


GENOME REPORT

Open Access



Complete genome of *Enterobacter sichuanensis* strain SGAir0282 isolated from air in Singapore

Akira Uchida¹, Hie Lim Kim^{1,2}, Rikky W. Purbojati¹, Vineeth Kodengil Vettath¹, Anjali B. Gupta¹, James N. I. Houghton¹, Caroline Chénard¹, Anthony Wong¹, Megan E. Clare¹, Kavita K. Kushwaha¹, Deepa Panicker¹, Alexander Putra¹, Cassie E. Heinle¹, Balakrishnan N. V. Premkrishnan¹, Ana Carolina M. Junqueira³, Daniela I. Drautz-Moses¹ and Stephan C. Schuster^{1,4*} 

Abstract

Background: *Enterobacter cloacae* complex (ECC) bacteria, such as *E. cloacae*, *E. sichuanensis*, *E. kobei*, and *E. roggenkampii*, have been emerging as nosocomial pathogens. Many strains isolated from medical clinics were found to be resistant to antibiotics, and in the worst cases, acquired multidrug resistance. We present the whole genome sequence of SGAir0282, isolated from the outdoor air in Singapore, and its relevance to other ECC bacteria by in silico genomic analysis.

Results: Complete genome assembly of *E. sichuanensis* strain SGAir0282 was generated using PacBio RSII and Illumina MiSeq platforms, and the datasets were used for de novo assembly using Hierarchical Genome Assembly Process (HGAP) and error corrected with Pilon. The genome assembly consisted of a single contig of 4.71 Mb and with a G+C content of 55.5%. No plasmid was detected in the assembly. The genome contained 4371 coding genes, 83 tRNA and 25 rRNA genes, as predicted by NCBI's Prokaryotic Genome Annotation Pipeline (PGAP). Among the genes, the antibiotic resistance related genes were included: Streptothricin acetyltransferase (SatA), fosfomycin resistance protein (FosA) and metal-dependent hydrolases of the beta-lactamase superfamily I (BLI).

Conclusion: Based on whole genome alignment and phylogenetic analysis, the strain SGAir0282 was identified to be *Enterobacter sichuanensis*. The strain possesses gene clusters for virulence, disease and defence, that can also be found in other multidrug resistant ECC type strains.

Keywords: Airborne bacteria, *Enterobacter cloacae* complex, Whole genome sequencing, Multidrug resistant

Background

Species belonging to *Enterobacter cloacae* complex (ECC) are commonly found in environment [1, 2], and are widely known to be opportunistic pathogens. In the past decades, ECC such as *E. hormaechei*, *E. sichuanensis*, *E. asburiae*, *E. kobei*, and *E. roggenkampii*, have been

a global health concern because of wide-spread antibiotic resistances and new acquisition of multidrug resistances [3–6]. They are intrinsically resistant to beta-lactam antibiotics, which are the most commonly used antibiotics such as ampicillin, amoxicillin and first generation cephalosporins [7].

Carbapenems are reserved to treat infections caused by multidrug resistant bacteria including ECC. An increasing number of patients that have been infected with these nosocomial pathogens can no longer be

*Correspondence: SCSchuster@ntu.edu.sg

¹ Singapore Centre for Environmental Life Sciences Engineering, Nanyang

Technological University, Singapore, Singapore

Full list of author information is available at the end of the article



© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

successfully treated due to the ECC's acquired resistances to carbapenems and other classes of antibiotics after long term exposure to the drugs. It has been reported that the gut microbiome community of healthy volunteers administered with cefprozil, one of beta-lactam antibiotics, drifts towards a higher abundance of antibiotic resistant bacteria including ECC [8]. The affected gut microbiota can end up in the sewage of hospitals, and turn the wastewater facilities into a reservoir for antibiotic resistant bacteria [9, 10]. As such, there is an increasing risk of spreading multidrug resistant strains and associated genes by releasing them into the environment through wastewater and other means [10].

In this study, we report the complete genome sequence of the strain SGAir0282, a bacterium that was isolated from outdoor air at a university sports field. Using the whole-genome dataset, we identified its taxon as a recently defined species, *E. sichuanensis*, which is a member of the ECC. To our knowledge, this is the first report that ECC was isolated from air. Considering the importance of the studies on ECC that were mostly isolated from medical clinics/hospitals and reported as multidrug resistant species, the high-quality genome sequence of strain SGAir0282 from our study will make a significant contribution to this field.

Methods

Isolation and culture conditions

We generated a large-scale isolate collection of airborne microorganisms using cultivation methods to produce whole-genome sequence data of those isolates. The genome data is to improve the taxonomy identification of metagenomic data of air microbiome [11]. To have the vast diversity of the airborne microbiota, we used various media for agar plates as many as possible, and sampled in different time and locations. Air samples were collected with an Single-staged Andersen-type air sampler (SKC, USA). A total volume of 23.4 L of air was drawn against the agar plates over 2 minutes.

Strain SGAir0282 was isolated from outdoor air (global position system coordinates 1.349° N, 103.689° E) by impacting air onto a Potato Dextrose Agar (PDA, Sigma-Aldrich, USA), and the plate was aerobically incubated until colony formation at 30 °C. A single colony was picked and further streaked on Tryptic Soy Agar (Becton Dickinson, USA) to obtain a pure clonal colonies. One of these successfully sub-cultured colonies was the strain SGAir0282, which was then cultured in Lysogeny Broth media (Merck, USA) overnight at 30 °C followed by DNA extraction.

DNA extraction and sequencing

DNA was extracted from an overnight culture of the isolate using the Wizard Genomic DNA Purification Kit (Promega, USA), according to the manufacturer's recommended protocol. Sequencing was performed on the Pacific Biosciences (PacBio) RSII sequencer, using a library that was prepared with SMRTbell Template Prep Kit 1.0 (Pacific Biosciences, USA). The library used for sequencing on the Illumina MiSeq was prepared with the TruSeq Nano DNA Library Preparation Kit (Illumina, USA).

Genome assembly and annotation

De novo assembly was performed with the Hierarchical Genome Assembly Process (HGAP) version 3 [12], which is part of PacBio SMRT Analysis 2.3.0 package and subsequently polished with Quiver [12]. Error correction was performed with Pilon version 1.16 using 300-bp MiSeq paired-end reads and following parameters (-tracks -changes -vcf -fix all -mindepth 0.1 -mingap 10 -minmq 30 -minqual 20 -K 47) [13]. Finally, the genome sequence was circularized using Circulator version 1.1.4 [14]. Gene prediction was carried out with NCBI's Prokaryotic Genome Annotation Pipeline (PGAP) version 4.2 [15]. Rapid Annotations using Subsystems Technology tool kit (RASTtk) was used for additional feature annotation of the assembled contig [16]. Default parameters were used for all steps unless otherwise stated above.

Genome sequence analysis for taxonomy identification

Average Nucleotide Identity (ANI) was calculated to identify the strain SGAir0282 to species level with a custom PERL script against all 10,744 bacterial genomes in the NCBI Reference Sequence Database (RefSeq; downloaded in April 2019) [17].

We also performed NCBI's BLASTn using query sequence of 16S rRNA of the SGAir0282 against nucleotide (nt) database with default settings. Based on the BLASTn result, we retrieved whole-genome sequence of every strain which met following two criteria. Firstly, the 16S rRNA gene sequence had high similarity to the one from strain SGAir0282 (higher than 99.5% and 100% query coverage). Secondly, the level of genome assembly was 'complete genome' or 'chromosome', according to the NCBI nucleotide database. As exceptional, we included genome sequence of *E. sichuanensis* WCHECL1597 (BLASTn: 99.4% similarity and 99% query coverage) to analysis, even the genome assembly quality was not high enough, because *E. sichuanensis* reference genome showed the highest value of ANI to strain SGAir0282. In total, 27 whole-genome datasets, including SGAir0282, were used for further genome analyses below.

Core genome alignment was performed on Parsnp [18] with SGAir0282 as a reference. Gingr [18] was used to visualise aligned genomes and to export the Newick format of a phylogenetic tree reconstructed by Parsnp. The tree was plotted by an online tool, phylogenetic tree viewer (<http://etetoolkit.org/treeview/>) [19]. MASH was used to calculate ANI between SGAir0282 and other strains [20]. The genome datasets were individually processed by sketching prior to estimating the distance from SGAir0282. We applied kSNP3 [21] to identify single nucleotide polymorphisms (SNPs) and to obtain a distance matrix. The k-mer size was optimized as 19 nucleotides, calculated by a kSNP3-packaged program, kChooser. For annotation of SNPs, SGAir0282 was used as a reference sequence. Variant call format (VCF) generated by kSNP3 was attached as an Additional file 1. Parameters which were not mentioned above were kept at the default value.

Sequence Types (STs) were assigned to the ECC complex strains using multilocus sequence type (MLST) databases (<https://pubmlst.org/ecloacae/>) [22].

Quality assurance

The assembly produced only a single contig and no trace of contamination of other microorganisms was found. Whole genome sequencing was carried out on two sequencing platforms (PacBio RSII and Illumina MiSeq) and led to assembly completion with a circularized chromosome.

Results

HGAP de novo assembly was performed with 53,503 PacBio subreads. The assembly was polished with Quiver and corrected by aligning 858,191 paired-end short reads using Pilon version 1.16. The resulting assembly consisted of a single contig with a total size of 4,711,389 bp, and a mean genome coverage of 84.8- and 110-fold by PacBio and Illumina data, respectively. Chromosomal G+C content was 55.5%. Species identification with ANI analysis resulted in 98.7% similarity to *E. sichuanensis*. BLASTn alignment analysis using 16S rRNA gene of SGAir0282 showed high similarity to *E. cloacae* strain A1137 with 99.8% identity [23] and *E. sichuanensis* WCHECL1597 [3] with 99.4% identity (Additional file 2: Table S1).

We selected the 26 genomes, which are most closely related strains to SGAir0282, based on 16S rRNA genes (Additional file 2: Table S1) and quality of their genome assembly, as described in the method section. The resulting 27 genomes, including SGAir0282, were aligned, and the genetic distance between them were estimated using three different methods.

The result produced by Parsnp [18] was presented as a phylogenetic tree (Fig. 1). In the tree, SGAir0282 was

closest to *E. sichuanensis* WCHECL1597, and second closest to *E. cloacae* A1137 (Fig. 1). The estimated ANI [20] between SGAir0282 and each of the other 26 strains was consistent with the tree. The strain WCHECL1597 showed the smallest distance (0.014) to the strain SGAir0282, and the second shortest strain was the strain A1137 (0.019) (Table 1, Additional file 3: Table S2). Lastly, the distance matrix calculated by kSNP3 [21] also clearly supported the close relationships with the strains WCHECL1597 and A1137 (Additional file 4: Table S3). The robust and high similarity with *E. sichuanensis* WCHECL1597 suggests that SGAir0282 belongs to the species, *E. sichuanensis*.

In the previous report [23], the strain A1137 belonged to ECC, based on a phylogenetic tree reconstructed for a single gene. Our analyses suggested the strain A1137 to be *E. sichuanensis*, as well as SGAir0282.

MLST was assigned for the strain SGAir0282 (Table 1). SGAir0282 had new alleles of *leuS* and *pyrG* genes, while for the other three genes, *dnaA*, *fusA* and *rplB*, it shared the same allele with the strain WCHECL1597. Therefore, SGAir0282 created a new ST (1330) in the database.

The SGAir0282 genome includes a total of 4611 genes, of which 4371 were coding genes. Ribosomal RNAs were reported as 9 copies of 5S and 8 copies each of 16S and 23S. Transfer RNAs were annotated with 83 genes. No evidence of plasmid DNA sequence was found.

As shown in Fig. 2a, RASTtk gene classification was estimated that SGAir0282 contained 53 genes, which were related to virulence, disease and defence. The 53 genes were shared across pathogenic ECC type strains, and their loci are indicated in Fig. 2b. The gene products conferring antibiotic resistance in the SGAir0282 genome were predicted: streptothricin acetyltransferase (SatA), fosfomycin resistance protein (FosA) and metal-dependent hydrolases of the beta-lactamase superfamily I (BLI). Strain SGAir0282 is potentially pathogenic due to the presence of virulence genes that are commonly found in pathogenic ECC type strains.

Discussion

Enterobacter sichuanensis strain SGAir0282 was isolated from air at a side of sports and recreation field where high human activity was observed. The fact that this genus commonly found in soil [1, 2, 24], may suggest that this isolate was stirred up from ground and dispersed into the air prior to being captured onto the agar plate. Hence, our study indicates that environmental bacteria can be aerosolized, underlying that environmental microorganisms can be transported through wind in a form of aerosols.

The strains WCHECL1597 [3] and A1137 [23] were isolated from a chronic renal insufficiency patient's urine

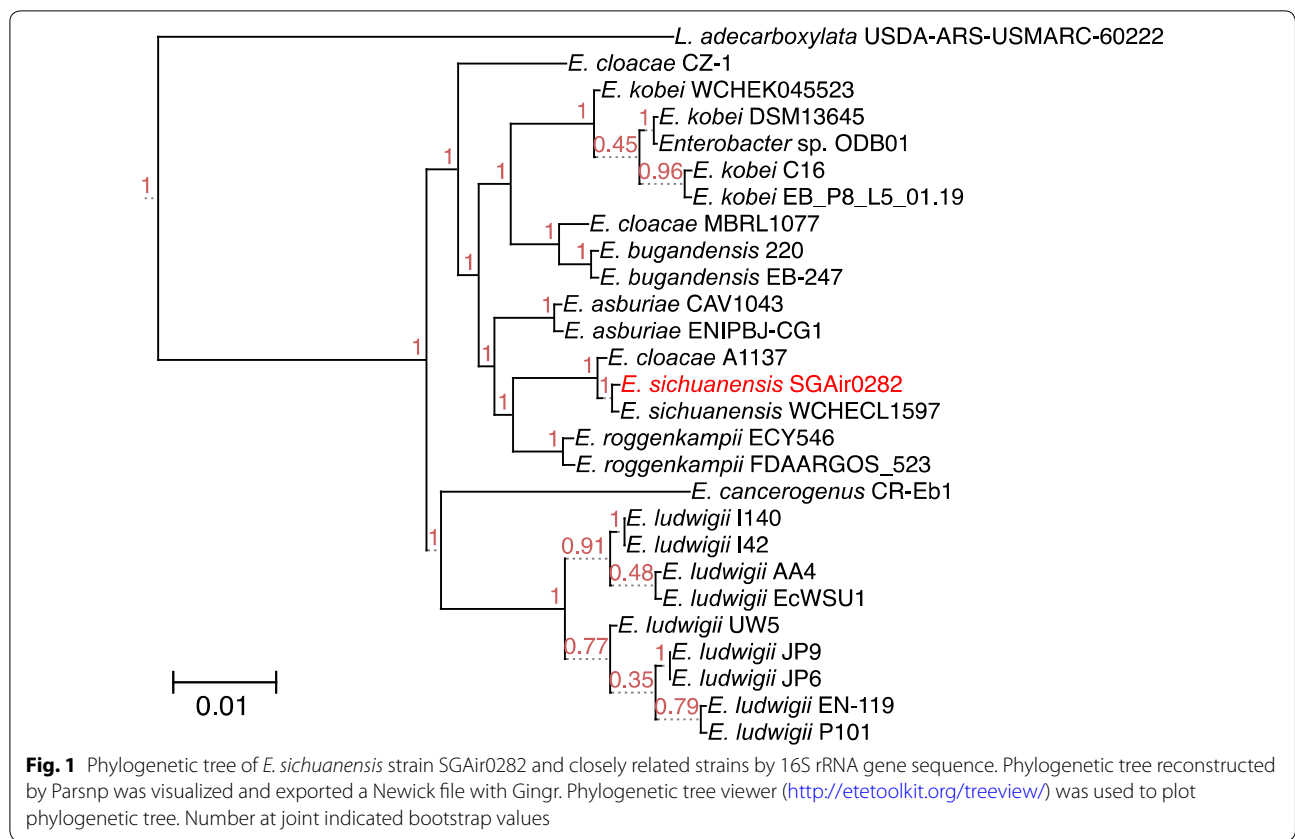


Table 1 Genetic distance by MASH and allele and sequence type number by MLST Sequence type of the house-keeping genes were defined through PubMLST.org using database for *E. cloacae*. One strain from each species were listed

| | Distance from ref. ^d | p-value | <i>dnaA</i> | <i>fusA</i> | <i>gyrB</i> | <i>leuS</i> | <i>pyrG</i> | <i>rplB</i> | <i>rpoB</i> | ST |
|-----------------------------------|---------------------------------|---------|-------------|-------------|-------------------|------------------|-------------------|-------------|-------------------|-------------------|
| SGAir0282 | – | – | 153 | 98 | 170 | 440 ^a | 344 ^a | 68 | 217 | 1330 ^a |
| <i>E. sichuanensis</i> WCHECL1597 | 0.014 | 0 | 153 | 98 | n.d. ^c | 159 | n.d. ^c | 68 | n.d. ^c | 472 ^b |
| <i>E. cloacae</i> A1137 | 0.019 | 0 | 241 | 98 | 170 | 287 | 237 | 68 | 109 | 738 |
| <i>E. asburiae</i> ENIPBJ-CG1 | 0.083 | 0 | 347 | 15 | 39 | 134 | 47 | 10 | 113 | n.d. ^c |
| <i>E. ludwigii</i> I140 | 0.103 | 0 | 13 | 2 | n.d. ^c | 286 | 51 | 2 | 13 | n.d. ^c |
| <i>E. kobei</i> WCHEK045523 | 0.096 | 0 | 43 | 3 | 66 | 314 | 3 | 16 | 168 | 910 |
| <i>E. ludwigii</i> AA4 | 0.104 | 0 | 15 | 158 | 365 | 416 | 55 | 105 | 222 | 1281 |
| <i>E. bugandensis</i> 220 | 0.091 | 0 | 322 | 110 | 200 | 30 | 120 | 76 | 28 | 1140 |
| <i>E. roggkampii</i> ECY546 | 0.077 | 0 | 120 | 25 | 127 | 90 | 49 | 12 | 20 | 466 |

^a New allele number and sequence type were given by curation

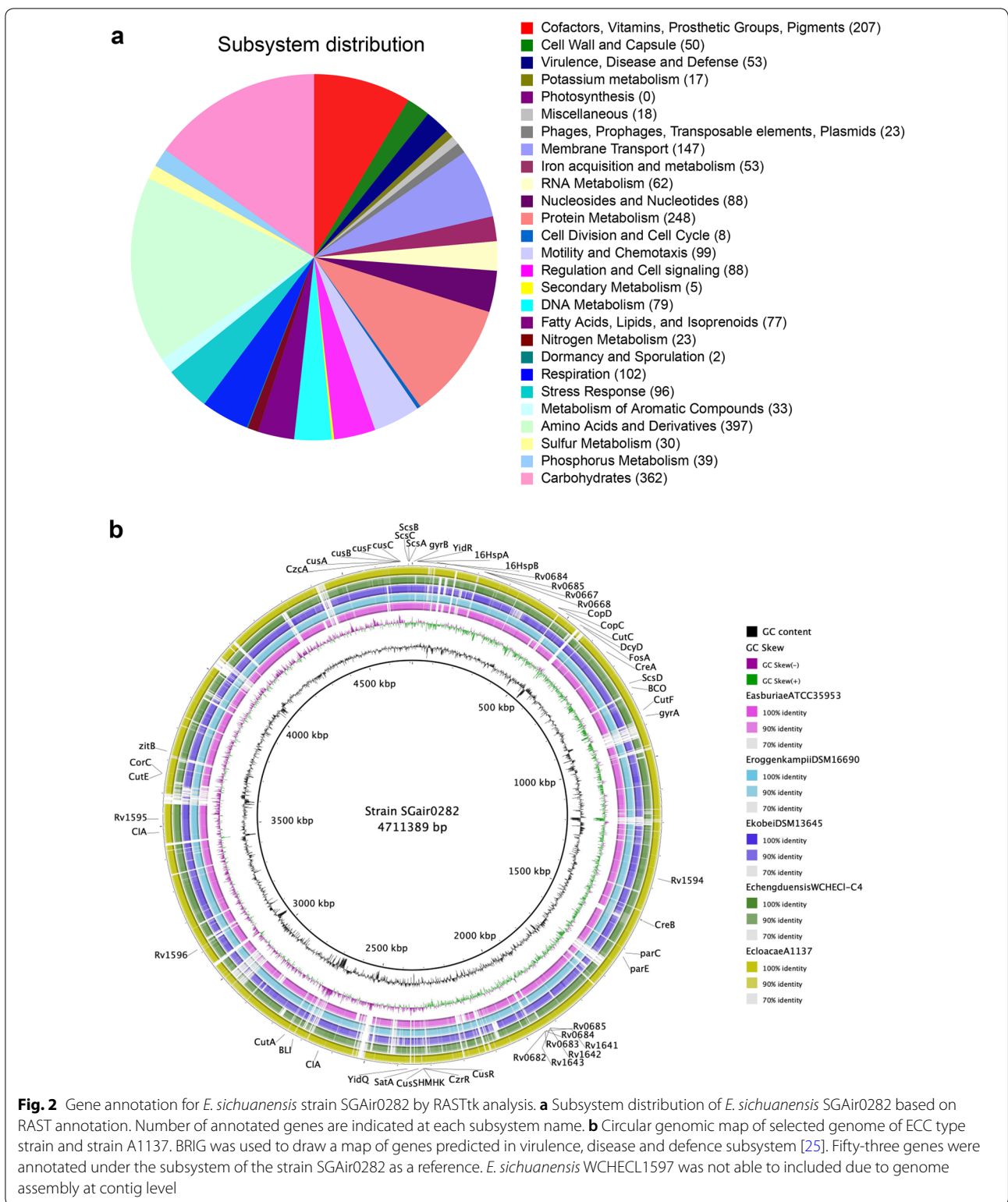
^b The profile was suggested for nearest match with combination of allele by PubMLST.org

^c n.d.: not defined. Database did not have exact allele and ST matches

^d SGAir0282 was used as reference and each strain was used as a query sequence on MASH software

(2013) and pneumonia patient’s blood (2016), respectively, in the hospitals in China. Those two patients were individually administered with antibiotics for the infections by other bacteria, and the two strains managed to acquire multidrug resistance through sets of antibiotic

treatments. Multidrug resistant bacteria that have been released through facilities such as wastewater handling facilities of hospitals [10], are potentially aerosolized and transported through air or humans that host the strain of non-virulent bacteria as part of their microbiome.



The high similarity of SGAir0282 to these two strains clearly suggests that the strain SGAir0282 is a potential multidrug-resistant pathogen. This hypothesis is

also supported by fact that SGAir0282 possesses metal-dependent hydrolases of the beta-lactamase superfamily I gene which confers resistance to beta-lactam

antibiotics. Although the carbapenemase gene was not detected in the SGAir0282 genome, the strain could acquire a gene encoding carbapenemase by mutation and/or recombination, in the presence of carbapenem.

Our analysis emphasized that the importance of using whole genome data for the taxonomy identification. The most commonly used method for the taxonomy identification with a single gene showed low resolution due to the limited amount of data (Additional file 2: Table S1). With the whole genome sequences, our analyses revealed that the strain SGAir0282, *E. cloacae* A1137 and *E. sichuanensis* WCHECL1597 are closely related with a sequence similarity that is high enough to support the claim that these are the same species (Fig. 1, Additional file 3: Table S2 and Additional file 4: Table S3).

The genome assembly for *E. sichuanensis* WCHECL1597 is highly fragmented and consists of 204 contigs as of December 2019. On the other hand, the genome assembly of *E. sichuanensis* SGAir0282 consists of a single, circular chromosome with high coverage sequence quality, and we therefore suggest that this genome should be used as the reference sequence for *E. sichuanensis* species.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s13099-020-00350-z>.

Additional file 1. The variant call was performed by kSNP3. Dataset used here was defined in the method section. Strain SGAir0282 was used as reference genome.

Additional file 2: Table S1. Similarity search result using 16S rRNA gene of strain SGAir0282. NCBI's BLASTn search was used for submitting query sequence, 16S rRNA gene. The alignment search was performed against nucleotide collection. Other parameters were kept at default.

Additional file 3: Table S2. Pairwise distance from strain SGAir0282 estimated by MASH. Pairwise distance of Reference sequence (Ref seq) against Query sequence (Query seq) was estimated. The same dataset was used as Fig. 1. Default setting was used for this analysis.

Additional file 4: Table S3. Pairwise distance matrix showing distance between genomes on the Neighbour-joining tree. k-mer based kSNP3 analysis was performed on our dataset. k-mer was set at 19 nucleotides. Genome sequence of SGAir0282 was used as reference.

Authors' contributions

AU, HLK and SCS designed the research project, analysed data, and wrote the manuscript. RWP and VKV assembled genome sequencing data. CC, MEC, JNHH contributed in microbial isolation. KKK, AP, DP, DID contributed in extracting DNA and sequencing. AW, ABG, ACMJ contributed in reviewing the manuscript. CEH and BNVP contributed to sequencing data analysis. All authors read and approved the final manuscript.

Funding

The work was supported by Singapore Ministry of Education Academic Research Fund Tier 3 Grant (MOE2013-T3-1-013).

Availability of data and materials

The complete genome sequence of *Enterobacter sichuanensis* SGAir0282 has been deposited in DDBJ/EMBL/GenBank under the accession numbers CP027986.

Ethics approval and consent to participants

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, Singapore, Singapore. ² Asian School of the Environment, Nanyang Technological University, Singapore, Singapore. ³ Departamento de Genética, Instituto de Biologia, Universidade Federal Do Rio de Janeiro, Rio de Janeiro, Brazil. ⁴ School of Biological Sciences, Nanyang Technological University, Singapore, Singapore.

Received: 28 August 2019 Accepted: 15 February 2020

Published online: 27 February 2020

References

- de Andrade FM, de Assis PT, Souza TP, Guimarães PHS, Martins AD, Schwan RF, et al. Beneficial effects of inoculation of growth-promoting bacteria in strawberry. *Microbiol Res.* 2019;223–225:120–8.
- Abraham J, Silambarasan S. Plant growth promoting bacteria *Enterobacter asburiae* JAS5 and *Enterobacter cloacae* JAS7 in mineralization of endosulfan. *Appl Biochem Biotechnol.* 2015;175:3336–48.
- Wu W, Feng Y, Zong Z. *Enterobacter sichuanensis* sp. nov., recovered from human urine. *Int J Syst Evol Microbiol.* 2018;68:3922–7. <https://doi.org/10.1099/ijsem.0.003089>.
- Zhu B, Wang S, Li O, Hussain A, Hussain A, Shen J, et al. High-quality genome sequence of human pathogen *Enterobacter asburiae* type strain 1497–78T. *J Glob Antimicrob Resist.* 2017;8:104–5.
- Chavda KD, Chen L, Fouts DE, Sutton G, Brinkac L, Jenkins SG, et al. Comprehensive genome analysis of carbapenemase-producing *Enterobacter* spp.: new insights into phylogeny, population structure, and resistance mechanisms. *MBio.* 2016;7:1–16.
- Hoffmann H, Schmoldt S, Trülzsch K, Stumpf A, Bengsch S, Blankenstein T, et al. Nosocomial urosepsis caused by *Enterobacter kobei* with aberrant phenotype. *Diagn Microbiol Infect Dis.* 2005;53:143–7.
- Davin-Regli A, Pagès JM. *Enterobacter aerogenes* and *Enterobacter cloacae*: versatile bacterial pathogens confronting antibiotic treatment. *Front Microbiol.* 2015;6:1–10.
- Raymond F, Ouameur AA, Déraspe M, Iqbal N, Gingras H, Dridi B, et al. The initial state of the human gut microbiome determines its reshaping by antibiotics. *ISME J.* 2016;10:707–20.
- Daoud Z, Farah J, Sokhn ES, El Kfoury K, Dahdouh E, Masri K, et al. Multidrug-resistant Enterobacteriaceae in Lebanese Hospital Wastewater: implication in the one health concept. *Microb Drug Resist [Internet].* 2018;24:166–74. <https://www.liebertpub.com/doi/10.1089/mdr.2017.0090>
- Gomi R, Matsuda T, Yamamoto M, Chou PH, Tanaka M, Ichiyama S, et al. Characteristics of carbapenemase-producing Enterobacteriaceae in wastewater revealed by genomic analysis. *Antimicrob Agents Chemother.* 2018;62:1–11.
- Gusareva ES, Acerbi E, Lau KJX, Luhung I, Premkrishnan BNV, Kolundžija S, et al. Microbial communities in the tropical air ecosystem follow a precise diel cycle. *Proc Natl Acad Sci.* 2019;116:23299–308. <https://doi.org/10.1073/pnas.1908493116>.
- Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, et al. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods.* 2013;10:563–9.
- Walker BJ, Abeel T, Shea T, Priest M, Boueilliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One.* 2014. <https://doi.org/10.1371/journal.pone.0112963>.

14. Hunt M, De SN, Otto TD, Parkhill J, Keane JA, Harris SR. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol.* 2015;16:1–10. <https://doi.org/10.1186/s13059-015-0849-0>.
15. Tatusova T, Dicuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, et al. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* 2016;44:6614–24.
16. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, et al. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep.* 2015;5:8365. <https://doi.org/10.1038/srep08365>.
17. Shen L, Liu Y, Xu B, Wang N, Zhao H, Liu X, et al. Comparative genomic analysis reveals the environmental impacts on two *Arcticibacter* strains including sixteen Sphingobacteriaceae species. *Sci Rep* [Internet]. 2017;7:2055. <https://www.nature.com/articles/s41598-017-02191-4>
18. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol.* 2014;15:1–15.
19. Huerta-Cepas J, Serra F, Bork P. ETE 3: Reconstruction, analysis, and visualization of phylogenomic data. *Mol Biol Evol.* 2016;33:1635–8.
20. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, et al. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol.* 2016;17:1–14. <https://doi.org/10.1186/s13059-016-0997-x>.
21. Gardner SN, Slezak T, Hall BG. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. *Bioinformatics.* 2015;31:2877–8.
22. Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res.* 2018;3:1–20. <https://doi.org/10.12688/wellcomeopenres.14826.1>.
23. Zhan Z, Hu L, Jiang X, Zeng L, Feng J, Wu W, et al. Plasmid and chromosomal integration of four novel bla_{IMP}-carrying transposons from *Pseudomonas aeruginosa*, *Klebsiella pneumoniae* and an *Enterobacter* sp. *J Antimicrob Chemother* [Internet]. 2018;73:3005–15. <https://academic.oup.com/jac/article/73/11/3005/5059526>
24. Khalifa AYZ, Alsyeed A-M, Almalki MA, Saleh FA. Characterization of the plant growth promoting bacterium, *Enterobacter cloacae* MSR1, isolated from roots of non-nodulating *Medicago sativa*. *Saudi J Biol Sci.* 2016;23:79–86. <https://doi.org/10.1016/j.sjbs.2015.06.008>
25. Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. BLAST ring image generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics.* 2011. <https://doi.org/10.1186/1471-2164-12-402>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

