Data in Brief

Dataset on endophytic and rhizoplane fungi on the roots of wild grapevine in Croatia and Bosnia and Herzegovina

Tomislav Radić\textsuperscript{a,1}, Matevž Likar\textsuperscript{b,1,*}, Katarina Hančević\textsuperscript{a}, Marjana Regvar\textsuperscript{b}, Mate Čarija\textsuperscript{a}, Goran Zdunić\textsuperscript{a}

\textsuperscript{a}Institute for Adriatic Crops and Karst Reclamation, Put Duilova 11, 21000 Split, Croatia
\textsuperscript{b}Biotechnical Faculty, Department of Biology, University of Ljubljana, Večna pot 111, Ljubljana, Slovenia

A R T I C L E   I N F O

Article history:
Received 29 November 2020
Revised 15 December 2020
Accepted 21 December 2020
Available online 24 December 2020

Keywords:
Plant-microbe interaction
Endophyte
Ectomycorrhizal fungi
Arbuscular mycorrhizas
Pathogenic fungi

A B S T R A C T

The data presented in this article are related to the research article entitled “Root-associated community composition and co-occurrence patterns of fungi in wild grapevine”. This dataset documents the diversity of endophytic and rhizoplane fungi found on the roots of 38 wild grapevine plants growing at four locations (Krka, Neretva, Psunj and Paklenica) in Croatia and Bosnia and Herzegovina. We record here 48 identified fungal operational taxonomic units (OTUs) from 3 phyla, 23 families and 30 genera. The material in this Data in Brief paper comprised the data on the identification of OTUs with their corresponding ecological niche, abundance distribution of each fungal OTU recorded in each of the host and location, restrictedness and with fit to the Sloan neutral community model.

© 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

DOI of original article: 10.1016/j.funeco.2020.101034
* Corresponding author.
\textsuperscript{1} These two authors contributed equally to this study

https://doi.org/10.1016/j.dib.2020.106692
2352-3409/© 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Specifications Table

| Subject                  | Biology; Plant Science: Plant Microbe Interaction |
|--------------------------|--------------------------------------------------|
| Specific subject area    | Plant Microbe Interaction, Fungal Endophytes, Arbuscular mycorrhizal fungi, Ectomycorrhizae, Pathogen Fungi |
| Type of data             | Table                                            |
| How data were acquired   | Raw data: DNA was extracted from collected young grapevine roots amplified using fungi specific primers and sequenced. The individual sequences were sequenced both ways to ensure correct sequencing of individual samples. All sequences were inspected for correct reading of the sequencer traces into nucleotide sequences. We collected the data on number of species and abundance for each individual grapevine plant. Analysed data: raw data was used to calculate the restrictedness for individual OTU and it fit to the Sloan’s neutral model in R. |
| Data format              | Raw numbers in matrices of abundances for individual OTU for individual sampling location. Analysed data on restrictedness of OTUs and their fit to Sloan’s neutral model. |
| Parameters for data collection | We examined 38 wild grapevine plants from four sampling locations (Krka, Neretva, Psnunj and Paklenica) in Croatia and Bosnia and Herzegovina. Young roots and the surrounding bulk soil (~500 cm) were sampled at a depth of 0 cm to 30 cm, following from the base of the grapevine stems to the terminal sections of the roots, to ascertain correct root origin. The samples were kept frozen at -20°C prior to analyses. DNA was extracted from collected young grapevine roots using DNeasy Plant Mini kits (Invitrogen, San Diego, USA), according to the manufacturer instructions. The internal transcribed spacer (ITS) rDNA regions were amplified using fungi specific primers (ITS1F and ITS4) and cloned using the pGEM-T Easy Vector system (Promega). Suitable PCR products from the pGEM-T Easy vectors were sequenced (ABI 3730xI DNA analyser; Macrogen Co., Netherlands). The individual sequences were sequenced both ways to ensure correct sequencing of individual samples. |
| Description of data collection | DNA was extracted from collected young grapevine roots using DNeasy Plant Mini kits (Invitrogen, San Diego, USA), according to the manufacturer instructions. The internal transcribed spacer (ITS) rDNA regions were amplified using fungi specific primers (ITS1F and ITS4) and cloned using the pGEM-T Easy Vector system (Promega). Suitable PCR products from the pGEM-T Easy vectors were sequenced (ABI 3730xI DNA analyser; Macrogen Co., Netherlands). The individual sequences were sequenced both ways to ensure correct sequencing of individual samples. |
| Data source location     | Four sampling locations from Croatia and Bosnia: Krka: 43.8019, 15.96597; Neretva: 42.90113, 17.92035; Paklenica: 44.30387, 15.47147; Psnunj: 45.39077, 17.45578 |
| Data accessibility       | Repository name: Mendeley Datasets Data identification number: http://dx.doi.org/10.17632/tjt8kkk66h.1 Direct URL to data: https://data.mendeley.com/datasets/tjt8kkk66h.1 |
| Related research article | T. Radić, M. Likar, K. Hančević, M. Regvar, M. Čarija, G. Zdunić, Root-associated community composition and co-occurrence patterns of fungi in wild grapevine. Fungal Ecology. https://doi.org/10.1016/j.funeco.2020.101034. |

Value of the Data

- These data are essential for further phylogenetical, ecological, and biogeographical studies. They will assist in the examination of spatial and temporal variations in the community structure of wild grapevines. They can be useful to compare population or community characteristics of wild grapevine vs. cultivated grapevine (vineyards) to assess the potential provided in sustainable agriculture. They can also be useful to examine interactions between ectomycorrhizal fungi and fungal pathogens on plant roots.
- These kinds of data could be useful for general biologists, biogeographers, ecologists, and plant pathologists. They can also be useful for owners of sustainable vineyards as it could provide them with information on potential disease incidence in connection with colonization by beneficial fungi. These data can be useful for regulatory agencies and stakeholders who seek to protect the public and their goods or values by limiting the adverse environmental impacts of development.
- These data can be used to further insights into interactions between ectomycorrhizal fungi and fungal pathogens on plant roots, which would benefit grapevine production. They can also be used for future efforts in designing guidelines for the protection of specific pristine environments that host wild grapevine.
1. Data Description

Tables can be found on the Mendeley dataset repository:

Table 1 contains 48 Operational taxonomic units (OTUs) represented by 52 fungal sequences obtained from the roots of wild grapevine at four sampling locations. We include the percent of the match with the nearest match in the NCBI Genbank, along with our accession number. For OTUs we provide their ecological niche (saprotroph, pathotroph, endophyte, or symbiotroph) according to available published literature. Taxonomic affiliation of OTUs to taxonomic family and phylum according to NCBI Taxonomy is also included.

Table 2 contains abundances for each fungal operational taxonomic unit (OTU) on the roots of wild grapevine at the individual sampling location. We also included the restrictedness index for each OUT. Each species (or individual) in the species pool has a restrictedness value, which ranges from 0 to 1, where 0 indicates that the focal species is present for all of the sites. Additionally, the information on the fit of individual OTU to Sloan’s neutral model is provided in column Partition. Samples were sorted into three partitions that depended on whether they occurred more frequently than (the ‘above’ partition), less frequently than (the ‘below’ partition), or within (the ‘neutral’ partition) the 95% confidence interval of the neutral model predictions. Also provided are calculated mean relative abundance (across all locations), observed and predicted frequency across all locations, as well as confidence intervals provided by the model.

2. Experimental Design, Materials and Methods

Three areas were identified as reservoirs of wild grapevines in Croatia: the National Park of Krka River and the National Park of Mount Paklenica in the middle coastal area; and Mount Psunj, in eastern continental Croatia. A further area was identified in south-western Bosnia and Herzegovina: the lower area of Neretva River. The roots of wild grapevines were sampled at 38 sites in total, with each represented by an individual plant. Young wild grapevine roots were sampled at a depth of 0–30 cm, following from the base of the grapevine stems to the terminal sections of the roots, to ascertain the correct root origin. All samples were collected during subsequent samplings in June of years 2013–2016. In total, fifteen wild grapevines were detected and sampled around Neretva river, ten at Paklenica, eight at Psunj, and five around the Krka river. Before analyses, roots for molecular analyses were frozen and kept at −20 °C.

The young grapevine roots collected from the 38 individual plants were washed and 150 mg roots fresh weight per sample was crushed in liquid nitrogen. This was the source for DNA extraction using DNeasy Plant Mini kits (Qiagen, Germantown, USA), according to the manufacturer’s instructions. PCR amplification was carried out for the internal transcribed spacer (ITS) rDNA regions, using the ITS1F (5’-CTTGGTCAT TTAGGGAAGTAA) and ITS4 (5’- TCCTCCGCTTATTGATATGC) primers [3,6]. Each PCR reaction mixture of 25 μL contained: 2.5 μL 10 × PCR buffer, 2.5 mM MgCl2, 200 μM of each dNTP, 500 nM of each primer, 0.75 U DNA polymerase, and 12.5 μL of a 100-fold diluted template. The PCR amplification program was as follows: 94 °C for 1 min, 35 cycles of denaturation at 94 °C for 35 s, annealing at 55 °C for 35 s, and elongation at 72 °C for 30 s. The elongation step was increased by 5 s in each following cycle. The final extension was at 72 °C for 10 min.

The extracted DNA was purified using Wizard SV (Promega, Madison, USA) gels and a PCR clean-up system (Promega, Madison, USA), according to the manufacturer’s instructions. It was then cloned using the pGEM-T Easy Vector system (Promega). PCR amplification was carried out for 10 positive colonies per sample, using the T7 and SP6 forward and reverse primers, respectively. The PCR program was: 2 min at 94 °C; 35 cycles of denaturation at 94 °C for 30 s, annealing at 50 °C for 30 s and elongation at 72 °C for 1 min; and finally 5 min at 72 °C. After visualization confirmation on 1% agarose gels, we sent the PCR products to Macrogen Co., Netherlands (ABI 3730xl DNA analyzer, Thermo Fisher Scientific, Waltham, USA) for Sanger sequencing. The individual sequences were sequenced both ways to ensure correct sequencing of individual sam-
ples. All sequences were inspected for correct reading of the sequencer traces into nucleotide sequences. Operational taxonomic units (OTUs) were formed from sequences assembled based on 97% similarity. These OTUs were compared to the NCBI archives through BLAST searches to assign putative taxonomic identities based on the sequence similarity. Ecological niches were selected based on a broad literature search with OTU as the search term. Taxonomic affiliation of OTUs to taxonomic family and phylum was selected according to NCBI Taxonomy.

Restrictedness was calculated using the Funrar package v1.2.2 [5]. This restrictedness defines how regionally rare a species is. The following function represents the species occupancy for a set of communities. Each species (or individual) in the species pool has a restrictedness value, which ranges from 0 to 1, where 0 indicates that the focal species is present for all of the sites.

To determine the potential importance of neutral processes for the community assembly, we assessed the fit of the Sloan neutral community model to the distributions of the fungal taxa (Sloan et al. 2006). The fitting of the migration rate parameter was performed in R using nonlinear least-squares fitting and the Minpack.lm package [2], with the code provided by Burns et al. [1]. The fit of the neutral model was also compared with the fit of a binomial distribution model, to determine whether incorporation of drift and dispersal limitations improved the fits of the model beyond just random sampling of the source metacommunity [4]. Sampling from a binomial distribution represents the case where the local communities are random subsets of the metacommunity in the absence of the processes of drift and dispersal limitations. The fits of the neutral and binomial models were compared using the Akaike information criterion (AIC) of each model. To analyze deviations from the neutral model predictions, we compared the composition and diversity of neutrally and non-neutrally distributed OTUs. To accomplish this, the samples were sorted automatically into three partitions that depended on whether they occurred more frequently than (the ‘above’ partition), less frequently than (the ‘below’ partition), or within (the ‘neutral’ partition) the 95% confidence interval of the neutral model predictions. The 95% confidence interval was determined using the Hmisc package v4.4-2 in R.

**Ethics Statement**

Not applicable.

**CRediT Author Statement**

Tomislav Radić: Conceptualization, Investigation, Formal analysis, Funding acquisition; Writing - Original Draft, Writing - Review & Editing; Matevž Likar: Conceptualization, Investigation, Formal analysis, Writing - Original Draft; Writing - Review & Editing, Data Curation; Katarina Hančević: Investigation, Formal analysis; Marjana Regvar: Writing - Original Draft; Matija Čarija: Writing - Original Draft; Goran Zdunić: Conceptualization, Funding acquisition, Supervision; .

**Declaration of Competing Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgments**

This study was funded by the Croatian Science Foundation [Grant agreement UIP-2014-09-9737] and supported by the Croatian-Slovenian Scientific Cooperation (“Wild grapevine root
symbionts – important factor of adaptation for this rare and endangered species”). The authors are grateful to C. Berrie, PhD, for improving the English in the manuscript.

References

[1] A.R. Burns, W.Z. Stephens, K. Stagaman, S. Wong, J.F. Rawls, K. Guillemin, B.J. Bohannan, Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development, ISME J. 10 (2016) 655–664.

[2] T.V. Elzhov, K.M. Mullen, A.N. Spiess, B Bolker, minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds, 2016 URL https://cran.r-project.org/package=minpack.lm.

[3] M. Gardes, T.D. Bruns, ITS primers with enhanced specificity of basidiomycetes: application to the identification of mycorrhizae and rusts, Mol. Ecol. 2 (1993) 113–118.

[4] W.T. Sloan, S. Woodcock, M. Lunn, I.M. Head, T.P. Curtis, Modeling taxa-abundance distributions in microbial communities using environmental sequence data, Microb. Ecol. 53 (2007) 443–455.

[5] C. Violle, W. Thuiller, N. Mouquet, F. Munoz, N.J.B. Kraft, M.W. Cadotte, S.W. Livingstone, D. Mouillot, Functional rarity: the ecology of outliers, Trends Ecol. Evol. 32 (2017) 356–367.

[6] T.J. White, T. Bruns, S. Lee, J. Taylor, Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics, in: M.A. Innis, D.H. Gelfand, J.J. Sninsky, T.J. White (Eds.), PCR Protocols: a Guide to Methods and Applications, Academic Press, London, 1990, pp. 315–322.