

1 **Extreme genome selection towards complete antimicrobial**
2 **resistance in a nosocomial strain of *Stenotrophomonas***
3 ***maltophilia* complex**

4 Sanjeet Kumar^{1,§,^}, Kanika Bansal^{1,^}, Prashant P. Patil^{1,#}, Amandeep Kaur¹, Satinder Kaur², Vikas
5 Gautam^{2,*} and Prabhu B. Patil^{1,*}

6

7 ¹Bacterial Genomics and Evolution Laboratory, CSIR-Institute of Microbial Technology,
8 Chandigarh, India.

9 ²Postgraduate Institute of Medical Education and Research, Chandigarh, India.

10 [§]Present address: Department of Archaeogenetics, Max Planck Institute for the Science of Human
11 History, Jena, Germany.

12 [#]Present address: Department of Microbiology, School of Medicine, University of Washington,
13 Seattle, WA, USA.

14 [^]Equal Contribution

15 ^{*} Corresponding author

16

17 **Correspondence:**

18 **1) Prabhu B. Patil**

19 Email: pbpatil@imtech.res.in

20 Senior Scientist

21 CSIR-Institute of Microbial Technology

22 Sector 39A, Chandigarh, India-160 036

23 **2) Vikas Gautam**

24 Email: r_vg@yahoo.co.uk

25 Professor

26 Department of Medical Microbiology

27 Post Graduate Institute of Medical Education & Research

28 Chandigarh, India - 160012

29

30 **Keyword**

31 Extreme drug resistance, *Stenotrophomonas maltophilia*, integrons, complete genome, *sul* gene
32 cluster.

33 **Abbreviations**

34 XDR, extreme drug resistance; ISCR, IS91-like elements common regions; ST, sequence type;
35 TMP/SXT, trimethoprim/sulfamethoxazole; DR, dynamic regions.

36 **Repositories:**

37 Complete genome sequence of *Stenotrophomonas maltophilia* SM866: CP031058

38 **ABSTRACT**

39 We report first complete genome sequence and analysis of an extreme drug resistance (XDR)
40 nosocomial *Stenotrophomonas maltophilia* that is resistant to the mainstream drugs i.e.
41 trimethoprim/sulfamethoxazole (TMP/SXT) and levofloxacin. Taxonogenomic analysis revealed it to
42 be a novel genomospecies of the *Stenotrophomonas maltophilia* complex (Smc). Comprehensive
43 genomic investigation revealed fourteen dynamic regions (DRs) exclusive to SM866, consisting of
44 diverse antibiotic resistance genes, efflux pumps, heavy metal resistance, various transcriptional
45 regulators etc. Further, resistome analysis of Smc clearly depicted SM866 to be an enriched strain,
46 having diversified resistome consisting of *sul1* and *sul2* genes. Interestingly, SM866 does not have
47 any plasmid but it harbors two diverse super-integrations of chromosomal origin. Apart from genes for
48 sulfonamide resistance (*sul1* and *sul2*), both of these integrations harbor an array of antibiotic resistance
49 genes linked to ISCR (IS91-like elements common regions) elements. These integrations also harbor
50 genes encoding resistance to commonly used disinfectants like quaternary ammonium compounds and
51 heavy metals like mercury. Hence, isolation of a novel strain belonging to a novel sequence type (ST)
52 and genomospecies with diverse array of resistance from a tertiary care unit of India indicates extent
53 and nature of selection pressure driving XDRs in hospital settings. There is an urgent need to employ
54 complete genome based investigation using emerging technologies for tracking emergence of XDR at
55 the global level and designing strategies of sanitization and antibiotic regime.

56 **Impact Statement**

57 The hospital settings in India have one of the highest usage of antimicrobials and heavy patient
58 load. Our finding of a novel clinical isolate of *S. maltophilia* complex with two super-integrations
59 harbouring array of antibiotic resistance genes along with antimicrobials resistance genes indicates
60 the extent and the nature of selection pressures in action. Further, the presence of ISCR type of
61 transposable elements on both integrations not only indicates its propensity to transfer resistome but
62 also their chromosomal origin suggests possibilities for further genomic/phenotypic complexities.
63 Such complex cassettes and strain are potential threat to global health care. Hence, there is an
64 urgent need to employ cost-effective long read technologies to keep vigilance on novel and extreme
65 antimicrobial resistance pathogens in populous countries. There is also need for surveillance for
66 usage of antimicrobials for hygiene and linked/rapid co-evolution of extreme drug resistance in
67 nosocomial pathogens. Our finding of the chromosomal encoding XDR will shed a light on the

68 need of hour to understand the evolution of an opportunistic nosocomial pathogen belonging to *S.*
69 *maltophilia*.

70 INTRODUCTION

71 According to World health organization (WHO) list *Stenotrophomonas maltophilia* is one of the
72 leading opportunistic multidrug-resistance pathogen in hospitals worldwide
73 (http://www.who.int/drugresistance/AMR_Importance/en/). *S. maltophilia* causes variety of
74 infections including respiratory tract infections, bloodstream infections, bone and joint infections,
75 urinary tract infections, endocarditis and meningitis (1). *S. maltophilia* is having intrinsic resistance
76 to multiple groups of widely used antibiotics such as: cephalosporins, carbapenems,
77 aminoglycosides and macrolides (2). Till now, trimethoprim/sulfamethoxazole (TMP/SXT) and
78 levofloxacin were considered as the mainstream therapy for such infections (3, 4). Elevated reports
79 of resistance to both of these drugs is a growing concern worldwide (1). Hence, their rapid
80 identification and investigation are key towards the successful treatment.

81 *S. maltophilia* has been reported to acquire resistance through integrons, transposons and plasmids
82 conferring multiple resistance genes (5). There are various PCR-based or draft genome based
83 studies for the identification of these resistant strains (6-8). Until now, ISCRs (IS91-like elements
84 common regions), considered as gene-capturing systems of 21st century are also found in the range
85 of extreme drug resistant (XDR) bacteria carrying antibiotic and antimicrobial resistance genes (9,
86 10). For instance, in *Shewanella xiamenensis* the ISCR was linked with antibiotic resistance genes
87 along with resistance genes for biocides heavily used in hospital settings such as: quaternary
88 ammonium compounds (*qac*), heavy metals (mercury resistance genes) etc. (11). Interestingly,
89 ISCR in *S. maltophilia* strains are also associated with *sul* genes. These are majorly reported from
90 plasmids or with very low confidence, predicted to be chromosomally encoded (5, 12). Lack of
91 complete genome based studies of XDR *S. maltophilia*, limits our understanding of the genomic
92 context of the resistance genes cassettes. Such an information will be critical to understand genome
93 dynamics particularly for new strains carrying multiple alleles that are spreading by integrons
94 which carry array of antibiotic resistance genes linked to ISCR elements (13).

95 Developing countries like India have high selection pressure because of heavy antibiotic and
96 biocide usage and high patient density coupled with tropical climate setting. Recent reports
97 revealed rapid emergence of extreme drug resistance in India in sulfonamide resistant *S.*
98 *maltophilia* and carbapenem resistance conferred by New Delhi metallo-beta-lactamase 1 (NDM-1)
99 in *Enterobacteriaceae* strains (1, 14). Furthermore, studies have reported emergence of biocides
100 (absolute ethanol, povidone iodine, sodium hypochlorite, and QACs) resistance among multidrug
101 resistance bacteria such as: *Escherichia coli*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* and
102 *Staphylococcus aureus* and diverse biofilm forming bacteria of dairy niche (15, 16). Use of

103 emerging cost-effective nanopore technologies are revolutionizing investigation of such extreme
104 antimicrobial resistance by allowing rapid access of complete genome information. In the present
105 study, we report isolation and complete genome based investigation of a novel sequence type (ST-
106 366) of extreme drug resistant strain belonging to Smc from a tertiary care center in northern India.
107 Complete genome level study of the strain indicates its highly dynamic nature and found to harbor
108 fourteen dynamic regions (DRs). The strain also harbours two diverse chromosomal integrons
109 linked to an array of antibiotic resistance genes along with IS elements and antimicrobial resistance
110 genes. Overall, the study indicated serious long term consequences of high selection pressure for
111 rapid emergence and evolution towards complete antimicrobial resistance strains in Indian
112 hospitals. At the same time, the study highlights the need to employ rapid and long read
113 technologies to carry out complete genome based investigations to understand and manage extreme
114 drug resistance at global level.

115 **METHODS**

116 **Bacterial isolation, identification and antimicrobial susceptibility testing**

117 Bacterium was isolated from the respiratory specimen of hospitalized patient from
118 Postgraduate Institute of Medical Education and Research (PGIMER), Chandigarh in 2013.
119 Ethics approval and each patient's written consent was not required as it was a part of routine
120 clinical testing. Bacterial cultures were grown in the nutrient broth at constant shaking of 200
121 rpm at 37 °C. DNA isolation was performed with ZR fungal/ bacterial DNA MiniPrep Kit
122 (Zymo research) in accordance to manufacturer instructions. In order to identify the bacterial
123 species partial 16S rRNA sequencing was performed using genomic DNA. Primers used for
124 16S rRNA amplification were 27F and 1492R. For Bacterial identification, EZ Biocloud (17)
125 (<http://www.ezbiocloud.net/identify>) was used. Further, antimicrobial susceptibility testing
126 was performed by the Kirby-Bauer disc diffusion method as per the Clinical Laboratory
127 Standards Institute (CLSI), 2017 guidelines against trimethoprim-sulfamethoxazole (25 mcg),
128 chloramphenicol (30 mcg), levofloxacin (5 mcg) and ceftazidime (30 mcg) antibiotics.

129 **Genome sequencing, assembly and annotation**

130 **a) Nanopore sequencing:** Bacterial cultures were grown overnight in 15 ml of Luria-
131 Bertani broth. Bacterial cells were harvested and high quality bacterial genomic DNA
132 was isolated using Qiagen DNeasy Blood & Tissue Kit. DNA was purified using
133 0.45x Agencourt AMPure XP beads (Beckman Coulter, High Wycombe, UK) and
134 quantified using Nanodrop 1000 v3.8 (Thermo Fisher Scientific) and Qubit 2.0

135 (Thermo Fisher Scientific). Purified DNA sample was end-repaired (NEBnext ultra II
136 end repair kit, New England Biolabs, MA, USA), cleaned up with 1x AMPure XP
137 beads (Beckmann Coulter, USA). Native barcode ligation was performed with NEB
138 blunt/ TA ligase (New England Biolabs, MA, USA) and cleaned with 0.5x AMPure
139 XP beads. Qubit quantified and barcode ligated DNA samples were pooled at equi-
140 molar concentration to attain 1 µg pooled sample. Adapter ligation (BAM adapters)
141 was performed for 15 minutes using NEBNext Quick Ligation Module (New England
142 Biolabs, MA, USA). Library mix was cleaned up using 0.4X AMPure XP beads
143 (Beckmann Coulter, USA) and finally sequencing library was eluted in 15 µl of
144 elution buffer.

145 **b) Illumina Sequencing:** Genomic DNA was extracted from SM866 using ZR
146 Fungal/Bacterial DNA MiniPrep Kit (Zymo Research, Irvine, CA, USA) and
147 quantified using Qubit 2.0 Fluorometer (Life technologies). Short reads libraries was
148 sequenced on Illumina Miseq (Illumina, Inc., San Diego, CA, USA). Illumina
149 sequencing library was prepared using Nextera XT sample preparation kit (Illumina,
150 Inc., San Diego, CA, USA) with dual indexing adapters and sequencing using 250*2
151 paired end sequencing kit. The adapters were trimmed while sequencing by internal
152 software Illumina MiSeq Control v.2.5.0.5 (Illumina, Inc., San Diego, CA, USA).

153 Genome assembly was first performed using the Unicycler v0.4.6 (18) with ONT long reads
154 with bold mode. The complete genome obtained using ONT reads was then polished for
155 multiple rounds of Pilon v1.22 (19) using Illumina raw reads of the sample SM866. The
156 genome was assembled in a closed circular single chromosome. Annotation was performed
157 using NCBI-PGAP (20) annotation pipeline. Assembly quality in terms of completeness and
158 contamination was accessed using the CheckM v1.0.11 (21). Genome coverage of the SM866
159 was calculated using BMAP v36.92 (22).

160 **16S rRNA gene Phylogeny:** Complete 16S rRNA gene sequence from the type strains of the
161 genus *Stenotrophomonas* was fetched from their corresponding INSDC number
162 (<http://www.insdc.org/>) from LPSN (23) given for their species definition
163 (<http://www.bacterio.net/stenotrophomonas.html>). 16s rRNA sequence from SM866 was
164 fetched from complete assembled genome using standalone version of RNAMmer v1.2 (24).
165 All the complete 16S rRNA gene sequences were aligned using ClustalW (25). Phylogenetic
166 tree was constructed based on neighbor-joining method was created with 1000 bootstrap.

167 *Xanthomonas campestris* ATCC33913 (26) was used as outgroup to genus
168 *Stenotrophomonas*.

169 **Multi-locus sequencing typing (MLST):** Sequence type analysis of the strain SM866 was
170 done using center for genomic epidemiology MLST 2.0
171 (<https://cge.cbs.dtu.dk/services/MLST/>) and PubMLST (<https://pubmlst.org/bigsub>).

172 **Phylogenomic and taxonogenomics based identification and characterization**

173 More robust phylogeny was constructed using > 400 conserved gene using PhyloPhlAn (27).
174 Complete genome proteome of all the type species of genus *Stenotrophomonas* was used for
175 the construction of the phylogenetic tree. *X. campestris* ATCC 33913 was used as outgroup.
176 Further, taxonogenomic analysis of the strain was performed using orthoANI (28) and dDDH
177 values (29).

178 **Dynamic regions and resistome analysis**

179 The unique DRs of SM866 were detected using BRIG v0.95 (30) which were further
180 confirmed using blastn package of BLAST++ (31). A well characterized nucleotide
181 sequence of the antibiotic resistance gene clusters and efflux pumps were retrieved from
182 the complete genome of *S. maltophilia* K279a (32). A standalone blast was performed with
183 these gene as query with SM866 and the previously reported strains of Smc (33). A heat
184 map showing the presence and absence of the gene cluster was generated using GENE-E
185 v3.0.215 (<http://www.broadinstitute.org/cancer/software/GENE-E>). Further, schematic
186 diagram of the cassettes I and II were generated using Easyfig v2.2.2 (34).

187 **Plasmid profiling**

188 The plasmid profiling was performed for isolate SM866 using FosmidMAX DNA
189 Purification kit (Epicenter, Illumina). For this the bacterial sample was grown in 10 ml of LB
190 and incubated for 28 °C for 16 hours. Bacterial cells were harvested by centrifugation at 8000
191 rpm for 5 minutes. Plasmid isolation was performed as per the manufacturer
192 recommendations. 5µl of sample was loaded on 0.7% agarose gel to check the presence of
193 plasmid (Supplementary figure 2). To reconfirm the absence of plasmid, Plasmid-Safe™
194 ATP-Dependent DNase was performed in accordance to manufacturer recommendation and
195 *S. maltophilia* ATCC13637 (T) was used as a control.

196 **RESULTS**

197 **Bacterial isolation, identification and antimicrobial susceptibility**

198 The strain SM866 was isolated in 2013 from respiratory specimen from an intensive care unit
199 patient admitted in a tertiary care hospital in Northern India (20). SM866 exhibited 99.5% partial
200 16S rRNA nucleotide sequence identity with the type strain *S. maltophilia* ATCC 13637(T). The
201 16S rRNA gene phylogeny using all the type strains of the genus *Stenotrophomonas* clearly showed
202 that our isolate was in Smc (supplementary figure 1(A)). For the phylogeny, we have considered
203 all the known novel genomospecies of the Smc (SM5815, SM3123, SM10507 and
204 SM16975) (33). Further, antibiotic resistance profiling revealed that SM866 is resistant to all the
205 antibiotics tested which includes trimethoprim/ sulfamethoxazole, levofloxacin, chloramphenicol
206 and ceftazidime. The zone of inhibition for the antibiotics used is in table 1.

207 **Multi-locus sequence typing and population structure analysis**

208 Multilocus sequence typing of SM866 revealed that out of seven housekeeping genes (*atpD*, *gapA*,
209 *guaA*, *mutM*, *nuoD*, *ppsA* and *recA*), it harbors three novel housekeeping alleles i.e. *guaA*, *mutM* and
210 *nuoD* with newly assigned accession numbers: 272, 147 and 141 respectively. Hence, SM866 was
211 assigned a novel sequence type, ST-366. Further, population structure analysis was carried out
212 considering all the STs of *S. maltophilia*, clearly depicting that SM866 is a singleton ST and it does
213 not belong to clonal complex of any of the already known STs (supplementary figure 1(B)).

214 **Genome sequencing and annotation**

215 The complete circular genome of SM866 with one chromosome (and no plasmid) of 5.08 MB with
216 60X coverage and 66.03% GC content was obtained by assembly of the Oxford Nanopore long
217 reads and polishing with Illumina short reads. SM866 genome has been submitted to the public
218 repository NCBI with accession number CP031058. PGAP annotation resulted in 4,715 CDS, 73
219 tRNA and 4 ncRNAs. Further, CheckM estimated 99.76% completeness and 0.72% contamination
220 in the assembled genome (<https://figshare.com/s/72aee1707e6a6182ac6d>). Further, plasmid profile
221 of the strain SM866 was confirmed by plasmid screening resulted in no plasmid detection in
222 SM866 and the *S. maltophilia* ATCC13637 (T) (supplementary figure 2(A)). Although, there was
223 presence of very faint band in SM866, which when treated with ATP-dependent DNase method was
224 degraded, inferring that to be a genomic DNA contamination (supplementary figure 2(B)). Hence,
225 this confirms that SM866 is devoid of plasmid, as also indicated by complete genome.

226 **Phylogenomic and taxonogenomic analysis**

227 Taxonogenomic analysis of SM866 was carried out to clearly know its species status. ANI and dDDH
228 values for SM866 were below the cut-off values for species delineation. Within Smc, values ranged
229 from 87%-93% and 34%-48% for ANI and dDDH respectively (figure1). Hence, this analysis
230 suggested SM866 to be a novel genomospecies in Smc. Furthermore, phylogenomic analysis done by
231 taking *Xanthomonas campestris* ATCC33913 (T) as an outgroup suggested that SM866 falls within
232 the Smc complex (figure 1).

233 **Understanding genome dynamics in evolution of extreme drug resistance**

234 SM866 was found to be trimethoprim/sulfamethoxazole and levofloxacin resistant, unlikely of
235 previous genomospecies of the Smc. We focused on the unique regions to SM866, which were absent
236 from other Smc complex strains. Interestingly, we found 14 dynamic regions (DRs) exclusively in
237 SM866 (supplementary table 1). Functional classification of the DRs along with their genomic
238 locations is provided in the figure 2 and supplementary table 1. Here, a range of antibiotic resistance
239 genes (tetracycline, chloramphenicol, beta-lactamases, sulfonamide-resistance, aminoglycosides),
240 efflux pumps (RND, ABC, cation transporter, Na⁺/H⁺ antiporter), heavy metal resistance (copper,
241 mercury, arsenate), quaternary ammonium compound (QacE), glycosyltransferases,
242 methyltransferases, acetyltransferases, DEAD/DEAH box helicase, disulfide bond formation protein
243 DsbA, DNA repair protein, transcriptional regulators (MerR, ArsR, XRE, CopG, Lrp/AsnC, LysR,
244 TolC, TetR/AcrR, LuxR), toxin-antitoxin system HipA, lipoproteins, lysozyme, redox-sensitive
245 transcriptional activator SoxR, integrases, sensor kinases, response regulators, IS elements, phage
246 proteins, transposases, conjugal transfer proteins (TrbBCDEFGIJK, TraGJ) and proteases etc. were
247 found in the DRs unique to SM866.

248 **Resistome analysis of Smc genomospecies**

249 Resistome of Smc genomospecies is represented in figure 3. Here, among all the Smc
250 genomospecies, only SM866 was found to have *sul1* and *sul2* genes. Here, drug resistance genes
251 for aminoglycosides, beta-lactams, chloramphenicol, fluoroquinolones; efflux pumps like RND,
252 SMR, MFS, MATE and ABC transporters were examined. This reveals the underlying genes for
253 intrinsic antibiotic resistance of the Smc strains. Interestingly, in the vast resistome of SM866, *sul1*
254 and *sul2* are exclusively present in SM866 only, which have been also detected in R11 and R9
255 respectively.

256 **Diverse super-integrations conferring extreme drug resistance**

257 Genome mining of the SM866 revealed presence of two *sul* alleles (*sul1* and *sul2*) each on two
258 distantly located integrations of chromosomal origin (figure 4(A)). To our knowledge, this is the first
259 complete genome of *Stenotrophomonas* carrying *sul* genes. In addition to the *sul* genes, both the

260 integrons were harbouring array of other drug resistance (tetracycline, phenicol, fluoroquinolone
261 and macrolide antibiotics) and biocide resistance (*mer* operon, *qac* gene) genes linked to gene
262 capturing machinery i.e. ISCR (IS91-like elements) element (figure 4(B)). Cassette I (3448861-
263 3472206) (DUW70_16515 to DUW70_16640) was carrying *sul1*, truncated *qacE*, *bla(PME)*,
264 *aac(3)*, *aph(3)*, *floR2* and *tet(G)* genes in the ISCR. Cassette II (3052377-3082576)
265 (DUW70_14405 to DUW70_14605) was carrying *sul2*, *floR*, *aph(3)*, *aph(6)* genes in the ISCR
266 along with *mer* operon. Additionally, SM866 is also having *smeDEF* (4366825- 4372654)
267 (DUW70_20855 to DUW70_20865) in the chromosomal region.

268 **DISCUSSION**

269 *S. maltophilia* complex strains demonstrate evolution of an opportunistic minor pathogen to
270 novel genomospecies and can be a major threat in health care settings. Since *S. maltophilia*
271 genomes displayed extreme fluidity at inter-strain level (35), there was need to understand
272 extreme drug resistance of emerging strains at complete genome level. Till now, *S.*
273 *maltophilia* has been studied at gene level (PCR-based) or draft genome level to understand
274 extreme drug resistance (6-8). But, such studies are not giving the genomic context such as
275 flanking region and origin of the resistance genes. In the present study, we have analyzed the
276 complete genome of extreme drug resistant novel genomospecies of Smc. Our finding of two
277 *sul* genes on two separate integrons in the chromosome indicates the systemic and rapid
278 evolution of XDRs. This clearly depicts extreme selection pressure towards acquisition,
279 maintenance and spread of multiple *sul* genes. At the same time, integrons carrying *sul* genes
280 along with other commonly used antibiotics and biocide resistance in a highly evolved ISCR
281 is alarming. Furthermore, till date, there is no complete genome available for a
282 sulfamethoxazole resistant strain of Smc. This is also first complete genome sequence of an
283 XDR strain from Smc complex (35). Hence, genome dynamics, considering the repetitive
284 elements and mobilome of a strain resistant to main-stream drugs will be highly significant in
285 understanding evolution of extreme drug resistance in the tertiary care units of Indian
286 hospitals.

287 Genome dynamics is the key for survival and evolution of the new-age superbugs. Heavy
288 usage of disinfectants like quaternary ammonium compounds and heavy metals is major
289 activity in hospital settings (15). Yet, hospital acquired infections with multi-drug resistant
290 bacteria are persistent and prevalent. Hence, the effectiveness of cleanliness is always under
291 question as the hospital settings are shown to be rapidly contaminated by deadly pathogens
292 with multidrug resistant organisms like: *Acinetobacter* spp., *Klebsiella pneumonia* etc. (36,

293 37). These nosocomial pathogens have been shown to persist in the hospital environment
294 from some days to even months (38). In the present study, we have proposed that it is
295 continuous biocide exposure, which might have favored the pathogen towards acquisition and
296 maintenance of resistance through integron mediated rapid evolution. Since, ISCR element,
297 along with antibiotic resistant genes, also carried genes for biocide and heavy metal
298 resistance. Hence, the overdose of disinfectants and reduced susceptibility to them has acted
299 as a potential selection for antibiotic resistance (39). Such co-evolution of antibiotic and
300 biocide resistance has already been reported in literature for deadly pathogens (40, 41).
301 Interestingly, biocide exposure of nosocomial pathogens have been reported to select for
302 multidrug resistance (42).

303 Complete genome dynamics of SM866 revealed exclusive DRs encoding for various efflux
304 pumps, mobile elements, transcriptional regulators etc. Multidrug efflux transporters of RND
305 family, MFS family, MATE and ABC transporters families etc. have been already reported in
306 *Salmonella enterica*, *Mycobacterium tuberculosis* contributing towards evolution of extreme
307 drug resistance (43-45). In the present study, we are reporting similar condition for the first
308 time in a novel genomospecies of *S. maltophilia* complex in context of chromosomally
309 encoded exclusive dynamic regions. Furthermore, high patient density in developing country
310 like India complicates the situation as seen by the potential of the integrons to rapidly
311 diversify at chromosomal level and also to move across distant pathogens. Such a scenario
312 results in evolution of complete drug resistance in nosocomial settings challenging their
313 successful treatment.

314 Hence, in the present study, by investigation of the complete genome obtained using
315 transformative long read technology, we are able to successfully reveal the cryptic
316 evolutionary radiation in the emerging novel genomospecies of *S. maltophilia*. More such
317 complete genome investigations of the superbugs at global level are required for a better
318 understanding of the clinical risks of reduced susceptibility towards biocides, which has got
319 little attention than antibiotic resistance. Further, present study emphasizes on awareness of
320 targeted use of disinfectants restricting to the clinical benefit and not in indiscriminate
321 manner.

322 **Conclusion**

323 Complete genome sequencing allowed us to establish conclusively that *sul* genes are of
324 chromosomal origin in our XDR strain. However, presence of two chromosomal integrons in a

325 nosocomial strain indicated importance of chromosomal route and extent of selection leading to
326 emergence of such a complex strain even in clinical settings. Further, Gillings noted in his review
327 that chromosomal integrons in contrast to transient plasmid borne integrons, can become site to
328 generate genomic and phenotypic complexity. In clinical setting, integrons are reported to be of
329 plasmid origin and thought to be the result of human intervention (46). In this context, finding of a
330 nosocomial pathogenic strain with two chromosomal integrons with complex genetic cassette
331 suggests that clinical settings, like in environmental counterparts, have emerged as hotspots for
332 integron evolution. Presence of antimicrobial genes in both integrons of a novel sequence type
333 suggests nature of extreme selection happening in the hospital settings and further possibility of
334 many more novel sequence types. Particularly, Indian hospitals are tropical clinical settings, having
335 high patient density and heavy use of antimicrobials. Hence, a coordinated global research effort
336 by integrating emerging long read technologies is needed for surveillance of usage of biocides for
337 hygiene and linked/ rapid co-evolution of extreme drug resistance in nosocomial pathogens.

338 **Author statements**

339 **Conflicts of interest**

340 The authors declare that there are no conflicts of interest.

341 **Funding information**

342 All the work in the present study was supported by “GEAR- Genomic Evolutionary and Big-Data
343 Analytic Strategies to address Antimicrobial Resistance” (MLP0016). The authors declare that the
344 research was conducted in the absence of any commercial or financial relationships that could be
345 construed as a potential conflict of interest.

346 **Ethical approval**

347 Ethics approval and each patient’s written consent was not required as it was a part of routine
348 clinical testing.

349

350 **References**

- 351 1. Kaur P, Gautam V, Tewari R. Distribution of class 1 integrons, sul1 and sul2 genes among
352 clinical isolates of *Stenotrophomonas maltophilia* from a tertiary care hospital in North India.
353 *Microbial Drug Resistance*. 2015;21(4):380-5.
- 354 2. Nicodemo A, Paez JG. Antimicrobial therapy for *Stenotrophomonas maltophilia* infections.
355 *European journal of clinical microbiology & infectious diseases*. 2007;26(4):229-37.

- 356 3. Wang YL, Scipione MR, Dubrovskaya Y, Papadopoulos J. Monotherapy with
357 fluoroquinolone or trimethoprim-sulfamethoxazole for treatment of *Stenotrophomonas maltophilia*
358 infections. *Antimicrobial agents and chemotherapy*. 2014;58(1):176-82.
- 359 4. Herrera-Heredia SA, Pezina-Cantú C, Garza-González E, Bocanegra-Ibarias P, Mendoza-
360 Olazarán S, Morfín-Otero R, et al. Risk factors and molecular mechanisms associated with
361 trimethoprim-sulfamethoxazole resistance in *Stenotrophomonas maltophilia* in Mexico. *Journal of*
362 *medical microbiology*. 2017;66(8):1102-9.
- 363 5. Barbolla R, Catalano M, Orman BE, Famiglietti A, Vay C, Smayevsky J, et al. Class 1
364 integrons increase trimethoprim-sulfamethoxazole MICs against epidemiologically unrelated
365 *Stenotrophomonas maltophilia* isolates. *Antimicrobial agents and chemotherapy*. 2004;48(2):666-9.
- 366 6. Hu L-F, Chang X, Ye Y, Wang Z-X, Shao Y-B, Shi W, et al. *Stenotrophomonas maltophilia*
367 resistance to trimethoprim/sulfamethoxazole mediated by acquisition of *sul* and *dfrA* genes in a
368 plasmid-mediated class 1 integron. *International journal of antimicrobial agents*. 2011;37(3):230-4.
- 369 7. Morsi S, Sharaf H, Gerges M. Association of *sul* genes and class 1 integron with
370 trimethoprim-sulfamethoxazole Resistance in *Stenotrophomonas maltophilia* clinical isolates in
371 Zagazig University, Egypt. *African Journal of Clinical and Experimental Microbiology*.
372 2016;17(3):158-65.
- 373 8. Maravić A, Skočibušić M, Fredotović Ž, Cvjetan S, Šamanić I, Puizina J. Characterization of
374 environmental CTX-M-15-producing *Stenotrophomonas maltophilia*. *Antimicrobial agents and*
375 *chemotherapy*. 2014;58(10):6333-4.
- 376 9. Boerlin P, Reid-Smith RJ. Antimicrobial resistance: its emergence and transmission. *Animal*
377 *Health Research Reviews*. 2008;9(2):115-26.
- 378 10. Bennett P. Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic
379 resistance genes in bacteria. *British journal of pharmacology*. 2008;153(S1):S347-S57.
- 380 11. Yousfi K, Touati A, Lefebvre B, Fournier É, Côté J-C, Soualhine H, et al. A novel plasmid,
381 pSx1, harboring a new Tn 1696 derivative from extensively drug-resistant *Shewanella xiamenensis*
382 encoding OXA-416. *Microbial Drug Resistance*. 2017;23(4):429-36.
- 383 12. Toleman MA, Bennett PM, Bennett DM, Jones RN, Walsh TR. Global emergence of
384 trimethoprim/sulfamethoxazole resistance in *Stenotrophomonas maltophilia* mediated by acquisition
385 of *sul* genes. *Emerging infectious diseases*. 2007;13(4):559.
- 386 13. Toleman MA, Bennett PM, Walsh TR. ISCR elements: novel gene-capturing systems of the
387 21st century? *Microbiology and molecular biology reviews*. 2006;70(2):296-316.
- 388 14. Kumarasamy KK, Toleman MA, Walsh TR, Bagaria J, Butt F, Balakrishnan R, et al.
389 Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular,
390 biological, and epidemiological study. *The Lancet infectious diseases*. 2010;10(9):597-602.

- 391 15. Gupta P, Bhatia M, Gupta P, Omar BJ. Emerging biocide resistance among multidrug-
392 resistant bacteria: Myth or reality? A pilot study. *Journal of pharmacy & bioallied sciences*.
393 2018;10(2):96.
- 394 16. Makwana M, Grover CR, Kumar N. Biocides resistance profiles of biofilm forming bacteria
395 of dairy niche and their control. *Journal homepage: <http://www.ijemas.com>*. 2018;7(02):2018.
- 396 17. Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y, Seo H, et al. Introducing EzBioCloud: A
397 taxonomically united database of 16S rRNA and whole genome assemblies. *International Journal of*
398 *Systematic and Evolutionary Microbiology*. 2016.
- 399 18. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: resolving bacterial genome assemblies
400 from short and long sequencing reads. *PLoS computational biology*. 2017;13(6):e1005595.
- 401 19. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an
402 integrated tool for comprehensive microbial variant detection and genome assembly improvement.
403 *PloS one*. 2014;9(11):e112963.
- 404 20. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, et al. NCBI
405 prokaryotic genome annotation pipeline. *Nucleic acids research*. 2016;44(14):6614-24.
- 406 21. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the
407 quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome*
408 *research*. 2015;25(7):1043-55.
- 409 22. Bushnell B. BBMap: a fast, accurate, splice-aware aligner. 2014.
- 410 23. Parte AC. LPSN-List of Prokaryotic names with Standing in Nomenclature (bacterio. net), 20
411 years on. *International journal of systematic and evolutionary microbiology*. 2018;68(6):1825-9.
- 412 24. Lagesen K, Hallin P, Rødland EA, Stærfeldt H-H, Rognes T, Ussery DW. RNAmmer:
413 consistent and rapid annotation of ribosomal RNA genes. *Nucleic acids research*. 2007;35(9):3100-8.
- 414 25. Thompson JD, Gibson TJ, Higgins DG. Multiple sequence alignment using ClustalW and
415 ClustalX. *Current protocols in bioinformatics*. 2003(1):2.3. 1-2.3. 22.
- 416 26. da Silva AR, Ferro JA, Reinach FdC, Farah CS, Furlan LR, Quaggio RB, et al. Comparison of
417 the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature*.
418 2002;417(6887):459.
- 419 27. Segata N, Börnigen D, Morgan XC, Huttenhower C. PhyloPhlAn is a new method for
420 improved phylogenetic and taxonomic placement of microbes. *Nature communications*. 2013;4:2304.
- 421 28. Lee I, Kim YO, Park S-C, Chun J. OrthoANI: an improved algorithm and software for
422 calculating average nucleotide identity. *International journal of systematic and evolutionary*
423 *microbiology*. 2016;66(2):1100-3.
- 424 29. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence-based species
425 delimitation with confidence intervals and improved distance functions. *BMC bioinformatics*.
426 2013;14(1):60.

- 427 30. Alikhan N-F, Petty NK, Zakour NLB, Beatson SA. BLAST Ring Image Generator (BRIG):
428 simple prokaryote genome comparisons. *BMC Genomics*. 2011;12(1):402.
- 429 31. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+:
430 architecture and applications. *BMC bioinformatics*. 2009;10(1):421.
- 431 32. Youenou B, Favre-Bonte S, Bodilis J, Brothier E, Dubost A, Muller D, et al. Comparative
432 genomics of environmental and clinical *Stenotrophomonas maltophilia* strains with different antibiotic
433 resistance profiles. *Genome biology and evolution*. 2015;7(9):2484-505.
- 434 33. Patil PP, Kumar S, Midha S, Gautam V, Patil PB. Taxonogenomics reveal multiple novel
435 genomospecies associated with clinical isolates of *Stenotrophomonas maltophilia*. *Microbial
436 genomics*. 2018;4(8).
- 437 34. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer.
438 *Bioinformatics*. 2011;27(7):1009-10.
- 439 35. Patil PP, Kumar S, Midha S, Gautam V, Patil PB. Taxonogenomics reveal multiple novel
440 genomospecies associated with clinical isolates of *Stenotrophomonas maltophilia*. *Microbial
441 genomics*. 2018.
- 442 36. Chung DR, Song J-H, Kim SH, Thamlikitkul V, Huang S-G, Wang H, et al. High prevalence
443 of multidrug-resistant nonfermenters in hospital-acquired pneumonia in Asia. *American journal of
444 respiratory and critical care medicine*. 2011;184(12):1409-17.
- 445 37. Abdelaziz A, Sonbol F, Elbanna T, El-Ekhnawy E. Exposure to Sublethal Concentrations of
446 Benzalkonium Chloride Induces Antimicrobial Resistance and Cellular Changes in *Klebsiellae
447 pneumoniae* Clinical Isolates. *Microbial Drug Resistance*. 2019.
- 448 38. Kramer A, Schwebke I, Kampf G. How long do nosocomial pathogens persist on inanimate
449 surfaces? A systematic review. *BMC infectious diseases*. 2006;6(1):130.
- 450 39. Harbarth S, Soh ST, Horner C, Wilcox M. Is reduced susceptibility to disinfectants and
451 antiseptics a risk in healthcare settings? A point/counterpoint review. *Journal of Hospital Infection*.
452 2014;87(4):194-202.
- 453 40. Buffet-Bataillon S, Tattevin P, Maillard J-Y, Bonnaure-Mallet M, Jolivet-Gougeon A. Efflux
454 pump induction by quaternary ammonium compounds and fluoroquinolone resistance in bacteria.
455 *Future microbiology*. 2016;11(1):81-92.
- 456 41. Gnanadhas DP, Marathe SA, Chakravorty D. Biocides–resistance, cross-resistance
457 mechanisms and assessment. *Expert opinion on investigational drugs*. 2013;22(2):191-206.
- 458 42. Webber MA, Whitehead RN, Mount M, Loman NJ, Pallen MJ, Piddock LJ. Parallel
459 evolutionary pathways to antibiotic resistance selected by biocide exposure. *Journal of antimicrobial
460 chemotherapy*. 2015;70(8):2241-8.
- 461 43. Balganes M, Dinesh N, Sharma S, Kuruppath S, Nair AV, Sharma U. Efflux pumps of
462 *Mycobacterium tuberculosis* play a significant role in antituberculosis activity of potential drug
463 candidates. *Antimicrobial agents and chemotherapy*. 2012;56(5):2643-51.

- 464 44. Sharma S, Kumar M, Sharma S, Nargotra A, Koul S, Khan IA. Piperine as an inhibitor of
465 Rv1258c, a putative multidrug efflux pump of Mycobacterium tuberculosis. Journal of Antimicrobial
466 Chemotherapy. 2010;65(8):1694-701.
- 467 45. Willers C, Wentzel JF, Du Plessis LH, Gouws C, Hamman JH. Efflux as a mechanism of
468 antimicrobial drug resistance in clinical relevant microorganisms: the role of efflux inhibitors. Expert
469 opinion on therapeutic targets. 2017;21(1):23-36.
- 470 46. Gillings MR. Integrins: past, present, and future. Microbiol Mol Biol Rev. 2014;78(2):257-
471 77.

472

473 **Data bibliography**

474 Complete genome resource of SM866 used for the cluster generation. Kumar S, Bansal K, Patil PP,
475 Kaur A, Kaur S, Gautam V, Patil PB. **CP031058**.

476

477 The genome information used in phylogenetic analysis and resistome analysis.

- 478 1) Patil PP, Midha S, Patil PB. GenBank.
479 https://www.ncbi.nlm.nih.gov/nuccore/NZ_JALV00000000 (2016).
- 480 2) Patil PP, Midha S, Patil PB. GenBank.
481 https://www.ncbi.nlm.nih.gov/nuccore/NZ_LLXW00000000.1 (2016).
- 482 3) Kyrpides N, Huntemann M, Han J, Chen A, Mavromatis K et al. GenBank.
483 <https://www.ncbi.nlm.nih.gov/nuccore/485091714?report=genbank> (2013).
- 484 4) Patil PP, Midha S, Patil PB. GenBank.
485 https://www.ncbi.nlm.nih.gov/nuccore/NZ_LLXV00000000.1 (2016).
- 486 5) Patil PP, Midha S, Patil PB. GenBank.
487 https://www.ncbi.nlm.nih.gov/nuccore/NZ_LDJN00000000.1 (2016).
- 488 6) Patil PP, Midha S, Patil PB. GenBank.
489 https://www.ncbi.nlm.nih.gov/nuccore/NZ_LLXT00000000.1 (2016).
- 490 7) Patil PP, Midha S, Patil PB. GenBank.
491 https://www.ncbi.nlm.nih.gov/nuccore/NZ_LDJK00000000.1 (2016).
- 492 8) Alavi P, Starcher MR, Thallinger G, Zachow C, Mueller H et al. GenBank.
493 <https://www.ncbi.nlm.nih.gov/nuccore/627787876?report=genbank> (2014).
- 494 9) Patil PP, Midha S, Patil PB. GenBank.
495 https://www.ncbi.nlm.nih.gov/nuccore/NZ_LLXU00000000.1 (2016).
- 496 10) Patil PP, Midha S, Patil PB. GenBank.
497 https://www.ncbi.nlm.nih.gov/nuccore/NZ_LDJH00000000.1 (2016).
- 498 11) Patil PP, Midha S, Patil PB. GenBank.
499 https://www.ncbi.nlm.nih.gov/nuccore/NZ_LDJM00000000.1 (2016).

- 500 12) Patil PP, Midha S, Patil PB. GenBank.
501 https://www.ncbi.nlm.nih.gov/nuccore/NZ_LDJO00000000.1 (2016).
- 502 13) Patil PP, Midha S, Patil PB. GenBank.
503 https://www.ncbi.nlm.nih.gov/nuccore/NZ_LDJP00000000.1 (2016).
- 504 14) Patil PP, Midha S, Patil PB. GenBank.
505 <https://www.ncbi.nlm.nih.gov/nuccore/LLXS00000000> (2016).
- 506 15) Patil PP, Midha S, Patil PB. GenBank.
507 https://www.ncbi.nlm.nih.gov/nuccore/NZ_LDJI00000000.1 (2016).
- 508 16) Patil PP, Midha S, Patil PB. GenBank.
509 https://www.ncbi.nlm.nih.gov/nuccore/NZ_LDJG00000000.1 (2016).
- 510 17) Patil PP, Midha S, Patil PB. GenBank.
511 https://www.ncbi.nlm.nih.gov/nuccore/NZ_LDJJ00000000.1 (2016).
- 512 18) Patil PP, Midha S, Patil PB. GenBank.
513 https://www.ncbi.nlm.nih.gov/nuccore/NZ_LDJL00000000.1 (2016).
- 514 19) Patil PP, Midha S, Kumar S, Gautam V, Patil PB. GenBank.
515 https://www.ncbi.nlm.nih.gov/nuccore/NZ_LXXY00000000.1 (2018).
- 516 20) Patil PP, Midha S, Kumar S, Gautam V, Patil PB. GenBank.
517 <https://www.ncbi.nlm.nih.gov/nuccore/LXXU00000000.1> (2018).
- 518 21) Patil PP, Midha S, Kumar S, Gautam V, Patil PB. GenBank.
519 <https://www.ncbi.nlm.nih.gov/nuccore/LXXR00000000.1> (2018).
- 520 22) Patil PP, Midha S, Kumar S, Gautam V, Patil PB. GenBank.
521 <https://www.ncbi.nlm.nih.gov/nuccore/LXXZ00000000.1> (2018).
- 522 23) Patil PP, Midha S, Kumar S, Gautam V, Patil PB. GenBank.
523 <https://www.ncbi.nlm.nih.gov/nuccore/LXXM00000000.1> (2018).
- 524 24) da Silva AR, Ferro JA, Reinach FD, Farah CS, Furlan LR, Quaggio RB, Monteiro-Vitorello
525 CB, Van Sluys MA, Almeida NA, Alves LM, Do Amaral AM et al. GenBank.
526 https://www.ncbi.nlm.nih.gov/nuccore/NC_003902.1 (2001).
- 527 The well-characterised resistance genes were fetched from the genome of *S. maltophilia* K279a and
528 used as a query in the BLAST analysis with sequenced *S. maltophilia* isolates.
- 529 25) Crossman LC, Gould VC, Dow JM, Vernikos GS, Okazaki A et al. GenBank.
530 <https://www.ncbi.nlm.nih.gov/nuccore/190572091?report=genbank> (2008).

531 **Figures and tables**

532 **Figure 1: PhyloPhlAn phylogeny and taxonogenomic analysis** of representative strains of
533 *Stenotrophomonas* genus along with SM866 (in red box) and Smc is highlighted in blue box.
534 ANI and dDDH values for all the representative strains taking SM866 as reference are shown
535 in the heat maps next to the tree generated on the scale as shown in the figure.

536 **Figure 2: Circular representation of SM866 along with other Smc genomospecies.** Rings
537 are color coded for different strains. The innermost two rings represent GC content (%) and
538 GC skew. Here, R1-R14 are DRs exclusively present in SM866, annotation for the regions
539 are indicated adjacent to regions and genomic locations of the DRs is provided in
540 supplementary table 1.

541 **Figure 3: Resistome heat map of Smc** distribution of antibiotic resistance and efflux
542 pump genes across Smc. Genes and strains are indicated in x-axis and y-axis respectively.
543 Here, presence and absence of gene is represented by green and yellow colors respectively.
544 Here, SM866 is in highlighted in red box.

545 **Figure 4 (A) Circular representation of SM866** indicating chromosomal location of
546 cassette I and cassette II. **(B) Schematic representation of antimicrobial genes in**
547 **cassettes I and II** ISCR elements are shown in red color and antibiotic resistance genes are
548 in green color.

549 **Table 1:** SM866 zone of inhibitions for antibiotics: TMP-SMX: trimethoprim-
550 sulfamethoxazole (25mcg), CHL: chloramphenicol (30 mcg), LEV: levofloxacin (5 mcg)
551 and CAZ: ceftazidime (30 mcg).

552

553 **Supplementary data**

554 **Supplementary figure 1(A) 16S rRNA phylogeny.** Here, SM 866 is in red box and Smc is
555 in blue box. **(B) eBURST analysis of Smc.** Here, ST-366 is highlighted.

556 **Supplementary figure 2(A) Electrophoretic gel image** of plasmid screening with Fosmid
557 as a positive control and *S. maltophilia* ATCC 13637 (Type strain) as negative control.
558 Loading order: Lane1: λ /hind III ladder; Lane 2: Fosmid; Lane 3: Fosmid induced; Lane 4: *S.*
559 *maltophilia* SM866; Lane 5: *S. maltophilia* ATCC13637; Lane 6: λ monocot ladder. **(B)**
560 **Electrophoretic gel image** of the plasmid screen after treatment with ATP-dependent
561 DNase. Loading order: Lane1: λ /hind III ladder; Lane 2: Fosmid; Lane 3: Fosmid (digested);
562 Lane 4: Fosmid induced; Lane 5: Fosmid induced (digested); Lane 6: steno SM866; Lane 7:
563 steno SM866 (digested); Lane 8: *S. maltophilia* ATCC13637; Lane 9: *S. maltophilia*
564 ATCC13637 (digested); Lane 10: λ monocot ladder.

565 **Supplementary table 1: Dynamic regions** Genomic coordinates for the DRs exclusive to
566 SM866.

567

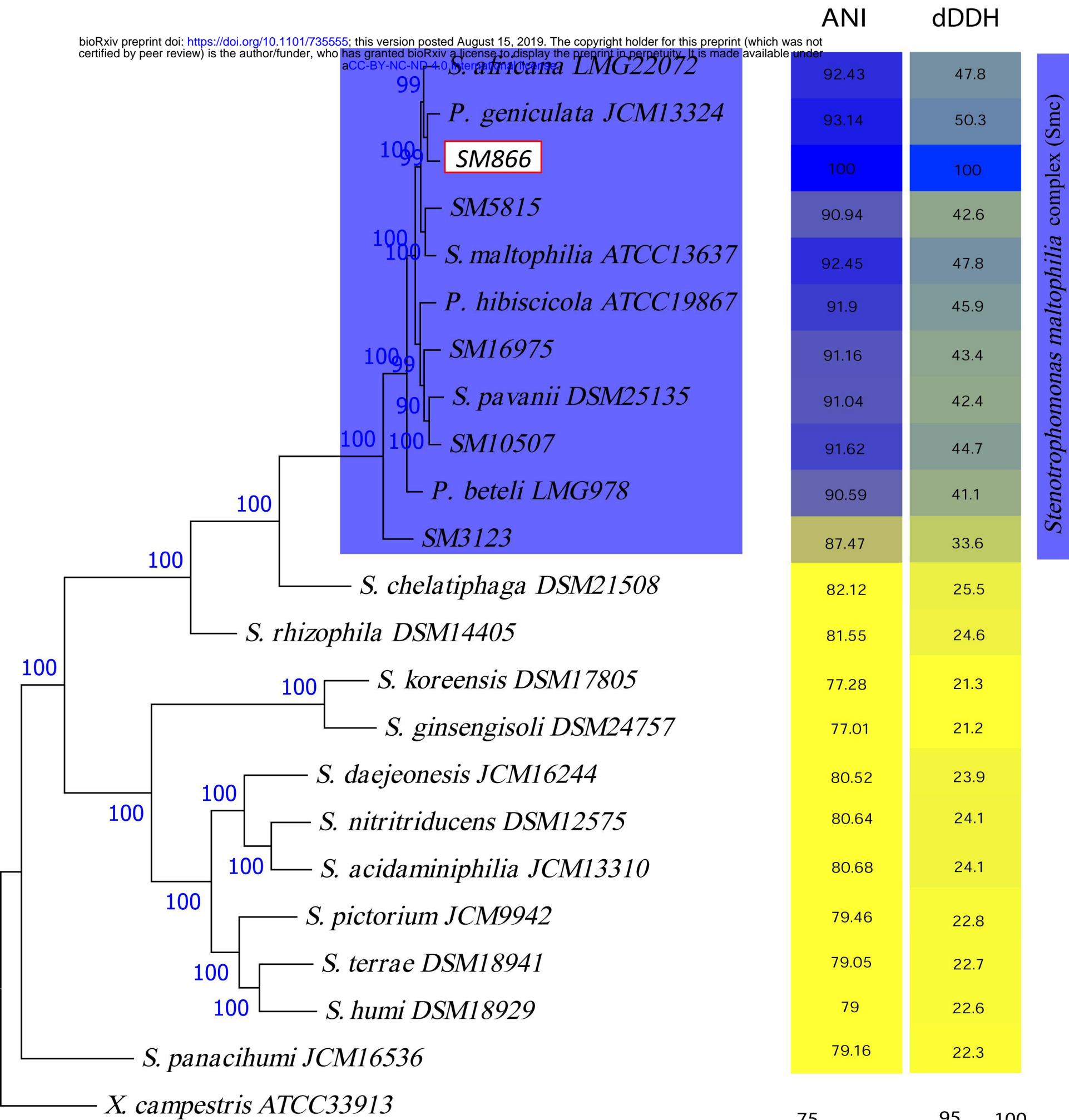
568 **Table 1:** SM866 zone of inhibitions for antibiotics: TMP-SMX: trimethoprim-
569 sulfamethoxazole (25mcg), CHL: chloramphenicol (30 mcg), LEV: levofloxacin (5 mcg)
570 and CAZ: ceftazidime (30 mcg).

571

Antibiotics	SM866 (zone diameter in mm)	Result interpretation		
		(\geq) Sensitive (S) (in mm)	Intermediate (I) (in mm)	(\leq) Resistant (R) (in mm)
TMP-SMX	0 (R)	16	11 to 15	10
CHL	0 (R)	18	13 to 17	12
LEV	11 (R)	17	14 to 16	13
CAZ	0 (R)	21	18 to 20	17

572

573



0.20

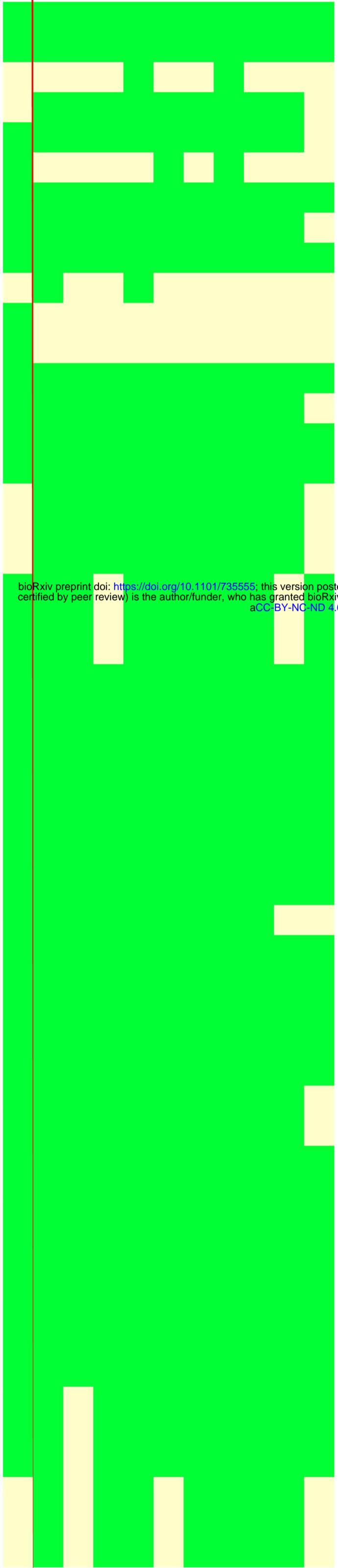
ANI 75 95 100
dDDH 20 70 100



global



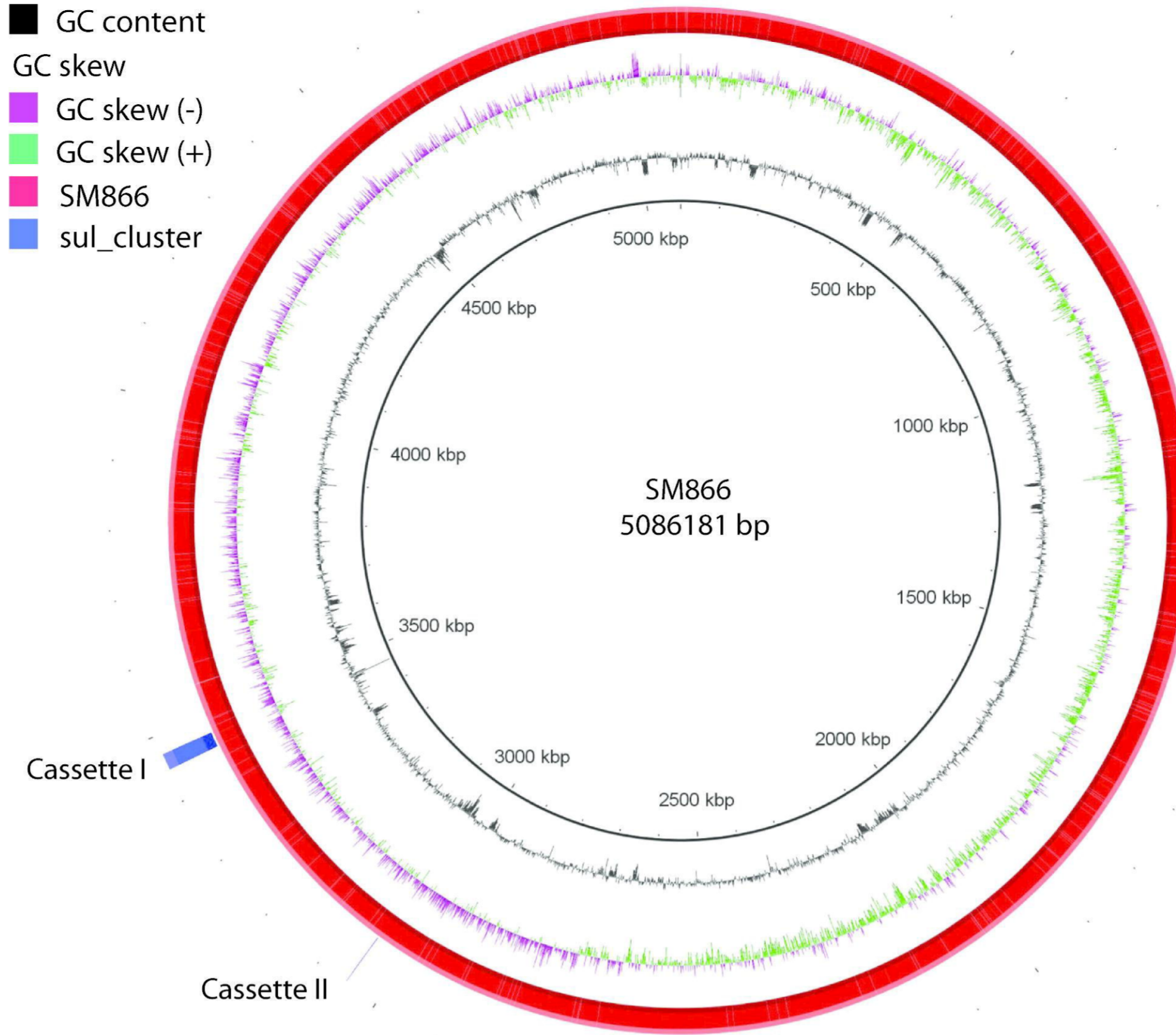
SM866
 S. africana LMG22072
 P. geniculata JCM13324
 SM5815
 S. maltophilia ATCC13637
 S. hibscicola ATCC16867
 SM16975
 S. pavanii DSM25135
 SM10507
 P. beteli LMG978
 SM3123



Dimethyladenosine transferase (ksgA)
 Streptomycin 3''-phosphotransferase
 Aminoglycoside 6'-N-acetyltransferase (aac(6')-iz)
 ampC
 Metallo-beta-lactamase I1 (blaL1)
 Beta-lactamase class C
 Metallo-beta-lactamase family protein
 Fluoroquinolone resistance protein qnrB (smqnr)
 Phosphomannomutase/phosphoglucomutase (spgM)
 Chloramphenicol acetyltransferase (cat)
 sul1
 sul2
 smrA
 macA
 OMF
 macB
 fuaA
 fuaB
 fuaC
 fuaC
 fuaC
 fuaC
 pmpM
 norM
 OMF
 emrA
 emrB
 mdtD
 bcr
 sugE
 emrE
 smeV
 smeW
 smeX
 smeY
 smeZ
 smeG
 smeH
 smeN
 smeM
 smeP
 smeO
 tolC
 smeF
 smeE
 smeD
 smel
 smeJ
 smeK
 smeC
 smeB
 smeA

bioRxiv preprint doi: <https://doi.org/10.1101/735555>; this version posted August 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

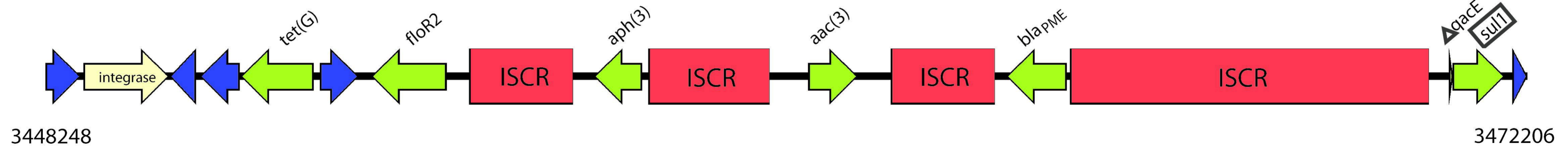
■ GC content
 GC skew
 ■ GC skew (-)
 ■ GC skew (+)
 ■ SM866
 ■ sul_cluster



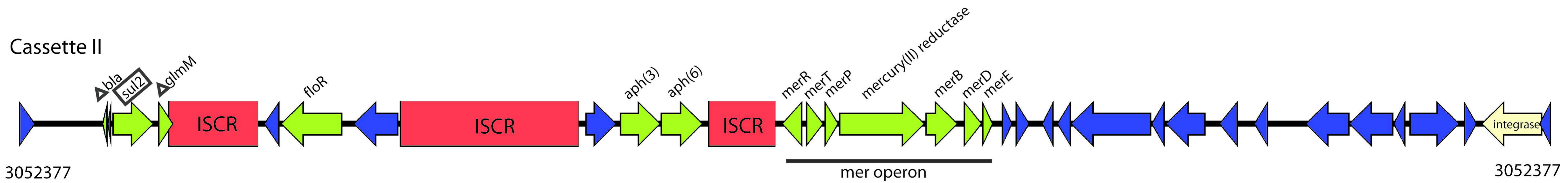
bioRxiv preprint doi: <https://doi.org/10.1101/735555>; this version posted August 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(A)

Cassette I



Cassette II



(B)