

Draft Genome Sequence of *Erwinia mallotivora* BT-MARDI, Causative Agent of Papaya Dieback Disease

R. Ahmad Redzuan,^a N. Abu Bakar,^a L. Rozano,^a R. Badrun,^a N. Mat Amin,^a M. F. Mohd Raih^b

Biotechnology Research Centre, Malaysian Agricultural Research and Development Institute (MARDI) Headquarters, Persiaran MARDI-UPM, Serdang, Selangor, Malaysia^a; School of Biosciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia^b

***Erwinia mallotivora* was isolated from papaya trees infected with dieback disease, which were planted at the Malaysian Agricultural Research and Development Institute (MARDI), Malaysia. Here, we report a draft genome sequence of *E. mallotivora* BT-MARDI, which offers an important source of information for understanding pathogen and host interaction during papaya dieback development.**

Received 7 April 2014 Accepted 14 April 2014 Published 8 May 2014

Citation Redzuan RA, Abu Bakar N, Rozano L, Badrun R, Mat Amin N, Mohd Raih MF. 2014. Draft genome sequence of *Erwinia mallotivora* BT-MARDI, causative agent of papaya dieback disease. *Genome Announc.* 2(3):e00375-14. doi:10.1128/genomeA.00375-14.

Copyright © 2014 Redzuan et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to R. Ahmad Redzuan, rohaiza@mardi.gov.my.

Erwinia mallotivora is a Gram-negative bacterium classified as a member of the *Enterobacteriaceae* family (1). *E. mallotivora* was first reported as the causal agent of bacterial leaf spot in *Mallotus japonicus* in Japan (2). In recent studies that included phenotypic observations, biochemical analysis, and genotypic information, *E. mallotivora* was identified as a close relative of *E. papayae* and the causal agent of the papaya dieback disease in Malaysia (3, 4). Thus, it is of great interest to uncover the genome sequence of the isolate obtained from papaya trees infected with dieback. Here, we report the draft genome sequence of *E. mallotivora* BT-MARDI isolated from papaya infected with dieback disease at the Malaysian Agricultural Research and Development Institute (MARDI), Malaysia.

E. mallotivora BT-MARDI was identified using Gram stain and biochemical tests. For biochemical tests, an API 20E kit (Biomerieux SA Co., France) was used to confirm the identity of this bacterium. The identity of this particular bacterium was also confirmed using 16S rRNA gene cloning. The genome was sequenced using Illumina and Solexa HiSeq 2000 technology (Illumina, San Diego, CA). The quality of the sequencing output was determined using FastQC. The sequencing data were generated from a paired-end, 101-bp run. The sequencing produced an output of 53,801,258 reads with 550× coverage. FastQC revealed the average G+C content of this bacterium to be about 52.4%. The G+C contents of other reference genomes are 53.6%, 53.4%, 55%, and 53.4% for *E. amylovora*, *E. pyrofiliae*, *E. bilingiae*, and *E. tasmaniensis*, respectively. After trimming, 28,912,327 reads were produced, with 54% trimming efficiency. The reads were assembled (*de novo* assembly) and scaffolded using MSR-CA assembler (5). A total of 80 contigs were produced, which covered a total genome size of 4,824,879 bp.

Fifty scaffolds were generated and remapped against the raw sequence (101 bp) to validate the assembly results. The total numbers of gaps predicted was 50. The annotation was performed using Glimmer3 and aligned to the UNIPROT database using Blastx. The annotation results were validated using Rapid Annotations using Subsystems Technology (RAST) (6). The scaffolds were annotated to have 498 subsystems, 4,855 coding sequences, 4,136 protein-encoding genes, and 81 RNAs. The draft

genome sequence of *E. mallotivora* BT-MARDI showed differences from the reference sequences of other *Erwinia* spp. Based on these annotated results, several putative virulence genes have been identified in the genome of *E. mallotivora* BT-MARDI, which include genes involved in secretion systems, sucrose metabolism, capsular polysaccharides, clustered regularly interspaced short palindromic repeats (CRISPRs), iron acquisition and metabolism, motility and chemotaxis, and cell regulation and signaling. These in turn will offer important information to increase our understanding of the host-pathogen interaction during dieback infection in papaya plants.

Nucleotide sequence accession numbers. The genome sequence of *E. mallotivora* BT-MARDI has been deposited in DDBJ/EMBL/GenBank under accession number **JFHN00000000**. The version described in this paper is version JFHN0100000000.

ACKNOWLEDGMENT

This study was supported by a grant from the SciencesFund, Ministry of Science Technology and Innovation (MOSTI) of Malaysia (2/3/08-SF0407).

REFERENCES

1. Agrios GN. 2005. Plant pathology, 5th ed. p 398–401. Elsevier Academic, London, United Kingdom.
2. Goto M. 1976. *Erwinia mallotivora* sp. nov., the causal organism of bacterial leaf spot of *Mallotus japonicus* Muell. *Int. J. Syst. Bacteriol.* 26:467–473. <http://dx.doi.org/10.1099/00207713-26-4-467>.
3. Maktar NH, Kamis S, Mohd Yusof FZ, Hussain NH. 2008. *Erwinia papayae* causing papaya dieback in Malaysia. *J. Plant Pathol.* 57:774. <http://dx.doi.org/10.1111/j.1365-3059.2008.01876.x>.
4. Amin NM, Bunawan H, Redzuan RA, Jaganath IB. 2010. *Erwinia mallotivora* sp., a new pathogen of Papaya (*Carica papaya*) in Peninsular Malaysia. *Int. J. Mol. Sci.* 12:39–45. <http://dx.doi.org/10.3390/ijms12010039>.
5. Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA. 2013. The MaSuRCA genome assembler. *Bioinformatics* 29:2669–2677. <http://dx.doi.org/10.1093/bioinformatics/btt476>.
6. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. <http://dx.doi.org/10.1186/1471-2164-9-75>.