Genome Resource of a Hypervirulent Strain LN4 of *Xanthomonas oryzae* pv. *oryzae* Causing Bacterial Blight of Rice

Zhengyin Xu, Sai Wang, Liang Liu, Yangyang Yang, Bo Zhu, Lifang Zou, and Gongyou Chen[†]

School of Agriculture and Biology/State Key Laboratory of Microbial Metabolism, Shanghai Jiao Tong University, Shanghai 200240, China

Abstract

Xanthomonas oryzae pv. *oryzae* is the causative agent of bacterial blight of rice and causes severe harvest loss and challenges to a stable food supply globally. In this study, a hypervirulent strain, LN4, compatible in rice varieties carrying *Xa3*, *Xa4*, *xa13*, and *xa25* resistance genes, was used to generate DNA for nanopore sequencing. After assembly, the genome comprises a single chromosome of 5,012,583 bp, consisting of a total of 6,700 predicted coding sequences. Seventeen transcription activator-like effectors (TALEs) were encoded in the genome, of which two (*Tal7* and *Tal6c*) were major TALEs. The approach and genome data provide information for the discovery of new virulence effectors and understanding of the virulence mechanism of TALEs in rice.

Funding

This research was supported by the National Natural Science Foundation of China (31830072), the National Key Research and Development Program of China (2016YFD0100601), and the National Transgenic Major Program (2016ZX08001-002).

Keywords

genome, *Xanthomonas oryzae* pv. *oryzae*, TALE

Genome Announcement

Bacterial blight (BB) of rice is a systemic disease caused by gram-negative bacterium *Xanthomonas oryzae* pv. *oryzae* (*Xoo*). The disease is one of the most destructive diseases in rice and is widespread in rice-growing areas in many countries of Asia, Latin America, and west Africa (Niño-Liu et al. 2006). BB damage was reported to range from 10 to 30%, and sometimes more than 50% or even no yield under environmental conditions favorable to *Xoo* infection (Schornack et al. 2013). BB is also one of the most important model systems in plant pathology and plant immunology (Mansfield et al. 2012; Niño-Liu et al. 2006). *Xoo* genomes encode significant numbers of transcription activator-like effector (TALE) proteins that function as transcriptional activators inside the plant cell to reprogram the host transcriptome (Boch and Bonas 2010; Boch et al. 2014; Hutin et al. 2015; Xu et al. 2017). Several TALEs are known as major TALEs that play an essential role during bacterial infection, including *AvrXa7*, *PthXo1*, *PthXo2*/*PthXo2*-like series, *PthXo3*, *TalB*, *TalC*, and *TalF* (Streubel et al. 2013; Tran et al. 2018; Xu et al. 2019; Yang and White 2004; Yang et al. 2006; Yu et al. 2011; Zhou et al. 2015). Here, we describe the complete genome sequence of a Chinese *Xoo* strain, LN4, which carries two major TALEs and is compatible in rice varieties harboring *Xa3*, *Xa4*, *xa13*, and *xa25* resistance genes (data not shown).

Xoo strain LN4 was isolated from rice in Liaoning province, China, in 2014, and the pathogenicity test showed its high virulence on rice compared with other strains (Xu et al. 2019). DNA was extracted using the AxyPrep Bacterial Genomic DNA Miniprep Kit (Axygen, U.S.A.). Extracted DNA sample was prepared for long-read sequencing with the Oxford Nanopore Technologies (ONT) Ligation library preparation kit according to the manufacturer's standard protocol. The library was sequenced with the ONT PromethION 48 sequencer. The sequencer was controlled with MinKNOW version 2.2.12 software. A total of

[†]Corresponding author: G. Chen; gyouchen@sjtu.edu.cn

The author(s) declare no conflict of interest.

Accepted for publication 30 April 2020.

Fig. 1. Genome features and transcription activator-like effector (TALE) gene cluster of *Xanthomonas oryzae* pv. *oryzae* strain LN4. **A**, Circos diagram and features of closed, circular genomes of LN4 strain. The outermost ring of the Circos diagram represents the clusters of orthologous groups (COG) annotation of the genome. The middle and inner rings represent annotated noncoding RNA and GC content, respectively. The inner track displays GC skew. **B**, The TALE gene cluster of LN4 strain.

133,039 reads with 15,319-bp average read length was generated using Guppy version 2.3.5. A de novo genome assembly was performed with the software Flye version 2.6 (Kolmogorov et al. 2019). The genome was assembled into a single contig with 407× coverage depth, and the size was 5,012,583 bp, which is similar in size to other *Xoo* strain genomes (Booher et al. 2015; Lee et al. 2005; Mücke et al. 2019; Salzberg et al. 2008; Tran et al. 2018). The N50 and L50 are 5,012,583 and 1, respectively. The GC content is 63.59%. Genome components analysis with Prodigal version 2.6.3 (Hyatt et al. 2010) predicted a total of 6,700 coding sequences. Fifty tRNA genes and six rRNA genes were predicted using tRNAscan-SE version 1.3.1 (Lowe and Eddy 1997) and RNAmmer version 1.2 (Lagesen et al. 2007), respectively. Of the 6,700 predicted genes, 4,523 could be categorized into a clusters of orthologous groups (COG) category and 3,015 genes could be categorized to a KEGG category by DIAMOND (Buchfink et al. 2015). There were 309 secreted proteins predicted by signal peptide prediction tool SignalP version 4.1 (Petersen et al. 2011). The summary of the genome assembly is shown in Figure 1A.

Sequence analysis and Southern blot data (Xu et al. 2019) indicated that LN4 genome contains 17 TALEs consisting of 16 typical TALEs and one iTALE/truncTALE, TalAI (Fig. 1B, Table 1). Ten, identical to those of other *Xoo* strains, are assigned to classes TalAE, TalAO, TalAL, TalAB, TalAD, TalAF, TalAH, TalAR, TalAC, and TalBK (Fig. 1B, Table 1). The remaining six typical TALEs are similar to known TALEs from other *Xoo* strains and belong to classes TalAQ, TalAP, TalAN, TalAA, TalAG, and TalHQ (Grau et al. 2016; Lee et al. 2005; Salzberg et al. 2008). Two major TALEs (Tal7 and Tal6c) have been found in LN4 strain. Tal7 (Tal7_{LN4}) belongs to the PthXo2-like (TalBK) TALEs and activates the expression of

Table 1. Repeat variable di-residue (RVD) sequences of *Xanthomonas oryzae* pv. *oryzae* (Xoo) strain LN4

TALE class ^a	TALE gene in LN4	Xoo strain	RVD sequence alignment
TalAE	<i>tal1a</i>	LN4	NI NN NI HG HG HD NG HD HG HD HD NG
		KACC10331	NI NN NI HG HG HD NG HD HG HD HD NG
TalAO	<i>tal1b</i>	LN4	NI NN N* NG NS NN NN NI NN NI NG HD HD NI NG
		PXO61	NI NN N* NG NS NN NN NI NN NI NG HD HD NI NG
		PXO71	NI NN N* NG NS NN NN NI NN NI NG HD HD NI NG
TalAQ	<i>tal1c</i>	LN4	HD HD NN NN NI NG HD S* HG HD NG N* NG HD HD N* NI NI NN
			HD HI ND HD NG NN HG N*
		PXO99 ^A	HD HD NN NN NG NG HD NS HG HD NG N* HD HD HD N* NN NI
			NN HD HI ND HD HG NN HG N*
TalAP	<i>tal1d</i>	LN4	HD HD HD NG N* NN HD HD N* NI NI NN HD HI ND HD NI – NG NG
		PXO99 ^A	HD HD HD NG N* NN HD HD N* NI NI NN HD HI ND HD NI H NG NG
		KACC10331	HD HD HD NG N* – – – N* NI NI NN HD HI ND HD NI HD NG NG
TalAL	<i>tal2a</i>	LN4	NI NS HD NG NS NN HD N* NN NN NI NN HD NG HD HD HD NG
		KACC10331	NI NS HD NG NS NN HD N* NN NN NI NN HD NG HD HD HD NG
TalAB	<i>tal2b</i>	LN4	NI HG NI NI NI NN HD NS NN NS NN HD NN NI HD NN NI NG HD NG
		KACC10331	NI HG NI NI NI NN HD NS NN NS NN HD NN NI HD NN NI NG HD NG
TalAD	<i>tal2c</i>	LN4	NN HD NS NG HD NN N* NI HD NS HD NN HD NN HD NN NN NN
			NN NN NN NN HD NG
		KACC10331	NN HD NS NG HD NN N* NI HD NS HD NN HD NN HD NN NN NN
			NN NN NN NN HD NG
TalAN	<i>tal3a</i>	LN4	NI N* NI HG NI NI NS HD NN HD NN NG SS HD NI NI NN NI NN NI NG
		PXO99 ^A	NI HG NI HG NI NI NI HD NN HD NS NG SS HD NI NI NN NI NN NI NG
TalAF	<i>tal3b</i>	LN4	NI NN NN NI NI NI HD NS HG NN NN NN NI NI NG HD
		MAFF311018	NI NN NN NI NI NI HD NS HG NN NN NN NI NI NG HD
		PXO71	NI NN NN NI NI NI HD NS HG NN NN NN NI NI NG HD
TalAH	<i>tal4a</i>	LN4	NI N* NI NS NN NG NN NS N* NS NN NS N* HD HG HD HD HD NS
			N* HD HG HD HD HD HD NG
		KACC10331	NI N* NI NS NN NG NN NS N* NS NN NS N* HD HG HD HD HD NS
			N* HD HG HD HD HD HD NG
TalAA	<i>tal4b</i>	LN4	NI HG NI NG HG HD NS NG HD NN NG HG NG HD HG HD HD NS
			NN NG
		KACC10331	NI HG NI NG HG HD NS NG HD NN NG HG NG HD HG HD HD NI
			NN NG
TalAR	<i>tal4c</i>	LN4	NI H* NI NN NN NN NN NN HD NI NS HG HD NI N* NS NI NI HD N*
			NS NI N*
		KACC10331	NI H* NI NN NN NN NN NN HD NI NS HG HD NI N* NS NI NI HD N*
			NS NI N*
TalAI	<i>tal5</i>	LN4	NS HD NG NG NG NG HD HD HD HD NN HD NG HD HD HD
			NN H*
		PXO99 ^A	NS HD NG NG NG NG HD HD HD NN HD NG HD HD HD
			HD N*
TalAG	<i>tal6a</i>	LN4	NI NG NN NG NK NG NI NN NI NN NI NN NS NG NS NN NI N*
			NS NG
		KACC10331	NI NG NN NG NK NG NI NN NI NN NI NN NS NG NS NN NI N*
			NS HG
TalHQ	<i>tal6b</i>	LN4	NI HG NI NI HG HD NN HD HD HD NI NI NN NI HD HD HD HG NN
			NN HD NS NN HD N* NS N*
		PXO61	NI HG NI NI HG HD NN HD HD HD NI NI NN NI HD HD HD HG NN
			NN HD NS NN HD NG NS N*
TalAC (AvrXa7)	<i>tal6c</i>	LN4	NI HG NI NI NS HD NN HD HD HD NS N* N* HD HD NS NS NN NN
			NI NG NN NI N* NS N*
		PXO86	NI HG NI NI NS HD NN HD HD HD NS N* N* HD HD NS NS NN NN
			NI NG NN NI N* NS N*
		MAFF311018	NI HG NI NI NS HD NN HD HD HD NS N* N* HD HD NS NS NN NN
			NI NG NN NI N* NS N*
TalBK (PthXo2-like)	<i>tal7</i>	LN4	NI HG NI NN NI NN HD HD HD HD NS NS HD NI NI HD NG HD HD
			HD NG NG
		KACC10331	NI HG NI NN NI NN HD HD HD HD NS NS HD NI NI HD NG HD HD
			HD NG NG

^a Transcription activator-like effector (TALE) classes based on AnnoTALE (Grau et al. 2016).

OsSWEET13_{Kit} (Xu et al. 2019). Tal6c is identical to AvrXa7 (TalAC), which induces effector-triggered immunity mediated by Xa7 and induces the susceptibility gene *OsSWEET14* in rice (Antony et al. 2010; Yang and White 2004). The genomic resource of this strain provides information for the discovery of new virulence effectors and understanding of virulence mechanism of TALEs in rice.

The complete genome sequence of Chinese *Xoo* strain LN4 has been deposited at GenBank under the accession number CP045452.

Literature Cited

- Antony, G., Zhou, J., Huang, S., Li, T., Liu, B., White, F., and Yang, B. 2010. Rice *xa13* recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene *Os-11N3*. *Plant Cell* 22:3864-3876.
- Boch, J., and Bonas, U. 2010. *Xanthomonas* AvrBs3 family-type III effectors: Discovery and function. *Annu. Rev. Phytopathol.* 48:419-436.
- Boch, J., Bonas, U., and Lahaye, T. 2014. TAL effectors-pathogen strategies and plant resistance engineering. *New Phytol.* 204:823-832.
- Booher, N. J., Carpenter, S. C., Sebra, R. P., Wang, L., Salzberg, S. L., Leach, J. E., and Bogdanove, A. J. 2015. Single molecule real-time sequencing of *Xanthomonas oryzae* genomes reveals a dynamic structure and complex TAL (transcription activator-like) effector gene relationships. *Microb. Genom.* 1. doi: 10.1099/mgen.0.000032.
- Buchfink, B., Xie, C., and Huson, D. H. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12:59-60.
- Grau, J., Reschke, M., Erkes, A., Streubel, J., Morgan, R. D., Wilson, G. G., Koebnik, R., and Boch, J. 2016. AnnoTAL: Bioinformatics tools for identification, annotation, and nomenclature of TALEs from *Xanthomonas* genomic sequences. *Sci. Rep.* 6:21077.
- Hutin, M., Pérez-Quintero, A. L., Lopez, C., and Szurek, B. 2015. MorTAL Kombar: The story of defense against TAL effectors through loss-of-susceptibility. *Front. Plant Sci.* 6:535.
- Hyatt, D., Chen, G. L., LoCascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. 2010. Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinf.* 11:119.
- Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P. A. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* 37:540-546.
- Lagesen, K., Hallin, P., Rødland, E. A., Staerfeldt, H. H., Rognes, T., and Ussery, D. W. 2007. RNAMmer: Consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35:3100-3108.
- Lee, B. M., Park, Y. J., Park, D. S., Kang, H. W., Kim, J. G., Song, E. S., Park, I. C., Yoon, U. H., Hahn, J. H., Koo, B. S., Lee, G. B., Kim, H., Park, H. S., Yoon, K. O., Kim, J. H., Jung, C. H., Koh, N. H., Seo, J. S., and Go, S. J. 2005. The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. *Nucleic Acids Res.* 33:577-586.
- Lowe, T. M., and Eddy, S. R. 1997. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25:955-964.
- Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M., Verdier, V., Beer, S. V., Machado, M. A., Toth, I., Salmond, G., and Foster, G. D. 2012. Top 10 plant pathogenic bacteria in molecular plant pathology. *Mol. Plant Pathol.* 13:614-629.
- Mücke, S., Reschke, M., Erkes, A., Schwietzer, C. A., Becker, S., Streubel, J., Morgan, R. D., Wilson, G. G., Grau, J., and Boch, J. 2019. Transcriptional reprogramming of rice cells by *Xanthomonas oryzae* TALEs. *Front. Plant Sci.* 10: 162.
- Niño-Liu, D. O., Ronald, P. C., and Bogdanove, A. J. 2006. *Xanthomonas oryzae* pathovars: Model pathogens of a model crop. *Mol. Plant Pathol.* 7:303-324.
- Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. 2011. SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods* 8: 785-786.
- Salzberg, S. L., Sommer, D. D., Schatz, M. C., Phillippy, A. M., Rabinowicz, P. D., Tsuge, S., Furutani, A., Ochiai, H., Delcher, A. L., Kelley, D., Madupu, R., Puiu, D., Radune, D., Shumway, M., Trapnell, C., Aparna, G., Jha, G., Pandey, A., Patil, P. B., Ishihara, H., Meyer, D. F., Szurek, B., Verdier, V., Koebnik, R., Dow, J. M., Ryan, R. P., Hirata, H., Tsuyumu, S., Won Lee, S., Seo, Y. S., Sriariyanum, M., Ronald, P. C., Sonti, R. V., Van Sluys, M. A., Leach, J. E., White, F. F., and Bogdanove, A. J. 2008. Genome sequence and rapid evolution of the rice pathogen *Xanthomonas oryzae* pv. *oryzae* PXO99^A. *BMC Genomics* 9:204.
- Schornack, S., Moscou, M. J., Ward, E. R., and Horvath, D. M. 2013. Engineering plant disease resistance based on TAL effectors. *Annu. Rev. Phytopathol.* 51: 383-406.
- Streubel, J., Pesce, C., Hutin, M., Koebnik, R., Boch, J., and Szurek, B. 2013. Five phylogenetically close rice SWEET genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*. *New Phytol.* 200:808-819.
- Tran, T. T., Pérez-Quintero, A. L., Wonn, I., Carpenter, S. C. D., Yu, Y., Wang, L., Leach, J. E., Verdier, V., Cunnac, S., Bogdanove, A. J., Koebnik, R., Hutin, M., and Szurek, B. 2018. Functional analysis of African *Xanthomonas oryzae* pv. *oryzae* TALomes reveals a new susceptibility gene in bacterial leaf blight of rice. *PLoS Pathog.* 14:e1007092.
- Xu, Z., Xu, X., Gong, Q., Li, Z., Li, Y., Wang, S., Yang, Y., Ma, W., Liu, L., Zhu, B., Zou, L., and Chen, G. 2019. Engineering broad-spectrum bacterial blight resistance by simultaneously disrupting variable TALE-binding elements of multiple susceptibility genes in rice. *Mol. Plant* 12:1434-1446.
- Xu, Z., Zou, L., Ma, W., Cai, L., Yang, Y., and Chen, G. 2017. Action modes of transcription activator-like effectors (TALEs) of *Xanthomonas* in plants. *J. Integr. Agric.* 16:2736-2745.
- Yang, B., Sugio, A., and White, F. F. 2006. *Os8N3* is a host disease-susceptibility gene for bacterial blight of rice. *Proc. Natl. Acad. Sci. U.S.A.* 103:10503-10508.
- Yang, B., and White, F. 2004. Diverse members of the AvrBs3/PthA family of type III effectors are major virulence determinants in bacterial blight disease of rice. *Mol. Plant-Microbe Interact.* 17:1192-1200.
- Yu, Y., Streubel, J., Balzergue, S., Champion, A., Boch, J., Koebnik, R., Feng, J., Verdier, V., and Szurek, B. 2011. Colonization of rice leaf blades by an African strain of *Xanthomonas oryzae* pv. *oryzae* depends on a new TAL effector that induces the rice Nodulin-3 *Os11N3* gene. *Mol. Plant-Microbe Interact.* 24: 1102-1113.
- Zhou, J., Peng, Z., Long, J., Sosso, D., Liu, B., Eom, J. S., Huang, S., Liu, S., Vera Cruz, C., Frommer, W. B., White, F. F., and Yang, B. 2015. Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. *Plant J.* 82:632-643.