RESOURCE ANNOUNCEMENT

e-Xtra*

Complete Genome Sequence Resource for *Xanthomonas translucens* pv. *undulosa* MAI5034, a Wheat Pathogen from Uruguay

Felipe Clavijo,[1] Claudia Barrera,[2] Aleksander Benčič,[3,4] Valentina Croce,[1] Jonathan M. Jacobs,[5,6] Adriana J. Bernal,[2] Ralf Koebnik,[7,†] and Veronica Roman-Reyna[5,6]

[1] Laboratorio de Microbiología Molecular, Departamento de Biociencias, Facultad de Química, Universidad de la República, Montevideo, Uruguay
[2] Department of Biological Sciences, Universidad de Los Andes, Bogotá, Colombia
[3] National Institute of Biology, Ljubljana, Slovenia
[4] Jožef Stefan International Postgraduate School, Ljubljana, Slovenia
[5] Department of Plant Pathology, The Ohio State University, Columbus, OH, U.S.A.
[6] Infectious Diseases Institute, The Ohio State University, Columbus, OH, U.S.A.
[7] Plant Health Institute of Montpellier, University of Montpellier, CIRAD, INRAE, Institut Agro, IRD, Montpellier, France

Funding
Support was provided by the Institut de Recherche pour le Développement, France, for the North-South-South Network on Xanthomonads within the International Scientific Coordination Network-South (GDRI-Sud NSSN-X); the European Cooperation in Science and Technology (COST) Programme for support of A. Benčič and R. Koebnik through the EuroXanth COST Action CA16107, which was key in initiating this collaborative genome project; the American Malting Barley Association funding to J. M. Jacobs; the Faculty of Sciences at Universidad de los Andes to A. J. Bernal; and the Slovenian Research Agency to A. Benčič.

Keywords
evolution, genetic diversity, TAL effector, wheat, *Xanthomonas*

Genome Announcement

The bacterial species *Xanthomonas translucens* is responsible for bacterial leaf streak and black chaff of small grains and bacterial wilt of forage grasses (Egli et al. 1975; Jones et al. 1917; Sapkota et al. 2020; Vauterin et al. 1995). In particular, bacterial leaf streak of wheat represents the most limiting bacterial disease in wheat production worldwide (Duveiller et al. 1997). The *X. translucens* pathovar most commonly associated with wheat is pathovar *undulosa* (Smith et al. 1919).

Until recently, genome sequences of 61 strains were available for this species at NCBI GenBank, 16 of which are complete genomes. Only one *X. translucens* genome sequence from a South American strain, UPB787, is currently available, which was isolated from barley in Paraguay in 1990. To enlarge the geographic coverage of the pathogen's genomic resources, we present the first complete genome sequence of a South American strain of *X. translucens*, which reveals a surprising conservation of its repertoire of transcription activator-like effectors (TALEs) across continents.

Strain MAI5034 was isolated in October 2018 from symptomatic wheat leaf tissue obtained in Soriano, Uruguay (Clavijo et al. 2022). Pathogenicity on wheat was confirmed in greenhouse assays. The strain was identified as *X. translucens* pv. *undulosa* by multilocus sequence analysis of the concatenated partial sequences of four housekeeping genes (*dnaK*, *fyuA*, *gyrB*, and *rpoD*) and, through multilocus sequence typing, it was assigned to novel sequence type ST3 (Clavijo et al. 2022).

Strain MAI5034 was grown at 28°C on peptone-sucrose-agar medium (0.5% peptone, 2% sucrose, and 1.5% agar) for 24 h. Bacteria were then resuspended in 10 mM $MgCl_2$ and diluted to an optical density at 600 nm of 1.0. Cells from 2 ml were harvested by centrifugation and washed once with 10 mM $MgCl_2$, and genomic DNA was isolated using Qiagen Genomic tip 100/G (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

†Corresponding author: R. Koebnik; ralf.koebnik@ird.fr

R. Koebnik and V. Roman-Reyna contributed equally to this work.

*The *e*-Xtra logo stands for "electronic extra" and indicates that one supplementary table is published online.

The author(s) declare no conflict of interest.

Accepted for publication 28 March 2022.

Table 1. Repeat-variable diresidue (RVD) sequences of completely sequenced *Xanthomonas translucens* pv. *undulosa* strains[a]

TALE class	Strains	RVD pattern
TalCT	ICMP11055	HN HD HD HD NI NI NI HN HD HD NH NN NI NN HD
	LW16, XtFa1	NN HD HD HD NI NI NI HN HD HD NN NN NI NN HD
	MAI5034, P3, XtKm12, XtKm15, XtLr8, Xtu4699	HN HD HD HD NI NI NI HN HD HD NN NN NI NN HD
TalCZ	ICMP11055, LW16, MAI5034, P3, XtFa1, XtKm12, Xtu4699	NH NN HD NN HD NH HD YK NG NH Y* HD NN NI NG QD
TalDA	ICMP11055, MAI5034, P3, XtFa1, XtKm12, XtKm15, XtLr8, Xtu4699	HD YD NI NG NG NN YK NG HD NG NG ND NG QD NH HD
	LW16	HD YD NI NG NG NN YK NG HD NG NG ND NG QD NH QD
TalDB	LW16	NN HD KG HD HD HN NF NI NN HD HD HD HN HN HD
	P3	NN HD NG HD HD HN NF NI NN NN HD HD HN HN HD
	XtKm15,[b] XtLr8	NN HD NG HD HD HN NF NI NF HD HD HD HN HN HD
	Xtu4699	NN HD NG HD HD HN NF NI NH HD HD HD HN HN HD
TalDC	MAI5034, P3, XtFa1, XtLr8, Xtu4699	NN NG HD HD HD KG NN Y* NG HD HD QD HN
TalDD	ICMP11055	NN HD NG NN HN KG NI HD NI HN HD HN HD Y* NG HD HD HN
	LW16	NN HD NG NN HN KG NI HD NI HN HD HN HD HD NI HN HN HD
	MAI5034	NN HD NG NN HN KG NI HD NI HN HD HN HD HD NI HN HD QD
	P3, XtFa1, XtKm12, XtKm15, XtLr8, Xtu4699	NN HD NG NN HN KG NI HD NI NN HD HN HD HD NI HN HD QD
	LW16, XtKm12	NN HD NG NN HN NG NI HD NI NN HD HD NN NN NI HN HD
TalDE	LW16, MAI5034	NN HD NG NN HN HN NN NI NI NH NN HD NN NH HD HD
	P3, XtFa1, XtKm15,[b] Xtu4699	NN HD NG NN HN HN NI NI NI NH NN HD NN NH HD HD
	XtKm12	NN HD NG NN HN HN NI NI NI NH NN HD HN NH HD HD
TalDF	ICMP11055, LW16, MAI5034, P3, XtKm12, XtKm15, XtLr8, Xtu4699	HD HN HN HD NH NH HG HD KG NN Y* NG HD HD HN
	XtFa1	HD HN HN HD NH NH HG HD KG NN Y* NG HD NI NH NG HD HN
TalHM	ICMP11055	NN HD NG HD HD HG HD KG NN Y* NG NG HD HD QD HN
	XtKm15[b]	NN HD NG HD HD HG HD KG
TalHN	ICMP11055	NN HD NG HD NG HD HD HG HD KG NN KG HD HN QD HN
TalJD	XtLr8	NN HD NG NN Y* NG HD HD NN NH HD HD

[a] Alignment of RVDs of transcription activator-like effectors (TALEs) of *X. translucens* pv. *undulosa* strains. RVDs that differ between otherwise highly conserved TALEs are underlined. An asterisk (*) indicates that the second amino acid of the RVD is missing.

[b] Three TALE genes of strain XtKm15(*talDB*, *talDE*, and *talHM*) are recorded as pseudogenes due to in-frame stop codons or frameshift mutations (http://www.jstacs.de/downloads/List_of_classes.txt) (Grau et al. 2016).

For library construction and DNA sequencing, performed by ohmX.bio (Gent, Belgium), 1 μg of DNA was mechanically fragmented with g-Tubes (Covaris) at approximately 13 kb. A sequencing library was prepared with the Ligation Sequencing kit (SQK-LSK110; Oxford Nanopore Technologies [ONT]) and the Native Barcode Expansion (PCR-free) (EXP-NBD114; ONT) based on the manufacturer's protocol. For multiplexing with other samples, a unique barcode (barcode ID NB19, GTTCCTCGTGCAGTGTCAAGAGAT) was ligated to the sample type using the ONT direct-DNA (SQK-LSK109) library preparation kit in combination with the Native barcoding expansion (EXP-NBD104). Upon pooling of eight libraries, samples were sequenced on a GridION, R9.4 flow cell, and sequence reads were demultiplexed by ohmX.bio and provided as FASTQ files.

Sequence reads were trimmed with Porechop (v0.2.4) and assembled using three different algorithms: Flye (version 2.8.1-b1676), Shasta, and Miniasm (Kolmogorov et al. 2020; Li 2016; Shafin et al. 2020; Wick 2017). In addition, the Flye and Miniasm assemblies were polished using Racon (Vaser et al. 2017). A comparison of these assemblies revealed the superior performance of the Racon-polished Flye assembly, as indicated by its better contiguity, completeness, and quality (Supplementary Table). However, manual inspection revealed significant issues with homopolymeric nucleotide runs, resulting in four frame shifts per TALE gene. Therefore, we applied Homopolish (version 0.0.1) on the Racon-polished Flye assembly (Huang et al. 2021), which resolved all of the problems at the TALE genes and reduced the number of predicted pseudogenes from 315 to 202 (see below).

This procedure yielded one circular chromosome of 4,625,916 bp with a typical G+C content of 64.7%, corresponding to 139× sequencing coverage. The chromosome was annotated with GeneMarkS-2+ (Lomsadze et al. 2018), as implemented in the NCBI Prokaryotic Genome Annotation Pipeline (https://www.ncbi.nlm.nih.gov/genome/annotation_prok/), which predicted a total of 4,001 genes, including 3,736 coding genes, 202 pseudogenes, 53 transfer RNA genes, 4 noncoding RNAs, and 2 rRNA operons (5S, 16S, and 23S).

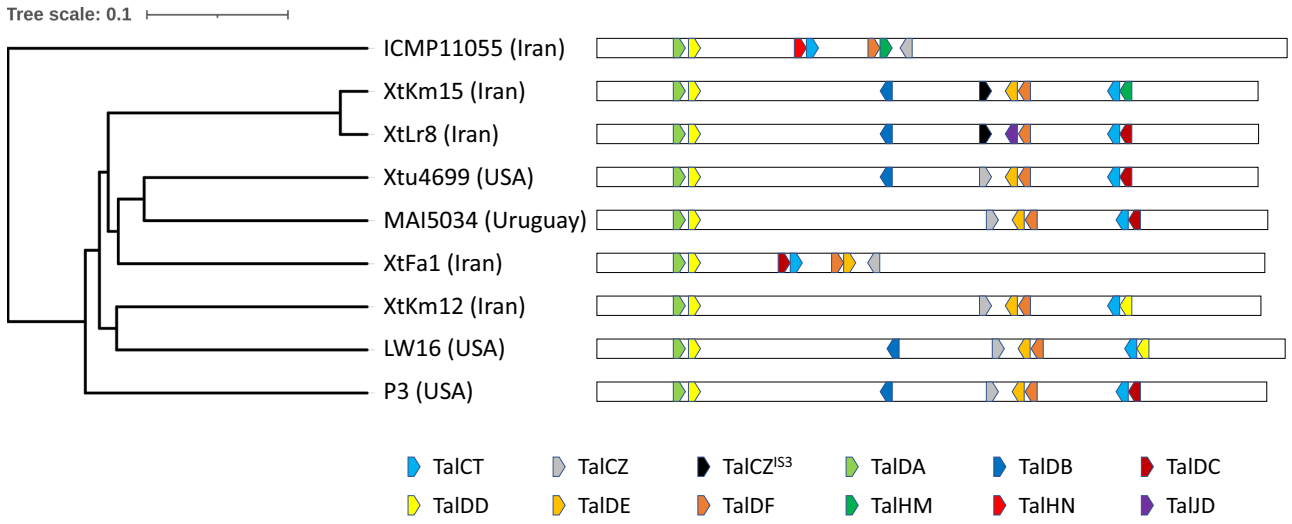

Fig. 1. Phylogenetic tree of *Xanthomonas translucens* pv. *undulosa* strains and repertoire of transcription activator-like (TAL) effector genes. Genome sequences were retrieved from GenBank (Sayers et al. 2022): ICMP 11055 (CP009750), XtKm15 (CP063997, CP063998, CP063999), XtLr8 (CP063993, CP063994, CP063995), Xtu4699 (CP008714), MAI5034 (CP089584), XtFa1 (CP063996), XtKm12 (CP064000), LW16 (CP043540), and P3 (CP043500). Calculation of genome-wide pairwise average nucleotide identities and phylogenetic analysis were performed on the enve-omics platform (http://enve-omics.ce.gatech.edu) (Rodriguez-R and Konstantinidis 2016). The unweighted pair-group method with arithmetic mean was used to build the phylogenetic tree (Sokal and Michener 1958). The interactive Tree Of Life suite was used for better visualization of the tree (https://itol.embl.de; Letunic and Bork 2019). Open rectangles symbolize the chromosomes of the *X. translucens* strains. For better comparability, genome sequences were rotated so that they start with the translation initiation codon of the *dnaA* gene. Position and orientation of TAL effector genes, as classified by the AnnoTale suite (Grau et al. 2016), are indicated by colored arrows. TAL effector genes are not drawn to scale. Genes encoding TalCZ in strains XtKm15 and XtLr8 are interrupted by the insertion of an IS element.

Type 3 effectors and, in particular, TALEs, are of specific interest when trying to understand the pathogenicity of xanthomonads and their host adaptation (Boch and Bonas 2010; Jacques et al. 2016; White et al. 2009). For this reason, TALEs were predicted and classified using the AnnoTale suite (Grau et al. 2016). In total, seven TALE genes were found, belonging to the classes TalCT, TalCZ, TalDA, TalDC, TalDD, TalDE, and TalDF, all of which are widely conserved in strains of *X. translucens* pv. *undulosa* (Table 1; Fig. 1) (Falahi Charkhabi et al. 2017; Peng et al. 2016; Shah et al. 2021), arguing for a recent worldwide expansion of this pathovar (Khojasteh et al. 2019). Notably, a homolog of the *tal8* gene in strain Xtu4699, which elevates expression of a 9-cis-epoxycarotenoid dioxygenase gene (*TaNCED-5BS*) in wheat, that encodes the rate-limiting step in the biosynthesis of the phytohormone abscisic acid for disease susceptibility (Peng et al. 2019), is present in strain MAI5034 (*talDC*). Two additional genes that contribute to virulence of strain ICMP 11055 are less conserved (Falahi Charkhabi et al. 2017). Whereas *tal4b* (*talHM*) is absent in strain MAI5034 and other strains of *X. translucens* pv. *undulosa* (Peng et al. 2016; Shah et al. 2021), strain MAI5034 encodes a novel allele of *tal2*, *talDD*.

The lower diversity of TALE repertoires in the pathovar *undulosa* as compared with the pathovar *translucens* is also mirrored by their lower genetic diversity, which was observed in a cohort of 178 strains of small-grain-infecting xanthomonads (Khojasteh et al. 2019) and may support the hypothesis that the pathovar *undulosa* emerged or originated from the pathovar *translucens*. Such a scenario is also supported by the observation that the *cbsA* gene, a genetic switch between vascular and nonvascular plant pathogenesis (Gluck-Thaler et al. 2020), is disrupted by an IS*1595*-family transposase in strain MAI5034 and other strains of *X. translucens* pv. *undulosa*, ultimately resulting in nonsystemic infection.

Data Availability

Raw reads and the complete genome were uploaded to the NCBI Sequence Read Archive and GenBank under BioProject accession PRJNA786744.

Acknowledgments

We thank J. Grau, Martin-Luther-Universität Halle-Wittenberg, Germany, for assistance with the AnnoTale suite.

Literature Cited

Boch, J., and Bonas, U. 2010. Xanthomonas AvrBs3 family-type III effectors: Discovery and function. Annu. Rev. Phytopathol. 48:419-436.

Clavijo, F., Curland, R. D., Croce, V., Lapaz, M. I., Dill-Macky, R., Pereyra, S., and Siri, M. I. I. 2022. Genetic and phenotypic characterization of Xanthomonas species pathogenic of wheat in Uruguay. Phytopathology 112:511-520.

Duveiller, E., Bragard, C., and Maraite, H. 1997. Bacterial leaf streak and black chaff caused by *Xanthomonas translucens*. Pages 25-32 in: The Bacterial Disease of Wheat: Concept and Methods of Disease Management. E. Duveiller, L. Fucikovskil, and K. Rudolph, eds. CIMMYT, D.F, Mexico.

Egli, T., Goto, M., and Schmidt, D. 1975. Bacterial wilt, a new forage grass disease. J. Phytopathol. 82:111-121.

Falahi Charkhabi, N., Booher, N. J., Peng, Z., Wang, L., Rahimian, H., Shams-Bakhsh, M., Liu, Z., Liu, S., White, F. F., and Bogdanove, A. J. 2017. Complete genome sequencing and targeted mutagenesis reveal virulence contributions of Tal2 and Tal4b of Xanthomonas translucens pv. undulosa ICMP11055 in bacterial leaf streak of wheat. Front. Microbiol. 8:1488.

Gluck-Thaler, E., Cerutti, A., Perez-Quintero, A. L., Butchacas, J., Roman-Reyna, V., Madhavan, V. N., Shantharaj, D., Merfa, M. V., Pesce, C., Jauneau, A., Vancheva, T., Lang, J. M., Allen, C., Verdier, V., Gagnevin, L., Szurek, B., Beckham, G. T., De La Fuente, L., Patel, H. K., Sonti, R. V., Bragard, C., Leach, J. E., Noël, L. D., Slot, J. C., Koebnik, R., and Jacobs, J. M. 2020. Repeated gain and loss of a single gene modulates the evolution of vascular plant pathogen lifestyles. Sci. Adv. 6:eabc4516.

Grau, J., Reschke, M., Erkes, A., Streubel, J., Morgan, R. D., Wilson, G. G., Koebnik, R., and Boch, J. 2016. AnnoTALE: Bioinformatics tools for identification, annotation, and nomenclature of TALEs from Xanthomonas genomic sequences. Sci. Rep. 6:21077.

Huang, Y. T., Liu, P. Y., and Shih, P. W. 2021. Homopolish: A method for the removal of systematic errors in nanopore sequencing by homologous polishing. Genome Biol. 22:95.

Jacques, M. A., Arlat, M., Boulanger, A., Boureau, T., Carrère, S., Cesbron, S., Chen, N. W., Cociancich, S., Darrasse, A., Denancé, N., Fischer-Le Saux, M., Gagnevin, L., Koebnik, R., Lauber, E., Noël, L. D., Pieretti, I., Portier, P., Pruvost, O., Rieux, A., Robène, I., Royer, M., Szurek, B., Verdier, V., and Vernière, C. 2016. Using ecology, physiology, and genomics to understand host specificity in Xanthomonas. Annu. Rev. Phytopathol. 54:163-187.

Jones, L. R., Johnson, A. G., and Reddy, C. S. 1917. Bacterial blight of barley. J. Agric. Res. 11:625-643.

Khojasteh, M., Taghavi, S. M., Khodaygan, P., Hamzehzarghani, H., Chen, G., Bragard, C., Koebnik, R., and Osdaghi, E. 2019. Molecular typing reveals high genetic diversity of Xanthomonas translucens strains infecting small-grain cereals in Iran. Appl. Environ. Microbiol. 85:e01518-19.

Kolmogorov, M., Bickhart, D. M., Behsaz, B., Gurevich, A., Rayko, M., Shin, S. B., Kuhn, K., Yuan, J., Polevikov, E., Smith, T. P. L., and Pevzner, P. A. 2020. metaFlye: Scalable long-read metagenome assembly using repeat graphs. Nat. Methods 17:1103-1110.

Letunic, I., and Bork, P. 2019. Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. Nucleic Acids Res. 47:W256-W259.

Li, H. 2016. Minimap and miniasm: Fast mapping and de novo assembly for noisy long sequences. Bioinformatics 32:2103-2110.

Lomsadze, A., Gemayel, K., Tang, S., and Borodovsky, M. 2018. Modeling leaderless transcription and atypical genes results in more accurate gene prediction in prokaryotes. Genome Res. 28:1079-1089.

Peng, Z., Hu, Y., Xie, J., Potnis, N., Akhunova, A., Jones, J., Liu, Z., White, F. F., and Liu, S. 2016. Long read and single molecule DNA sequencing simplifies genome assembly and TAL effector gene analysis of Xanthomonas translucens. BMC Genomics 17:21.

Peng, Z., Hu, Y., Zhang, J., Huguet-Tapia, J. C., Block, A. K., Park, S., Sapkota, S., Liu, Z., Liu, S., and White, F. F. 2019. Xanthomonas translucens commandeers the host rate-limiting step in ABA biosynthesis for disease susceptibility. Proc. Natl. Acad. Sci. U.S.A. 116:20938-20946.

Rodriguez-R, L. M., and Konstantinidis, K. T. 2016. The enveomics collection: A toolbox for specialized analyses of microbial genomes and metagenomes. PeerJ Preprints 4:e1900v1.

Sapkota, S., Mergoum, M., and Liu, Z. 2020. The translucens group of Xanthomonas translucens: Complicated and important pathogens causing bacterial leaf streak on cereals. Mol. Plant Pathol. 21:291-302.

Sayers, E. W., Cavanaugh, M., Clark, K., Pruitt, K. D., Schoch, C. L., Sherry, S. T., and Karsch-Mizrachi, I. 2022. GenBank. Nucleic Acids Res. 50:D161-D164.

Shafin, K., Pesout, T., Lorig-Roach, R., Haukness, M., Olsen, H. E., Bosworth, C., Armstrong, J., Tigyi, K., Maurer, N., Koren, S., Sedlazeck, F. J., Marschall, T., Mayes, S., Costa, V., Zook, J. M., Liu, K. J., Kilburn, D., Sorensen, M., Munson, K. M., Vollger, M. R., Monlong, J., Garrison, E., Eichler, E. E., Salama, S., Haussler, D., Green, R. E., Akeson, M., Phillippy, A., Miga, K. H., Carnevali, P., Jain, M., and Paten, B. 2020. Nanopore sequencing and the Shasta toolkit enable efficient de novo assembly of eleven human genomes. Nat. Biotechnol. 38:1044-1053.

Shah, S. M. A., Khojasteh, M., Wang, Q., Taghavi, S. M., Xu, Z., Khodaygan, P., Zou, L., Mohammadikhah, S., Chen, G., and Osdaghi, E. 2021. Genomics-enabled novel insight into the pathovar-specific population structure of the bacterial leaf streak pathogen Xanthomonas translucens in small grain cereals. Front. Microbiol. 12:674952.

Smith, E. F., Jones, L. R., and Reddy, C. S. 1919. The black chaff of wheat. Science 50:48.

Sokal, R. R., and Michener, C. D. 1958. A statistical method for evaluating systematic relationships. Univ. Kans. Sci. Bull. 38:1409-1438.

Vaser, R., Sović, I., Nagarajan, N., and Šikić, M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res. 27:737-746.

Vauterin, L., Hoste, B., Kersters, K., and Swings, J. 1995. Reclassification of Xanthomonas. Int. J. Syst. Bacteriol. 45:472-489.

White, F. F., Potnis, N., Jones, J. B., and Koebnik, R. 2009. The type III effectors of Xanthomonas. Mol. Plant Pathol. 10:749-766.

Wick, R. 2017. Porechop. Github. https://github.com/rrwick/Porechop