

## The *tpm* metabarcoding DNA sequence database for taxonomic allocations using the Mothur and DADA2 bio-informatic tools (Version 2.0.0)

A.C.M. Pozzi<sup>1</sup>, R. Bouchali<sup>1</sup>, L. Marjolet<sup>1</sup>, B. Cournoyer<sup>1</sup>

<sup>1</sup> *University of Lyon, UMR Ecologie Microbienne Lyon (LEM), CNRS 5557, INRAE 1418, Université Claude Bernard Lyon 1, VetAgro Sup, Research Team “Bacterial Opportunistic Pathogens and Environment” (BPOE), 69280 Marcy L’Etoile, France.*

### Corresponding authors:

- A.C.M. Pozzi, UMR Microbial Ecology, CNRS 5557, CNRS 1418, VetAgro Sup, Main building, aisle 3, 1st floor, 69280 Marcy-L’Etoile, France. Tel. (+33) 478 87 39 47. Fax. (+33) 472 43 12 23. Email: [adrien.meynier\\_pozzi@vetagro-sup.fr](mailto:adrien.meynier_pozzi@vetagro-sup.fr)
- B. Cournoyer, UMR Microbial Ecology, CNRS 5557, CNRS 1418, VetAgro Sup, Main building, aisle 3, 1st floor, 69280 Marcy-L’Etoile, France. Tel. (+33) 478 87 56 47. Fax. (+33) 472 43 12 23. Email: [and\\_benoit.cournoyer@vetagro-sup.fr](mailto:and_benoit.cournoyer@vetagro-sup.fr)

### Keywords:

BACtpm, Bacteria, *tpm*, thiopurine-S-methyltransferase EC:2.1.1.67, Nucleotide sequences, PCR products, Next-Generation-Sequencing, OTHU

### Description:

- The *tpm* gene codes for the thiopurine-S-methyltransferase (TPMT), an enzyme that can detoxify metalloid-containing oxyanions and xenobiotics (Cournoyer et al., 1998). Bacterial TPMTs radiated apart from human and animal TPMTs, and showed a vertical evolution in line with the 16S rRNA gene molecular phylogeny (Favre-Bonté et al., 2005).
- The *tpm* database, named BACtpm, was designed to apply the *tpm*-metabarcoding analytical scheme published in Aigle et al. (2021). It includes the full *tpm* identifiers, GenBank accession numbers, complete taxonomic records (domain down to strain code) of about 215 nucleotide-long *tpm* sequences of 840 unique taxa belonging to 139 genera.
- Nucleotide sequences of *tpm* (range: 190-233 nucleotides) were either retrieved from public repositories (GenBank) or made available by B. Cournoyer’s research group. Colin et al. (2020) described the PCR and high throughput Illumina Miseq DNA sequencing procedures used to produce *tpm* sequences.
- BACtpm v.2.0.0 (June 2021 release) is made available under the Creative Commons Attribution 4.0 International Licence. It can be used for the taxonomic allocations of *tpm* sequences down to the species and strain levels. Data is stored in the csv format enabling future user to reformat it to fit their specific needs.

## Acknowledgments:

We thank the worldwide community of microbiologists who made contributions to public databases in the past decades, and made possible the elaboration of the BACtpm database. We also thank the Field Observatory in Urban Hydrology (OTHU, [www.graie.org/othu/](http://www.graie.org/othu/)), Labex IMU (Intelligence des Mondes Urbains), the Greater Lyon Urban Community, the School of Integrated Watershed Sciences H2O'LYON, and the Lyon Urban School for their support in the development of this database. This work was funded by the French national research program for environmental and occupational health of ANSES under the terms of project "louqmer" EST 2016/1/120, l'Agence Nationale de la Recherche through ANR-16-CE32-0006, ANR-17-CE04-0010, ANR-17-EURE-0018 and ANR-17-CONV-0004, by the MITI CNRS project named Urbamic, and the French water agency for the Rhône, Mediterranean and Corsica areas through the Desir and Domic projects.

## Cite as:

A.C.M. Pozzi, R. Bouchali, L. Marjolet, B. Cournoyer **The *tpm* metabarcoding DNA sequence database for taxonomic allocations using the Mothur and DADA2 bio-informatic tools (Version 2.0.0)**, 2021, <https://zenodo.org/>, BACtpm v2.0.0, doi: 10.5281/zenodo.4492211

## References:

- Aigle, A., Colin, Y., Bouchali, R., Bourgeois, E., Marti, R., Ribun, S., Marjolet, L., Pozzi, A.C.M., Misery, B., Colinon, C., Bernardin-Souibgui, C., Wiest, L., Blaha, D., Galia, W., Cournoyer, B., 2021. Spatio-temporal variations in chemical pollutants found among urban deposits match changes in thiopurine S-methyltransferase-harboring bacteria tracked by the *tpm* metabarcoding approach. *Sci. Total Environ.* 767, 145425. <https://doi.org/10.1016/j.scitotenv.2021.145425>
- Colin, Y., Bouchali, R., Marjolet, L., Marti, R., Vautrin, F., Voisin, J., Bourgeois, E., Rodriguez-Nava, V., Blaha, D., Winiarski, T., Mermillod-Blondin, F., Cournoyer, B., 2020. Coalescence of bacterial groups originating from urban runoffs and artificial infiltration systems among aquifer microbiomes. *Hydrol. Earth Syst. Sci.* 24, 4257–4273. <https://doi.org/10.5194/hess-24-4257-2020>
- Cournoyer, B., Watanabe, S., Vivian, A., 1998. A tellurite-resistance genetic determinant from phytopathogenic pseudomonads encodes a thiopurine methyltransferase: evidence of a widely-conserved family of methyltransferases1The International Collaboration (IC) accession number of the DNA sequence is L49178.1. *Biochim. Biophys. Acta BBA - Gene Struct. Expr.* 1397, 161–168. [https://doi.org/10.1016/S0167-4781\(98\)00020-7](https://doi.org/10.1016/S0167-4781(98)00020-7)
- Favre-Bonté, S., Ranjard, L., Colinon, C., Prigent-Combaret, C., Nazaret, S., Cournoyer, B., 2005. Freshwater selenium-methylating bacterial thiopurine methyltransferases: diversity and molecular phylogeny. *Environ. Microbiol.* 7, 153–164. <https://doi.org/10.1111/j.1462-2920.2004.00670.x>