidases which attack the target cell wall, and associated Rap (Tai4) immunity proteins are encoded within the T6SS gene cluster (supplementary fig. S5, Supplementary Material online). Overall, the T6SS gene clusters of Db11 and SM39 are very similar with some significant differences: Although the Sp1–Rap1a effector–immunity pair is conserved, the “orphan” Rap1b immunity protein is missing from SM39 and Sp2 has been replaced by a new effector–immunity pair, belonging to the same Tae4–Tai4 family but clearly distinct. This strongly suggests that the T6SS antibacterial activity of \textit{S. marcescens} is highly specific and strain-dependent, as has been observed experimentally (Murdoch et al. 2011).

**Iron Uptake**

Iron uptake is essential for bacterial growth within host organisms because iron is sequestered by the host (Andrews et al. 2003). SM39 and Db11 possess multiple common sets of iron acquisition systems with only a few differences. Although some of the common systems are also found in other sequenced \textit{Serratia} species (supplementary table S8, Supplementary Material online), others, including a second gene cluster for siderophore synthesis, \textit{fecIRA} and \textit{fepBGDC} for transport of ferric compounds, and the \textit{has} operon, are restricted to \textit{S. marcescens}. These species-specific iron uptake systems might be associated with the higher virulence of \textit{S. marcescens} compared with other species of the genus.

**Quorum-Sensing System**

Quorum-sensing systems sense bacterial cell density and play important roles in regulating the behavior of bacterial populations by controlling various biological processes, including pathogenicity. None of the three \textit{N}-acyl homoserine lactone (AHL) quorum-sensing systems so far described in strains of \textit{S. marcescens} (SwrI/SwrR [Eberl, Christiansen et al. 1996; Eberl, Winson et al. 1996], SpnI/SpnR [Horng et al. 2002], and SmaI/SmaR [Coulthurst et al. 2006]) was found in SM39 or Db11. Instead, we identified a candidate SM39-specific AHL system comprising LuxI/LuxR family proteins (SM39_4838 and SM39_4837), which show a high similarity (about 80% amino acid sequence identity) to ExpR/ExpI of \textit{Erwinia tasmaniensis} (Muller et al. 2011). In contrast, no AHL quorum-sensing system was found in Db11. These data indicate a remarkable variation in the AHL quorum-sensing system among \textit{S. marcescens} strains.
Variation of Nonribosomal Peptide Synthetase-Dependent Secondary Metabolites and Weapons for Interbacterial Competition

Diffusible inhibitory molecules include antibiotics and bacteriocins (protein or peptide toxins targeting related organisms). SM39 and Db11 have one and two bacteriocins, respectively, not shared by the other, as well as two in common (including bacteriocin 28b).

An operon of six genes specific to Db11 (SMDB11_2293-2288, named alb1–alb6) encodes a hybrid nonribosomal peptide synthetase (NRPS)-polyketide synthase (PKS) enzyme, plus tailoring enzymes and a resistance-conferring efflux pump. This cluster was recently identified as required for the ability of Db10 to produce a diffusible antimicrobial compound and shown to direct the synthesis of a broad-spectrum antibiotic called althiomycin (Gerc et al. 2012). Two predicted NRPS enzymes are found only in SM39 and may also be involved in biosynthesis of antimicrobial molecules. These variations, along with those in T6SS and Cdi systems, suggest that interbacterial competition in microbial communities is one of the important drivers of genomic diversity in S. marcescens.

NRPS-dependent molecules also have other roles. In Db11, SMDB11_3680, or SwrA, produces the biosurfactant Serratwettin W2 (Pradel et al. 2007). In SM39, SM39_3884 shares only 56% identity with SwrA but, given that it is encoded in the same genomic locus and has a very similar domain organization, it probably synthesizes a related biosurfactant molecule. Three other NRPSs are shared by both strains, including two involved in synthesis of the two siderophores mentioned earlier.

In Vivo Screening Shows That Many Core S. marcescens or Serratia sp. Genes Are Involved in Virulence in an Invertebrate Model Host

Db10 was isolated from D. melanogaster (Flyg et al. 1980), but it is pathogenic in several other infection models, including C. elegans and mice. Previous work has shown that there is a substantial overlap between the genes required for full virulence of S. marcescens during infection of flies and C. elegans (Kurz et al. 2003). Thus, as the nematode is well-suited to automated screens (Kurz and Ewbank 2007; Gavis et al. 2009), we used it to look for new S. marcescens virulence factors and to determine whether genes associated with virulence were within the shared or accessory gene sets. We constructed a mini-Tn5-Sm transposon mutant library in Db10 and assayed clones in a high-throughput screen to identify mutants with a reduced ability to kill C. elegans. From the 12,480 individual Db10-derived clones that were tested, 12 mutants (0.1% of all the mutants screened) were selected as showing the most robust reduction in virulence.

The mini-Tn5-Sm insertion site in each mutant was identified by sequencing and mapping to the genomic sequence of S. marcescens Db11 (for a complete list, see supplementary table S9, Supplementary Material online). Of the 12 mutants, four loci had also been identified in a previous solid medium-based screen (Kurz et al. 2003). These include those involved in O antigen biosynthesis and siderophore production. The availability of the complete sequence combined with recent findings (Murdoch et al. 2011; English et al. 2012; Fritsch et al. 2013) revealed a hitherto unsuspected contribution of a T6SS to the pathogenic capability of S. marcescens, either directly or by providing a competitive fitness advantage against other bacteria in vivo. The previously isolated mutants 22D9 and 7D1 (Kurz et al. 2003) were found to correspond to SMDB11_2265 and SMDB11_2266, respectively, which are both T6SS immunity proteins, 8C7 to SMDB11_1112, a T6SS effector protein of unknown function, whereas the previously uncloned 23C11 corresponded to SMDB11_3455, a T6SS-related, minor Hcp homolog (English et al. 2012; Fritsch et al. 2013). As previously reported, this screen also led to the isolation of a mutant in the swrA gene responsible for the biosynthesis of the surfactant Serratwettin W2 (Pradel et al. 2007). Additional mutants of interest include one containing an insert in a wza homolog, potentially involved in capsule biosynthesis, and three deficient in respiration (JESM266, JESM268, JESM271; supplementary table S9, Supplementary Material online). None of these latter mutants showed a defect in growth under standard conditions. This raises the possibility that the nematode intestine represents a restrictive respiratory environment.

The combined results of this and the previously published screens (Kurz et al. 2003) identified 30 loci important for the in vivo virulence of Db10 and/or Db11. Among the 30 genes, 22 are conserved in SM39. Of these, 11 are conserved in all of the four other Serratia strains analyzed (supplementary table S9, Supplementary Material online). The proportions of virulence genes conserved between Db11 and SM39 (73%) and between this common set and the other Serratia strains (50%) are markedly lower than the equivalent proportions of the entire gene sets (84.5% and 71%, respectively). This presumably reflects the different tropisms of the different species and the host-specific nature of many virulence factors.

Origins of Antimicrobial Resistance in SM39 and Db11

Intrinsic Resistance

A large number of genes related to antimicrobial resistance are encoded on the chromosomes of both Db11 and the clinical S. marcescens isolate SM39. As listed in table 2, we identified many efflux pumps belonging to five families: The ABC superfamily, the major facilitator superfamily (MFS), the multidrug and toxic-compound extrusion (MATE) family, the small multidrug resistance (SMR) family, and the resistance nodulation division (RND) family. Among these, only smdAB (Matsuo et al. 2008), sdeAB (Kumar and Worobec 2005), sdeXY (Chen et al. 2003), smdY (Shahcheraghi et al. 2007), and smrE (Minato et al. 2008) had been characterized in previous
studies. Most of the efflux pumps are shared by the two strains although a few are strain-specific. In addition, a class C beta-lactamase AmpC, two aminoglycoside acetyltransferases each having distinct substrate specificities, and a fosfomycin-inactivating enzyme (FosA) are encoded on the chromosomes of the two strains (table 3). Thus, the overall repertoire of the chromosomally encoded resistance genes is largely similar in both strains despite their different origins, and is highly conserved accounting for the high and broad intrinsic resistance of \textit{S. marcescens} (table 3). One notable difference between the two strains is a mutation in the \textit{gyrA} gene of the clinically isolated SM39, generating a S83R substitution and conferring resistance to quinolone (Weigel et al. 1998).

### Table 2

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0414-0415</td>
<td>0354-0355</td>
<td>smdAB</td>
<td>smdA, smdB</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>1134-1135</td>
<td>0964-0965</td>
<td>macAB</td>
<td>macB</td>
<td>macA</td>
<td>—</td>
</tr>
<tr>
<td>1329-1331</td>
<td>1118-1120</td>
<td>etsABC</td>
<td>etsB</td>
<td>etsA, etsC</td>
<td></td>
</tr>
<tr>
<td>4783-4785</td>
<td>4552-4554</td>
<td>—</td>
<td>4784/4553</td>
<td>4783/4552</td>
<td>—</td>
</tr>
<tr>
<td>RND type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0448-0449</td>
<td>0369-0370</td>
<td>sdeXY</td>
<td>sdeY</td>
<td>sdeX</td>
<td>—</td>
</tr>
<tr>
<td>1281-1282</td>
<td>1059-1060</td>
<td>—</td>
<td>1281/1059</td>
<td>1282/1060</td>
<td>—</td>
</tr>
<tr>
<td>1400-1401</td>
<td>1196-1197</td>
<td>sdeAB</td>
<td>sdeB</td>
<td>sdeA</td>
<td></td>
</tr>
<tr>
<td>Not found</td>
<td>1254A-1255A</td>
<td>—</td>
<td>SMDB11_1254A</td>
<td>SMDB11_1255A</td>
<td>SMDB11_1256</td>
</tr>
<tr>
<td>1913-1915</td>
<td>Not found</td>
<td>SM39_1914</td>
<td>SM39_1913</td>
<td>SM39_1915*</td>
<td></td>
</tr>
<tr>
<td>1920-1922</td>
<td>1741-1743</td>
<td>—</td>
<td>1921/1742</td>
<td>1922/1743</td>
<td>SMDB11_1741*</td>
</tr>
<tr>
<td>Not found</td>
<td>1698-1699</td>
<td>—</td>
<td>SMDB11_1699</td>
<td>SMDB11_1698</td>
<td>—</td>
</tr>
<tr>
<td>1958</td>
<td>Not found</td>
<td>SM39_1958</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>3100</td>
<td>2891</td>
<td>acrD</td>
<td>acrD</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>3162-3164</td>
<td>2945-2947</td>
<td>sdeCDE</td>
<td>sdeD, sdeE</td>
<td>sdeC</td>
<td>—</td>
</tr>
<tr>
<td>MFS type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>467</td>
<td>390</td>
<td>fsr</td>
<td>fsr</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>1351</td>
<td>1140</td>
<td>mdtG</td>
<td>mdtG</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>1810</td>
<td>1629</td>
<td>—</td>
<td>1810/1629</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Not found</td>
<td>1759</td>
<td>tetA</td>
<td>tetA</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>2164</td>
<td>1961</td>
<td>smfY</td>
<td>smfY</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>2273</td>
<td>2069</td>
<td>mdtH</td>
<td>mdtH</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>2819</td>
<td>2580</td>
<td>bcr</td>
<td>bcr</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>3165</td>
<td>2948</td>
<td>mdtD</td>
<td>mdtD</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>3343-3344</td>
<td>3133-3134</td>
<td>emrAB</td>
<td>emrB</td>
<td>SMDB11_3132*</td>
<td></td>
</tr>
<tr>
<td>4622</td>
<td>4024</td>
<td>emrD</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>4391</td>
<td>4201</td>
<td>mdfA</td>
<td>mdfA</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>4497</td>
<td>4107</td>
<td>—</td>
<td>4497/4107</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>MATE type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1653</td>
<td>1463</td>
<td>mdtK</td>
<td>mdtK</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>2598</td>
<td>2377</td>
<td>—</td>
<td>2598/2377</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>3912</td>
<td>3713</td>
<td>dinF</td>
<td>dinF</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>SMR type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2035</td>
<td>1855</td>
<td>ssmE</td>
<td>2035/1855</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>2237-2238</td>
<td>2032-2033</td>
<td>mdtII</td>
<td>2237/2032</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>4693</td>
<td>4462</td>
<td>sugE</td>
<td>4693/4462</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>pSMC1_35</td>
<td>Not found</td>
<td>qacE1</td>
<td>qacE1</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>pSMC1_46</td>
<td>Not found</td>
<td>qacE2</td>
<td>qacE2</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

*Similar to the NodT family protein but contains no membrane-spanning domain.

Acquired Antimicrobial Resistance Genes

Unlike Db11, SM39 carries two plasmids. The overall GC contents of the two plasmids differ significantly from the
The plasmid pSMC2 carries a set of genes for conjugation (the tra and trb genes), and although lacking traJ and traT, it can be transferred by conjugation from SM39 to E. coli strain x1037 Rifr (data not shown). Although pSMC2 contains no gene related to virulence or drug resistance, pSMC1 carries an integron that contains the aadA gene encoding an aminoglycoside-3'-adenylyltransferase, the...
blaCMY-8 gene encoding a class C beta-lactamase which inactivates a much broader range of beta-lactams than the chromosomally encoded class C beta-lactamase (AmpC), and the blaIMP-1 gene encoding a metal beta-lactamase (MBL). In addition, pSMC1 contains two copies of the qacE gene, associated with resistance to a range of disinfectants (Kucken et al. 2000), and two copies of the sulfonamide resistance gene (sulI) as well as a transposon encoding the mer operon for mercury resistance (fig. 6A).

Comparison of a pSMC1-cured SM39 derivative (KS3) with the wild type showed that the high level of resistance displayed by SM39 to beta-lactams, aminoglycosides and mercury is accounted for by the presence of pSMC1. Compared with SM39, KS3 was great or equal to four more sensitive to...
22 of the 36 antimicrobial agents tested; 14 of the 16 beta-lactams, all of the seven aminoglycosides, and to mercury. The 14 beta-lactams included carbapenems, the resistance to which is attributable to MBL. The exception is resistance to fluoroquinolones, which, as mentioned above, is attributable to the chromosomal mutation in gyrA (table 3).

Plasmid pSMC1 exhibits a remarkable similarity to an IncP plasmid of Pseudomonas aeruginosa, pO7-406 (Li et al. 2008) (fig. 6A). The backbones of the two plasmids encoding the tra region and replication functions are nearly identical in gene compliment and with a GC content (64.1–66.1%) characteristic of pseudomonads. The main differences concern their genetic cargo encoding multiple drug resistance. This is consistent with the notion that Pseudomonas species are the original hosts for IncP plasmids with high GC backbones (Thorsted et al. 1998). However, as the integron-containing region of pSMC1 is also high similarity to that on plasmid pK29 of Klebsiella pneumoniae (Chen et al. 2007) although complex rearrangements including inversion and gene duplication have occurred, it suggests that these closely related plasmids have become widely disseminated in the Enterobacteriaceae (fig. 6B).

Concluding Remarks

We selected two S. marcescens strains from contrasting sources in an attempt to maximize our ability to capture the genomic diversity of this ubiquitous enteric species. Through our comparative analysis, we have defined a core genome of S. marcescens as well as that of the genus, although it should be noted that the latter gene set was generated from free-living Serratia species only. The former gene set defines the intrinsic metabolic capacities, virulence, and multidrug resistance of this opportunistic pathogen. Analysis of strain-specific genes or genomic regions has further revealed a high level of genomic diversity and plasticity of S. marcescens, which reflects the diversity of niches that this species can occupy. In fact, among the SM39-specific genes, many genes that are implicated in the high virulence potential of this clinical isolate have been identified, this includes many virulence-related genes found within the core genome of S. marcescens. Furthermore, analysis of the pSMC1 plasmid of SM39 has revealed that it encodes MBL and many other drug resistance determinants and is responsible for the extremely high level of multidrug resistance of the strain. We were able to propose a possible origin and evolution of pSMC1 on the basis of its genomic features. These data, combined with available tools for functional genomic analysis such as that described by Petty et al. (2006: 1701–1708), will accelerate research on S. marcescens in numerous domains and provide new insights into the genetic mechanisms responsible for the emergence of pathogens highly resistant to multiple antimicrobial agents.

Supplementary Material

Supplementary figures S1–S5 and tables S1–S9 are available at Genome Biology and Evolution online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

The authors thank many members of the Serratia research community, in particular C. Wandersman, for their support, Aurelie Blanc for worm sorting, and T. Nakamura and H. Takahashi for providing strain SM39. This work was supported partly by a Scientific Research Grant on Priority Areas from the University of Miyazaki and the Program to Disseminate Tenure Tracking System from the Japanese Ministry of Education, Culture, Sports, Science and Technology to A.I. (http://www.miyazaki-u.ac.jp/en/english/index.html), and Grant-in-Aid for Scientific Research (18390137 and 23590523) from the Japan Society for the Promotion of Science to N.G.

Sequencing of Db11 was made possible by the 2001 CNRS Séquençage à Grande Echelle programme, and that of SM39 was by the Research for the Future Program of the Japan Society for the Promotion of Science (JSPS-RFTF 00L01411).

Work in the Evbank lab was supported by institutional funding from INSERM and the CNRS, and program grants from the French Ministry of Research. This work was also supported by Wellcome Trust grant (098051).

Literature Cited


Cusa E, Obradors N, Baldoma L, Badia J, Aguilar J. 1999. Genetic analysis of a chromosomal region containing genes required for assimilation of


Associate editor: John McCutcheon