Complete and Draft Genome Sequences of Amino Acid-Producing Corynebacterium glutamicum Strains ATCC 21799 and ATCC 31831 and Their Genomic Islands

Kawaguchi, Hideo
Sazuka, Takashi
Kondo, Akihiko

(Citation)
Microbiology Resource Announcements,9(32):e00430-20

(Issue Date)
2020-08

(Resource Type)
journal article

(Version)
Version of Record

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete and Draft Genome Sequences of Amino Acid-Producing *Corynebacterium glutamicum* Strains ATCC 21799 and ATCC 31831 and Their Genomic Islands

Hideo Kawaguchi,[a] Takashi Sazuka,[b] Akihiko Kondo[a,c]

[a]Graduate School of Science, Technology, and Innovation, Kobe University, Kobe, Japan
[b]Bioscience and Biotechnology Center, Nagoya University, Nagoya, Japan
[c]Engineering Biology Research Center, Kobe University, Kobe, Japan

ABSTRACT We determined the complete and draft genome sequences of two strains of *Corynebacterium glutamicum* and revealed their genomic islands (GEIs). The two strains, ATCC 21799 and ATCC 31831, were found to have 3,079 and 3,109 coding sequences, respectively, with 13 GEIs each not present in the reference strain, ATCC 13032.

Corynebacterium glutamicum is a Gram-positive soil microorganism (1) widely used for industrial amino acid production (2, 3). Although the genomes of several *C. glutamicum* strains have been elucidated (4–7), their genomic islands (GEIs) have not been comprehensively reported. Therefore, we determined the complete and draft genome sequences of two *C. glutamicum* strains, ATCC 21799 and ATCC 31831, and predicted GEIs not present in the reference strain, ATCC 13032.

Two *C. glutamicum* strains purchased from ATCC (Manassas, VA) were cultivated aerobically in brain heart infusion broth (Merck, Darmstadt, Germany) at 30°C. After 16 h of cultivation, the genomic DNA was extracted from these strains using a Nucleobond AXG system (TaKaRa Bio, Shiga, Japan). DNA sequencing was performed using the PacBio RS II system (Pacific Biosciences, Menlo Park, CA) with DNA sequencing reagent 4.0 v2. A single SMRTbell library was prepared according to the manufacturer's instructions and was sheared to 15 kb using BluePippin size selection system v10 (Sage Science, Beverly, MA). The genomes were assembled with Hierarchical Genome Assembly Process v2.3.0 (8), and filtering was based on a threshold of 0.80 for minimum polymerase read quality. GEIs were predicted using GIPSy v1.1.2 software (9), and maps of the circular genomes of *C. glutamicum* showing GEIs were generated using BLAST Ring Image Generator (BRIG) v0.95 software (10) for the reference strain, ATCC 13032 (GenBank accession number BA000036.3). In all analyses, default parameters were used except when otherwise noted.

For strain ATCC 21799, 107,747 filtered reads with an N_{50} value of 14,464 bp were assembled into one contig, yielding a 3,332,273-bp complete sequence. However, for strain ATCC 31831, 144,465 filtered reads with an N_{50} value of 18,854 bp were assembled into two contigs, namely, contigs 1 and 2 (3,302,680 bp and 29,004 bp, respectively), yielding a 3,311,684-bp draft genome sequence. Coverage depths were 282× and 301× with average G+C contents of 54.3% and 54.0% for strains ATCC 21799 and ATCC 31831, respectively. Genome sequences were automatically annotated using the genome annotation pipeline DFAST (11), yielding 3,079 and 3,109 coding sequences (CDSs), 65 and 62 tRNAs, and 18 and 21 rRNAs, respectively, for strains ATCC 21799 and ATCC 31831.

The genomes of strains ATCC 21799 and ATCC 31831 had 13 GEIs each. In strain ATCC 21799, pathogenicity island 9 (PAI 9) was the largest GEI (65 kb) and had

Citation Kawaguchi H, Sazuka T, Kondo A. 2020. Complete and draft genome sequences of amino acid-producing *Corynebacterium glutamicum* strains ATCC 21799 and ATCC 31831 and their genomic islands. Microbiol Resour Announc 9:e00430-20. https://doi.org/10.1128/MRA.00430-20.

Editor Catherine Putonti, Loyola University Chicago

Address correspondence to Akihiko Kondo, akondo@kobe-u.ac.jp.

Received 13 May 2020
Accepted 3 July 2020
Published 6 August 2020

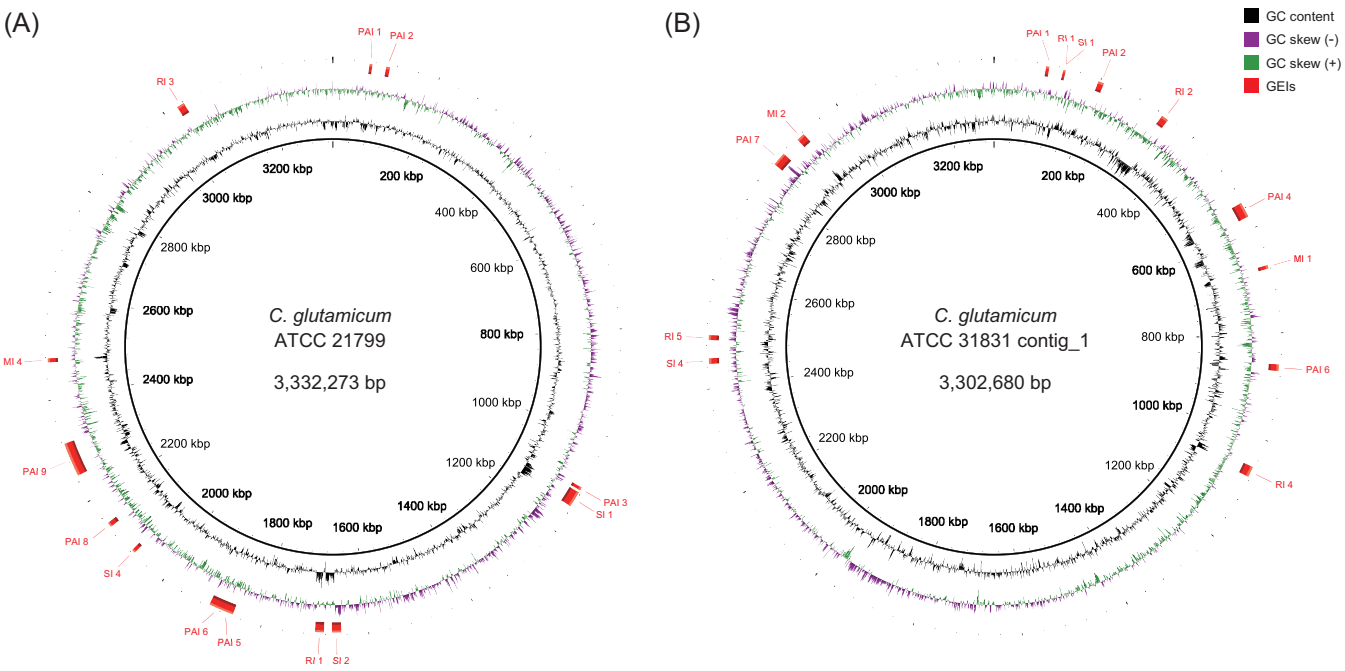

FIG 1 Maps of the circular genomes of *Corynebacterium glutamicum* ATCC 21799 (A) and ATCC 31831 (B) showing genomic islands (GEIs) predicted for the *C. glutamicum* type strain ATCC 13032 and regions of genome plasticity. The figure was generated with *C. glutamicum* ATCC 13032 as the reference strain using the software BLAST Ring Image Generator (BRIG) (10). GEIs are classified as putative pathogenicity islands (PAI), resistance islands (RI), metabolic islands (MI), and symbiotic islands (SI).

transposases at both ends (Fig. 1A). This region consisted of 55 coding DNA sequences (CDSs) (KaCgl20770 to KaCgl21310), and 80% of them showed a deviation in codon usage from the standard value of 0.95. In strain ATCC 31831, metabolic island 2 (MI 2) (KbCgl27180 to KbCgl27280) is involved in an exclusive gene cluster responsible for L-arabinose utilization (12) with distinct codon usage patterns (Fig. 1B). In conclusion, the two GEIs of *C. glutamicum* are unique and structurally discrete sequences that most likely arose independently during evolution by horizontal gene transfer.

Data availability. The complete genome sequence of strain ATCC 21799 and the draft genome sequences of strain ATCC 31831 have been deposited under DDBJ/ENA/GenBank accession number AP022856.1 and numbers BLRJ01000001.1 and BLRJ01000002.1, respectively. The raw reads were deposited under SRA accession numbers DRR232384 and DRR232383, respectively.

ACKNOWLEDGMENTS

We gratefully acknowledge the technical assistance provided by Shoko Miyazaki and Kumiko Yoshihara. We are also grateful to Hiromichi Araki (Kyoto University, Japan) and Christopher John Vavricka, Jr. (Kobe University, Japan), for their valuable advice regarding the whole-genome analysis.

This work was supported in part by a grant from the JST-Mirai Program, Japan (JPMJMI17EG), and by a KAKENHI Grant-in-Aid for Scientific Research (B) to H.K. (JP19KT0009) from the Japan Society for the Promotion of Science (JSPS), Japan.

REFERENCES

1. Wendisch VF, Bott M, Kalinowski J, Oldiges M, Wiechert W. 2006. Emerging *Corynebacterium glutamicum* systems biology. J Biotechnol 124:74–92. https://doi.org/10.1016/j.jbiotec.2005.12.002.
2. Nakayama K, Kitada S, Kinoshita S. 1961. Studies on lysine fermentation. I. The control mechanism on lysine accumulation by homoserine and threonine. J Gen Appl Microbiol 7:145–154. https://doi.org/10.2323/jgam.7.145.
3. Wittmann C, Kiefer P, Zelder O. 2004. Metabolic fluxes in *Corynebacterium glutamicum* during lysine production with sucrose as carbon source. Appl Environ Microbiol 70:7277–7287. https://doi.org/10.1128/AEM.70.12.7277-7287.2004.
4. Yang J, Yang S. 2017. Comparative analysis of *Corynebacterium glutamicum* genomes: a new perspective for the industrial production of amino acids. BMC Genomics 18:940. https://doi.org/10.1186/s12864-016-3255-4.

5. Nishio Y, Koseki C, Tonouchi N, Matsui K, Sugimoto S, Usuda Y. 2017. Analysis of strain-specific genes in glutamic acid-producing *Corynebacterium glutamicum* ssp. *lactofermentum* AJ 1511. J Gen Appl Microbiol 63:157–164. https://doi.org/10.2323/jgam.2016.09.004.
6. Ikeda M, Nakagawa S. 2003. The *Corynebacterium glutamicum* genome: features and impacts on biotechnological processes. Appl Microbiol Biotechnol 62:99–109. https://doi.org/10.1007/s00253-003-1328-1.
7. Kalinowski J, Bathe B, Bartels D, Bischoff N, Bott M, Burkovski A, Dusch N, Eggeling L, Eikmanns BJ, Gaigalat L, Goesmann A, Hartmann M, Huthmacher K, Krämer R, Linke B, McHardy AC, Meyer F, Möckel B, Pfefferle W, Pühler A, Rey DA, Rückert C, Rupp O, Sahm H, Wendisch VF, Wiegräbe I, Tauch A. 2003. The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. J Biotechnol 104:5–25. https://doi.org/10.1016/s0168-1656(03)00154-8.
8. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563–569. https://doi.org/10.1038/nmeth.2474.
9. Soares SC, Geyik H, Ramos RTJ, de Sa PHCG, Barbosa EGV, Baumbach J, Figueiredo HCP, Miyoshi A, Tauch A, Silva A, Azevedo V. 2016. GIPSy: genomic island prediction software. J Biotechnol 232:2–11. https://doi.org/10.1016/j.jbiotec.2015.09.008.
10. Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. 2011. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics 12:402. https://doi.org/10.1186/1471-2164-12-402.
11. Tanizawa Y, Fujisawa T, Nakamura Y. 2018. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. Bioinformatics 34:1037–1039. https://doi.org/10.1093/bioinformatics/btx713.
12. Kawaguchi H, Sasaki M, Vertes AA, Inui M, Yukawa H. 2009. Identification and functional analysis of the gene cluster for L-arabinose utilization in *Corynebacterium glutamicum*. Appl Environ Microbiol 75:3419–3429. https://doi.org/10.1128/AEM.02912-08.