

# Transcriptomic data sets for *Novosphingobium aromaticivorans* DSM12444 and a $\Delta$ SARO\_RS14285 mutant grown in the presence of glucose and either protocatechuic, vanillic, syringic, or 4-coumaric acid

Laura Rodriguez-Castro,<sup>1,2</sup> Kevin S. Myers,<sup>1,2</sup> Daniel R. Noguera,<sup>1,2,3</sup> Timothy J. Donohue<sup>1,2,4</sup>

**AUTHOR AFFILIATIONS** See affiliation list on p. 3.

**ABSTRACT** The SARO\_RS14285 gene, encoding a transcription factor, was deleted in *Novosphingobium aromaticivorans* DSM12444. The transcriptomes of the parent and  $\Delta$ SARO\_RS14285 strains were determined when grown in medium containing glucose with or without protocatechuic, vanillic, syringic, or 4-coumaric acid. We present the raw RNA sequencing data obtained from these cultures.

**KEYWORDS** transcriptomics, aromatics, *Novosphingobium aromaticivorans*

Several cellular processes are regulated by transcription factors (TFs) including aromatic compound degradation pathways (1). *Novosphingobium aromaticivorans* DSM12444 is an Alphaproteobacterial species known to metabolize aromatic compounds (2–7). In the *N. aromaticivorans* DSM12444 genome, more than 20 TFs that could be involved in aromatic catabolism were identified (8). *N. aromaticivorans* DSM12444 can be engineered to funnel aromatic compounds found in lignocellulosic biomass into the bioplastic precursor 2-pyrone-4,6-dicarboxylic acid via the characterized *meta*-cleavage pathway (2). Immediately upstream of genes encoding known or predicted enzymes to function in aromatic metabolism (SARO\_RS14260 to SARO\_RS14300) is one gene encoding a predicted LysR-type TF (SARO\_RS14285). To determine the role of SARO\_RS14285 in *N. aromaticivorans* DSM12444, a mutant containing an in-frame deletion of the gene for this TF was constructed in the  $\Delta$ *sacB* parent strain (7) by homologous recombination using previously described methods (6, 7, 9). *N. aromaticivorans* DSM12444 was isolated from a polyaromatic hydrocarbon-contaminated site by Fredrickson *et al.* (10). The *sacB* gene (SARO\_RS09410, formerly Saro\_1879) was deleted previously from *N. aromaticivorans* DSM12444 (7), and the mutant strain was stored at  $-80^{\circ}\text{C}$  to allow the use of *sacB*-suicide plasmids for genetic modifications. To identify the role of this TF, a transcriptomic analysis was performed with the parent and mutant strains cultured in glucose alone or glucose plus protocatechuic acid (PCA), vanillic acid (VA), syringic acid (SA), or 4-coumaric acid (4-CA). The  $\Delta$ *sacB* parent strain *N. aromaticivorans* (DSM12444 $\Delta$ 09410) and  $\Delta$ SARO\_RS14285 mutant strain (DSM12444 $\Delta$ 09410 $\Delta$ 14285) were cultured aerobically at  $30^{\circ}\text{C}$  in standard mineral base (SMB) medium (7) supplemented with 10 mM glucose and 5 mM of an aromatic compound (PCA, 4-CA, VA, or SA) until mid-log phase. The cells were pelleted, frozen in a dry ice bath, and stored at  $-80^{\circ}\text{C}$ . They were later thawed, lysed, and RNA was extracted using a hot acid phenol:chloroform extraction, as previously described (11). The samples were treated with RNase-free DNase, and RNA was purified using the RNeasy Kit (Qiagen, Hilden, Germany).

RNA-seq library preparation and sequencing was performed by the Joint Genome Institute (JGI) using standard protocols. rRNA in the samples was depleted using the

**Editor** Leighton Pritchard, University of Strathclyde, Glasgow, United Kingdom

Address correspondence to Timothy J. Donohue, tdonohue@bact.wisc.edu.

The authors declare no conflict of interest.

**Received** 14 March 2025

**Accepted** 9 July 2025

**Published** 31 July 2025

Copyright © 2025 Rodriguez-Castro *et al.* This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

TABLE 1 Summary of RNA-seq sample data

Sample	Strain	Carbon source(s)	Replicate number	Total filtered paired reads	SRA accession number	Raw FASTQ JGI file name
NG_1	Parent	Glucose	1	8,261,114	SRX27744256	52934.3.507152.AAGTCGAGAT-AAGTCGAGG.fastq.gz
NG_2	Parent	Glucose	2	7,998,812	SRX27744257	52934.3.507152.GGACTAGAAT-GGACTAGAG.fastq.gz
NPG_1	Parent	PCA + glucose	1	6,902,516	SRX27744258	52934.3.507152.CTCATCAGAT-CTCATCAGG.fastq.gz
NPG_2	Parent	PCA + glucose	2	7,400,166	SRX27744259	52934.3.507152.TCGTAGTCAT-TCGTAGTCG.fastq.gz
N4G_1	Parent	4-CA + glucose	1	8,765,942	SRX27744260	52934.3.507152.ACCTTCTCAT-ACCTTCTCG.fastq.gz
N4G_2	Parent	4-CA + glucose	2	11,861,436	SRX27744261	52934.3.507152.TGGACTCTAT-TGGACTCTG.fastq.gz
NVG_1	Parent	VA + glucose	1	11,170,954	SRX27744262	52934.3.507152.TCTAACGCAT-TCTAACGCG.fastq.gz
NVG_2	Parent	VA + glucose	2	10,620,082	SRX27744263	52934.3.507152.AGCTTGAGAT-AGCTTGAGG.fastq.gz
NSG_1	Parent	SA + glucose	1	9,324,358	SRX27744264	52934.3.507152.GGTATAGGAT-GGTATAGGG.fastq.gz
NSG_2	Parent	SA + glucose	2	12,992,624	SRX27744265	52934.3.507152.CCGGAATTAT-CCGGAATTG.fastq.gz
LG_1	Mutant	Glucose	1	10,416,308	SRX27744266	52934.3.507152.TGTTGTGGAT-TGTTGTGGG.fastq.gz
LG_2	Mutant	Glucose	2	8,867,776	SRX27744267	52934.3.507152.TCCTGCTAAT-TCCTGCTAG.fastq.gz
LPG_1	Mutant	PCA + glucose	1	10,786,372	SRX27744268	52934.3.507152.GGAAGGATAT-GGAAGGATG.fastq.gz
LPG_2	Mutant	PCA + glucose	2	9,115,976	SRX27744269	52934.3.507152.TGGAGTTGAT-TGGAGTTGG.fastq.gz
LPG_3	Mutant	PCA + glucose	3	10,065,984	SRX27744270	52934.3.507152.CAATGTGGAT-CAATGTGGG.fastq.gz
L4G_1	Mutant	4-CA + glucose	1	10,814,906	SRX27744271	52934.3.507152.CGAAGAACAT-CGAAGAACG.fastq.gz
L4G_2	Mutant	4-CA + glucose	2	11,394,736	SRX27744272	52934.3.507152.GACGATCTAT-GACGATCTG.fastq.gz
L4G_3	Mutant	4-CA + glucose	3	9,511,994	SRX27744273	52934.3.507152.AGATGAGGAT-AGATGAGGG.fastq.gz
LVG_1	Mutant	VA + glucose	1	9,905,952	SRX27744274	52934.3.507152.ATGACCAGAT-ATGACCAGG.fastq.gz
LVG_2	Mutant	VA + glucose	2	9,338,470	SRX27744275	52934.3.507152.ATCGGTGTAT-ATCGGTGTG.fastq.gz
LVG_3	Mutant	VA + glucose	3	12,642,560	SRX27744276	52934.3.507152.CTGGTTCTAT-CTGGTTCTG.fastq.gz
LSG_1	Mutant	SA + glucose	1	11,232,974	SRX27744277	52934.3.507152.TCCGTATGAT-TCCGTATGG.fastq.gz
LSG_2	Mutant	SA + glucose	2	9,211,142	SRX27744278	52934.3.507152.GGATCTCAT-GGATCTTCG.fastq.gz
LSG_3	Mutant	SA + glucose	3	8,357,162	SRX27744279	52934.3.507152.TGATACGCAT-TGATACGCG.fastq.gz

QIaseq FastSelect 5S/16S/23S, rRNA Plant, rRNA Yeast, and custom rRNA Algae Depletion kits (Qiagen, Hilden, Germany) and libraries were constructed using the TruSeq Stranded mRNA kit (Illumina, San Diego, CA, USA) following standard JGI protocols. Sequencing of the flow cell was performed on the Illumina NovaSeq sequencer using NovaSeq XP V1.5 reagent kits, S4 flow cell as a 2 × 151 indexed run. Reads were quality controlled using reported JGI protocols. BBDuk (version 39.01, JGI, default parameters) was used to remove adapter sequences, remove reads with multiple “N” bases, remove reads ≤49 bp or quality score <10. Reads were mapped with BBMap (version 39.01, default parameters) (12) to select organisms including human, cat, dog, mouse, and microbial contaminants as well as ribosomal RNA species. Reads were removed that aligned with any reference sequence set with ≥93% identity. We submitted the filtered FASTQ reads to NCBI Sequence Read Archive (SRA), and raw FASTQ reads are available directly from JGI (Table 1).

RNA-seq data sets have proven useful for studying metabolism in *N. aromaticivorans* (11). The reported transcriptomic data sets will be useful to decipher the regulatory network associated with the *N. aromaticivorans* LysR-type TF SARO\_RS14285.

## ACKNOWLEDGMENTS

This material is based upon work supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Biological and Environmental Research Program under Award Number DE-SC0018409. The work (award DOI: 10.46936/10.25585/60008808) conducted by the U.S. Department of Energy Joint Genome Institute (<https://ror.org/04xm1d337>) is supported by the Office of Science of the U.S. Department of Energy operated under contract no. DE-AC02-05CH11231.

## AUTHOR AFFILIATIONS

<sup>1</sup>DOE Great Lakes Bioenergy Research Center, Universities of Wisconsin, Madison, Wisconsin, USA

<sup>2</sup>Wisconsin Energy Institute, University of Wisconsin, Madison, Wisconsin, USA

<sup>3</sup>Department of Civil and Environmental Engineering, Universities of Wisconsin, Madison, Wisconsin, USA

<sup>4</sup>Department of Bacteriology, Universities of Wisconsin, Madison, Wisconsin, USA

## AUTHOR ORCID*s*

Laura Rodriguez-Castro  <http://orcid.org/0000-0001-9974-4752>

Kevin S. Myers  <http://orcid.org/0000-0003-3302-3877>

Daniel R. Noguera  <http://orcid.org/0000-0003-0372-3063>

Timothy J. Donohue  <http://orcid.org/0000-0001-8738-2467>

## AUTHOR CONTRIBUTIONS

Laura Rodriguez-Castro, Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review and editing | Kevin S. Myers, Data curation, Formal analysis, Investigation, Software, Writing – review and editing | Daniel R. Noguera, Conceptualization, Funding acquisition, Investigation, Supervision, Writing – review and editing | Timothy J. Donohue, Conceptualization, Funding acquisition, Investigation, Supervision, Writing – review and editing

## DATA AVAILABILITY

Filtered RNA-seq FASTQ files are available from NCBI SRA, BioProject Accession Number [PRJNA1226177](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1226177)). Raw RNA-seq FASTQ files and other sequencing files are available for download from the JGI Data Portal (<https://data.jgi.doe.gov/>, Proposal ID 510025).

## REFERENCES

- Baugh AC, Momany C, Neidle EL. 2023. Versatility and complexity: common and uncommon facets of LysR-type transcriptional regulators. *Annu Rev Microbiol* 77:317–339. <https://doi.org/10.1146/annurev-micro-050323-040543>
- Perez JM, Kontur WS, Alherech M, Coplien J, Karlen SD, Stahl SS, Donohue TJ, Noguera DR. 2019. Funneling aromatic products of chemically depolymerized lignin into 2-pyrone-4-6-dicarboxylic acid with *Novosphingobium aromaticivorans*. *Green Chem* 21:1340–1350. <https://doi.org/10.1039/C8GC03504K>
- Vilbert AC, Kontur WS, Gille D, Noguera DR, Donohue TJ. 2024. Engineering *Novosphingobium aromaticivorans* to produce *cis,cis*-muconic acid from biomass aromatics. *Appl Environ Microbiol* 90. <https://doi.org/10.1128/aem.01660-23>
- Perez JM, Kontur WS, Gehl C, Gille DM, Ma Y, Niles AV, Umana G, Donohue TJ, Noguera DR. 2021. Redundancy in aromatic O-demethylation and ring opening reactions in *Novosphingobium aromaticivorans* and their impact in the metabolism of plant derived phenolics. *Appl Environ Microbiol* 87:e02794–20. <https://doi.org/10.1128/AEM.02794-20>
- Linz AM, Ma Y, Perez JM, Myers KS, Kontur WS, Noguera DR, Donohue TJ. 2021. Aromatic dimer dehydrogenases from *Novosphingobium aromaticivorans* reduce monoaromatic diketones. *Appl Environ Microbiol* 87:e0174221. <https://doi.org/10.1128/AEM.01742-21>
- Kontur WS, Olmsted CN, Yusko LM, Niles AV, Walters KA, Beebe ET, Vander Meulen KA, Karlen SD, Gall DL, Noguera DR, Donohue TJ. 2019. A heterodimeric glutathione S-transferase that stereospecifically breaks lignin's  $\beta$ (R)-aryl ether bond reveals the diversity of bacterial  $\beta$ -etherases. *Journal of Biological Chemistry* 294:1877–1890. <https://doi.org/10.1074/jbc.RA118.006548>
- Kontur WS, Bingman CA, Olmsted CN, Wassarman DR, Ulbrich A, Gall DL, Smith RW, Yusko LM, Fox BG, Noguera DR, Coon JJ, Donohue TJ. 2018. *Novosphingobium aromaticivorans* uses a Nu-class glutathione S-transferase as a glutathione lyase in breaking the  $\beta$ -aryl ether bond of lignin. *J Biol Chem* 293:4955–4968. <https://doi.org/10.1074/jbc.RA117.001268>
- Myers KS, Linz AM, Baumgart LA, Zhang Y, Blaby I, Mewalal R, Ng V, Noguera DR, Donohue TJ. 2025. Using DNA affinity purification sequencing (DAP-seq) to identify *in vitro* binding sites of potential *Novosphingobium aromaticivorans* DSM12444 transcription factors. *Microbiol Resour Announc*. <https://doi.org/10.1128/mra.00260-25>
- Rodriguez-Castro L, Myers KS, Linz AM, Mettert EL, Camp W, Kiley PJ, et al. 2025. *Novosphingobium aromaticivorans* AtfA coordinates transcription of multiple aromatic metabolic pathways. *mSystems*
- Fredrickson JK, Brockman FJ, Workman DJ, Li SW, Stevens TO. 1991. Isolation and characterization of a subsurface bacterium capable of growth on toluene, naphthalene, and other aromatic compounds. *Appl Environ Microbiol* 57:796–803. <https://doi.org/10.1128/aem.57.3.796-803.1991>
- Metz F, Olsen AM, Lu F, Myers KS, Allemann MN, Michener JK, Noguera DR, Donohue TJ. 2024. Catabolism of  $\beta$ -5 linked aromatics by *Novosphingobium aromaticivorans*. *MBio* 15:e0171824. <https://doi.org/10.1128/mbio.01718-24>
- Bushnell B. 2014. BBMap: a fast, accurate, splice-aware aligner. Available from: <https://sourceforge.net/projects/bbmap>