



HAL
open science

Estimating the bias related to DNA recovery from hemp stems for retting microbial community investigation

Eliane Bou Orm, Stéphane Sauvagère, Janick Rocher, Jean-Charles Benezet, Sandrine Bayle, Christian Siatka, Anne Bergeret, Luc Malhautier

► To cite this version:

Eliane Bou Orm, Stéphane Sauvagère, Janick Rocher, Jean-Charles Benezet, Sandrine Bayle, et al.. Estimating the bias related to DNA recovery from hemp stems for retting microbial community investigation. *Applied Microbiology and Biotechnology*, 2023, 107, pp.4665-4681. 10.1007/s00253-023-12582-5 . hal-04106389

HAL Id: hal-04106389

<https://imt-mines-ales.hal.science/hal-04106389v1>

Submitted on 8 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Estimating the bias related to DNA recovery from hemp stems for retting microbial community investigation

Eliane Bou Orm^{1,2} · Stéphane Sauvagère³ · Janick Rocher² · Jean-Charles Benezet¹ · Sandrine Bayle² · Christian Siatka⁴ · Anne Bergeret¹ · Luc Malhautier²

Abstract

The industrial hemp plant *Cannabis sativa* is a source of vegetable fiber for both textiles and biocomposite applications. After harvesting, the plant stems are laid out on the ground and colonized by microorganisms (bacteria and fungi) naturally present in the soil and on the stems. By producing hydrolytic enzymes that degrade the plant wall polymers, the natural cement that binds the fiber bundles together is removed, thus facilitating their dissociation (retting process) which is required for producing high-performant fibers. To investigate temporal dynamics of retting microbial communities (density levels, diversity, and structure), a reliable protocol for extracting genomic DNA from stems is mandatory. However, very little attention has been paid to the methodological aspects of nucleic acid extraction, although they are crucial for the significance of the final result. Three protocols were selected and tested: a commercial kit (FastDNA™ Spin Kit for soil), the Gns-GII procedure, and a custom procedure from the Genosol platform. A comparative analysis was carried out on soil and two different varieties of hemp stem. The efficiency of each method was measured by evaluating both the quantity and quality of the extracted DNA and the abundance and taxonomy of bacterial and fungal populations. The Genosol protocol provides interesting yields in terms of quantity and quality of genomic DNA compared to the other two protocols. However, no major difference was observed in microbial diversity between the two extraction procedures (FastDNA™ SPIN Kit and Genosol protocol). Based on these results, the FastDNA™ SPIN kit or the Genosol procedure seems to be suitable for studying bacterial and fungal communities of the retting process. It should be noted that this work has demonstrated the importance of evaluating biases associated with DNA recovery from hemp stems.

Key points

- *Metagenomic DNA was successfully extracted from hemp stem samples using three different protocols.*
- *Further evaluation was performed in terms of DNA yield and purity, abundance level, and microbial community structure.*
- *This work exhibited the crucial importance of DNA recovery bias evaluation.*

Keywords Hemp · Retting · Microbial communities · Genomic DNA · DNA extraction · Next-generation sequencing

Introduction

Industrial hemp *Cannabis sativa* L. is one of the oldest non-food crops that has been continuously cultivated for over 6000 years. During the nineteenth and twentieth centuries, hemp cultivation declined due to competition with cotton and synthetic fibers (such as glass fibers) and the increase of governmental controls on the recreational use of cannabinoids. However, cotton cultivation is restricted to sub-tropical climates and requires high amounts of water and agrochemicals to ensure good fiber yields, leading to renewed interest in alternative bast fiber crops such as hemp. Currently, the hemp industry exploits mainly dual-purpose

✉ Luc Malhautier
luc.malhautier@mines-ales.fr

¹ Polymers, Composites and Hybrids (PCH), IMT Mines Alès, 6 avenue de Clavières, 30319 Alès Cedex, France

² Laboratoire des Sciences des Risques (LSR), IMT Mines Alès, 6 avenue de Clavières, 30319 Alès Cedex, France

³ École de l'ADN, Université de Nîmes, 19 Grand Rue BP 81295, 30015 Nîmes cedex 1, France

⁴ UPR CHROME, Université de Nîmes, Place Gabriel Péri, 30000 Nîmes cedex 1, France

varieties, for both seeds and low-added-value fiber production for paper or insulation industries. At the same time, an increasing demand either for high added-value textile grade fibers that require low fineness and high tenacity of the fibers or for biocomposites including plant fibers (De Fazio et al. 2020; Mazian et al. 2020; Venkatarajan and Athijayamani 2020) is occurring. Therefore, hemp cultivation seems to be suitable to address these future perspectives and initiatives (Zimmiewska 2022).

After cultivation and harvesting, hemp stems undergo several successive processing steps before obtaining high-added-value grade fibers. The first step consists in placing freshly harvested and cut stems on the soil. In this way, they are colonized by a complex microbial community (bacteria and fungi) from both the soil and the surface of the stems. This process is called the dew or field retting. It is well known that the enzymatic activities of this complex microbial community may remove the pectin cement that binds the fibers bundles together, facilitating the fibers decohesion within the stems and thus the fiber mechanical extraction (Mazian et al. 2020).

These retting microbial communities are subject to various biotic and abiotic signals and factors that influence their functional and structural dynamics (Comeau et al. 2020). Temporal microbial dynamics in terms of structure and activity can be affected by plant cell wall structure, plant growth stage, soil composition, climatic conditions, and plant species (Djemiel et al. 2017). Characterization of both bacterial and fungal communities during retting is then a crucial step in studying their temporal dynamics. This characterization is carried on by performing population density (real-time PCR) as well as diversity and structure (high throughput sequencing). For this, efficient DNA extraction and purification procedures are required.

Soils are one of the major reservoirs of biological diversity. Bacteria and fungi represent a significant fraction of this biodiversity (Plassart et al. 2012). Many works have evaluated the performance and limitations of DNA extraction methods in soil (Robe et al. 2003; Gobbi et al. 2019; Guerra et al. 2020). Nevertheless, microbial DNA extraction from plant stems in general, and from hemp stems, in particular, is poorly known and the number of papers is strictly limited (Plassart et al. 2012; Ribeiro et al. 2015; Djemiel et al. 2017; Mazian et al. 2020; Law et al. 2020).

In this study, methodological aspects of extracting nucleic acids from microbial communities that colonized hemp stem during field retting are investigated. Previously adapted protocols for studying soil microbial communities are therefore selected.

Three methods allowing quantitative and qualitative DNA characterization were selected from different protocols mentioned in the literature (Petric et al. 2011; Plassart et al. 2012; Knauth et al. 2013; Lim et al. 2017). These protocols

were used in previous works for genomic DNA extraction from soils based on a commercial kit (FastDNA™ Spin Kit for soil) (MP-Biomedicals, Solon, OH, USA), the GnS-GII protocol (Plassart et al. 2012), and a recent protocol used by the Genosol platform (INRAE, Dijon, France) (Lazarevic et al. 2013; Lelievre 2020). The main steps of each method are as follows: desorption and lysis of microbial cells by both chemical and mechanical actions and deproteinization, precipitation, and washing of extracted nucleic acids. In contrast, the purification step of the extracted DNA samples differs between the three protocols.

The aim of this work is then to optimize genomic DNA extraction and purification methods from hemp stems for further quantitative and qualitative characterizations of bacterial and fungal communities. The efficiency of each method was evaluated using relevant criteria: DNA yield, DNA purity, abundance levels of bacterial and fungal populations, and finally, representation of microbial community structure.

Materials and methods

Preparation of soil and stem samples

The comparison between the three methods was performed on independent soil and stem samples of the dioecious Finola hemp variety and the monoecious Futura 75 as described in Fig. 1.

A clay soil (pH 8.4) was collected on September 3, 2021, from a hemp retting plot (Mas de la Valus, Bouquet, France, 44°9'52", 4°16'56"). Soil samples were freeze-dried, sieved (1.5 mm), and stored at -80 °C. These samples were named S-R0.

Hemp stems of the dioecious variety Finola were harvested at the end of flowering (August 30, 2021) (Canebounes la Divine, Mérindol, France, 43°45'13", 5°13'51") (<https://canebounes.fr/>) (Supplemental Fig. S1). Stems were stored at -80 °C and cut into 5-cm-long pieces. The moisture content of the stems was 75%. These samples were named Fi-R0.

Hemp stems of the monoecious variety Futura 75 were harvested at the end of flowering (September 2, 2021) (Drome Chanvre, Mirabel et Blacons, France, 44°42'35", 5°5'29") (<http://www.dromechanvre.fr/>) which is 168 km north of Mérindol (Supplemental Fig. S1). These stems were retted from September 3, 2021, to October 17, 2021, in Mas de la Valus (Bouquet, France) on a similar clayey agricultural soil to that on which the hemp was grown for logistics reasons (close to the laboratory site). Unretted stems (R0) and retted stems for 2 weeks (R2) and 4 weeks (R4) were stored at -80 °C. Stems were cut into 5-cm-long pieces. The moisture content of the unretted stems was 40%, as it was 11% for the stems retted for 2 weeks and 43% for the

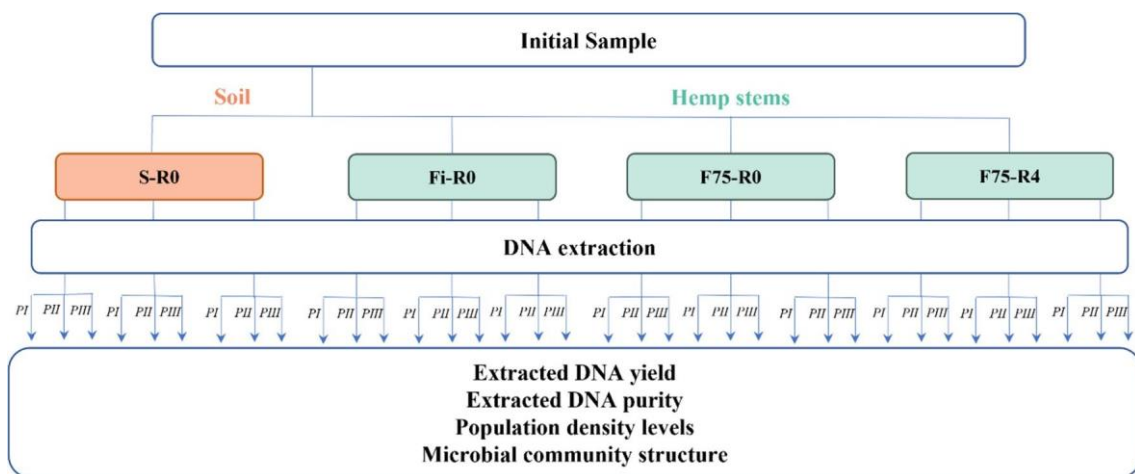


Fig. 1 Experimental design. S-R0, soil samples (triplicates); Fi-R0, unretted Finola stems (triplicates); F75-R0, unretted Futura 75 stems (triplicates); F75-R4, Futura 75 stems retted for 4 weeks (triplicates).

P I, Protocol I: commercial FastDNA™ Spin Kit for soil; P II, Protocol II: GnS-GII procedure; P III, Protocol III: custom procedure from the Genosol platform

stems retted for 4 weeks. These samples were named F75-R0 (unretted Futura 75 stems), F75-R2 (Futura 75 stems retted for 2 weeks), and F75-R4 (Futura 75 stems retted for 4 weeks).

Extraction and purification of genomic DNA from the soil and stems

The extraction of genomic DNA from soil and stems was performed according to the following methods:

- Protocol I: FastDNA™ Spin Kit for soil (MP-Biomedicals, Solon, OH, USA);
- Protocol II: GnS-GII + PVPP columns and Bio-Spin® P-6 Gel (Bio-Rad, Marnes-La-Coquette, France) and GeneClean® Turbo Kit (MP-Biomedicals, Solon, OH, USA) (Plassart et al. 2012);
- Protocol III: adapted from the GnS-GII protocol (Genosol platform, INRAe, Dijon, France) (Internal reference: G.MO-026.6, 2020).

For each protocol, triplicates of soil and stem samples were performed for this study. The quantities of soil and stems were 1 g and 2 g, respectively. There was no prior preparation of the stems; the thawed 5-cm pieces were used directly.

Protocol I. The FastDNA™ SPIN kit for soil is used to obtain genomic DNA directly from soil samples. This commercial kit can be used on plant and animal tissues, bacteria, algae and fungi, and other members of a soil population. Each sample is homogenized in the presence of a lysis matrix in a Fast-prep-24™ shaking system commercialized by MP-Biomedicals (Irvine, CA, USA). It has been reported

that the efficiency and yield of this shaking system are superior to the results of traditional extraction methods, such as enzymatic digestion, sonication, and vortex (Hemkemeyer et al. 2018). The procedure included chemical (244 µl MT buffer and 1956 µl of sodium phosphate buffer) and mechanical (mixture of 1 g beads: 100 mm silica; 1.4 mm ceramic beads and 4 mm glass beads) means. Both buffers are developed specifically to protect and solubilize nucleic acids and proteins during cell lysis with minimal RNA contamination. To perform both disruption and lysis of the membranes, each sample is placed in 15-ml Falcon tubes containing beads and buffers. This homogenization step is performed within the Fast-prepR-24™: 90 s at 4000 s⁻¹ shaking for soil samples and 60 s at 4000 s⁻¹ shaking for stems. The obtained suspension is then centrifuged at 7000 g for 5 min to pellet the soil, cell debris, and lysis matrix. The supernatant is then deproteinized in the presence of a protein precipitation solution and centrifuged at 14,000 g for 5 min. The recovered DNA (750 µl of supernatant) is then bound to a binding matrix suspension (silica-based), and the mixture is transferred to a spin filter for wash steps. The filtrate is removed, and the binding matrix is washed twice with a salt-ethanol wash solution to remove contaminants. Finally, the purified DNA is eluted with 50 µl of DNA elution solution buffer. This DNA suspension serves as a template for all analyses.

Protocol II. DNA extraction and purification are performed using the GnS-GII method. To allow mechanical lysis of the cells, each sample is shaken in 15-ml Falcon tubes containing a mixture of the beads described above and 4 ml of a lysis buffer by using the Fast-prep-24™ system. Chemical lysis is then performed by incubating each sample at 70°C and shaking at 300 rpm for 30 min. The lysis buffer is a mixture of 100 mM Tris-HCl, pH8; 100 mM EDTA,

pH8; and 100 mM NaCl 2% and SDS 2%. Tris-HCl stabilizes the pH and prevents DNA degradation, EDTA (ethylenediaminetetraacetic acid) mainly chelates divalent cations (Mg^{2+}) as DNases cofactors, NaCl protects the extracted DNA from possible denaturation at 70°C and weakens DNA-protein interactions to facilitate protein precipitation, and SDS (sodium dodecyl sulfate) is involved in membrane lysis and protein denaturation. After centrifugation (7000 g, 5 min), the supernatant is recovered. Denatured proteins and other suspended particles are precipitated by the addition of 100 µl potassium acetate (3 M, pH5.5). After centrifugation (14,000 g, 5 min), the supernatant is recovered, and DNA is precipitated with isopropanol (-20 °C) and washed with 70% ethanol (-20 °C). After centrifugation (14,000 g, 5 min), the supernatant is removed, and the pellet is dried and resuspended in 100 µl of water.

DNA extracts are then purified by two different successive techniques: PVPP (polyvinylpolypyrrolidone) columns and then Micro Bio-Spin^R columns with Bio-Gel P-6. Firstly, DNA is purified with PVPP which is a high molecular weight polymer capable of trapping phenolic and alkaloid compounds present in soil (Holben et al. 1988). PVPP columns are prepared by adding 300 µl of the PVPP powder and washed with sterile qPCR water to remove impurities. Micro Bio-Spin^R columns are used according to the manufacturer's instructions. Micro Bio-Spin^R columns with Bio-Gel P-6 are small chromatography columns. Bio-Gel P-6 is a variety of polyacrylamide gel that divides molecules according to their size. Smaller molecules can pass through the gel and move more slowly through the column, but larger molecules are excluded and pass through the column more quickly. Nucleic acids are separated and purified using this method. Finally, the extracted DNA is repurified using the commercial GeneClean[®] Turbo kit according to the manufacturer's instructions. This kit uses a patented Turbo Binding Matrix that binds and removes contaminants (salts, detergents, and PCR inhibitors), from the DNA sample. The purified DNA can then be used for downstream applications such as PCR and sequencing. As a result of the purification step, a total volume of 30 µl of DNA is obtained which serves as a template for all analyses.

Protocol III. For cell lysis, each sample is shaken in the FastPrep^R-24 in 15-ml Falcon tubes containing 2 g beads (ceramic beads, silica beads, glass beads) and 5 ml lysis buffer (EDTA 100 mM, pH8; Tris 100 mM, pH8; NaCl 100 mM; and SDS 2%). Chemical lysis is performed by incubating the samples at 70 °C for 30 min with agitation at 300 rpm. Samples are then centrifuged (7000 g, 5 min), and 1 ml of the lysate is collected. Deproteinization is performed by adding 100 µl of potassium acetate (3M, pH 5.5) to the lysate, and the supernatant is collected after centrifugation (14,000 g, 5 min). For DNA precipitation, 900 µl isopropanol (-20 °C) is added and the samples are incubated at

-20 °C (1 h for soil and 30 min for stems). The supernatant is removed after centrifugation (13,000 rpm, 30 min). The DNA pellet is washed by adding 400 µL of 70% ethanol (-20 °C), centrifuged (13,000 rpm, 2 min), and dried at 60 °C (10–15 min). The crude DNA pellet is resuspended in 200 µl of ultrapure water after 4 h of storage at 4 °C. The DNA suspension is stored at -20 °C.

Purification of the extracted DNA is performed using the Nucleospin Soil Kit (Macherey Nagel, Düren, Germany) and is performed in two steps: (i) molecular sieving on a NucleoSpin Inhibitor Removal Column and (ii) DNA binding to a silica membrane: NucleoSpin Soil Column. Purification is performed according to the manufacturer's instructions. As a result of the purification step, a total volume of 80 µl of DNA is obtained, which serves as a template for all analyses.

Determination of the amount and purity of extracted DNA

The amount of DNA extracted from each sample was determined by absorbance at 260 nm by using Nanodrop technology. However, this technique has several drawbacks, including interference from impurities in the nucleic acid preparation (salts, ethanol, chloroform, proteins, ...) and the inability to distinguish double-stranded DNA from single-stranded nucleic acids (including RNA). Subsequently, the extracted DNA was also quantified by absorbance at 535 nm after PicoGreen staining (Quant-iT[™] PicoGreen[™] dsDNA Assay kit, Thermo Fisher, Bleiswijk, Netherlands). This accurate quantification was used to further determine the abundance of bacterial and fungal communities by real-time PCR.

The NanoDrop[™] spectrophotometer (Thermo Scientific, Wilmington, USA) was used to determine the purity of the DNA samples by measuring the absorbance of 1.5 µL of DNA at 280 and 230 nm and determining the ratios of the 260/280 nm and 260/230 nm values. The absorbance at 280 nm was mainly due to protein contamination, and the absorbance at 230 nm was explained by the presence of organic compounds, proteins, or chaotropic substances. In addition, the extracted DNA was visualized by electrophoresis on a 1% agarose gel. The DNA (10 µl) was stained with 3 µl of charge buffer and visualized by UV light and photographed. The intensities of the bands were compared with the Smart Ladder DNA marker (Eurogentec, Seraing, Belgium) using Quantum ST4 software (Vilber, Marne-la-Vallée, France).

Fungal and bacterial population density levels

Quantitative real-time PCR was used to determine the abundance of bacteria and fungi in each sample by examining the 16S rDNA (bacteria) and 18S rDNA (filamentous fungi)

genes. All DNA extracts were amplified in triplicate. The efficiency of the qPCR performed was above 90% for all samples.

Fungal 18S rDNA

The primers used for targeting the fungal community are nu-SSU-0817-F (5'-TTAGCATGGAATAATRRAATAGG A-3', with R = A or G, melting temperature (T_m) 54.2 °C) and nu-SSU-1196-R (5'-TCTGGACCTGGT GAGTTTC-3', T_m 56.7 °C) (Borneman and Hartin 2000) (Supplemental Table S1). The Hot firepol EvaGreen kit (Solis BioDyne, Tartu, Estonia) was used in the Corbett Research Rotorgene 6000 thermal cycler (QIAGEN, Hilden, Germany) to perform quantifications according to the following program: initial activation at 95 °C for 15 min followed by 48 cycles of denaturation at 95 °C for 18 s, hybridization at 60 °C for 60 s, and extension at 72 °C for 20 s. A PCR product of 422 bp is expected.

Bacterial 16S rDNA

Primers used to target the bacterial community are BAC338F (5'-ACTCCTACGGGAGGCAG-3', T_m 57.6 °C), BAC805R (5'-GACTACCAGGGTATCTAATCC-3', T_m 57.9 °C) and the probe BAC516F (FAM-TGCCAG CAGCCGCGTAATAC-TAM) (Yu et al. 2005) (Supplemental Table S1). The qPCR was performed using the GoTaq Probe qPCR Master mix kit (Promega, Madison, USA) in a Corbett Research Rotorgene 6000 thermal cycler (QIAGEN, Hilden, Germany) according to the following program: initial activation at 95 °C for 4 min followed by 50 cycles of denaturation at 95 °C for 15 s and hybridization-extension at 60 °C for 60 s. A PCR product of 468 bp is expected.

Metabarcoding approach

For the bacterial population, the primers BAC-F (5'-ACTCCT ACGGGAGGCAG-3') and BAC-R (5'-GACTACCAGGGT ATCTAATCC-3') (Yu et al. 2005) were used. For the fungal population, the primers SSU-F (5'-TTAGCATGGAATAAT RRAATAGGA-3' with R = A or G) and SSU-R (5'-TCTGGA CCTGGT GAGTTTC-3') (Borneman and Hartin 2000) were selected.

Microbial diversity was determined for each sample by using MinION Mk1C[®] and the RPB004 Rapid PCR Barcoding Kit (Oxford Nanopore Technologies, Oxford, UK). One to 5 ng of the DNA sample was mixed with 1 µl of the fragmentation MIX. Samples were then incubated at 30 °C for 1 min followed by incubation at 80 °C for 1 min. The PCR reaction MIX was prepared with 2 µl nuclease-free water, 4

µl labeled DNA sample, 1 µl rapid barcode primer, and 25 µl LongAmp Tap 2× Master Mix.

PCR was conducted under the following conditions: 1 cycle of 95 °C for 3 min, 14 cycles of 95 °C for 15 s, 14 cycles of 56 °C for 15 s, 14 cycles of 65 °C for 6 min, followed by 1 cycle of 6 min at 65 °C. Samples were then purified by using the QIAquick PCR Purification Kit[®] (QIAGEN, Hilden, Germany), and elution was performed in 30 µl ultrapure water. The concentration of PCR products was measured with the Qubit[®] system using 10 µl of the eluate. Qubit is a DNA and RNA quantification system (Thermo Fisher Scientific, Waltham, MA). It uses fluorescence to accurately measure the concentration of the sample before subsequent steps such as sequencing and PCR. The Qubit system is more accurate and sensitive than other quantification methods, such as UV spectrophotometry. Adaptor ligation was performed by adding 2 µl of rapid adaptor plasmid (RAP) to the sample pool, which was then incubated for 5 min at room temperature. The RAP system (Oxford Nanopore Technologies, Oxford, UK) contains short DNA adapter sequences, which in turn contain specific recognition sites for the DNA fragments present in the sample pool.

Loading of samples into the sequencing cell was performed according to the following procedure: 34 µl of sequencing buffer, 25.5 µl of charging bead, 4.5 µl of nuclease-free water, and 26 µl of sample pool to be sequenced were added to a nuclease-free tube. The sequencing MIX (total volume of 90 µl) was then applied to the charging cell. Sequencing was performed over 24 h. Metabarcoding was performed with one replicate for each sample.

Sequencing results were analyzed using Galaxy[™] software (Sloggett et al. 2013) and EPI2ME[™] software Fastq 16S (Cuscó et al. 2019) Workflow (version 2021.09.09) (Nanopore, Oxford, UK) for bacterial community analysis and Fastq WIMP (version 2021.11.26) (Nanopore, Oxford, UK) Workflow for fungal community analysis to determine the microbial community structure of soil (S-R0) and stem samples (F75-R0 and F75-R4). The sequencing raw data were deposited into the NCBI Sequence Read Archive (SRA) database (Accession Number: PRJNA911961).

Alpha diversity analysis was performed using the Shannon diversity index and Pielou evenness index. The Shannon index was calculated according to the equation: $H' = -\sum ((ni/N) \times \log_2(ni/N))$. Pielou index was calculated according to the equation: $J = H'/H_{max} = H'/\log_2 S$. Analysis of the Bray–Curtis similarity matrix was performed using the Past 4.03 software (Hammer et al. 2001).

Statistical analysis

Statistical evaluation of the results was performed for DNA quantification, DNA purity, and bacterial and fungal abundances obtained from the different procedures (I, II, and III)

using a one-way test ANOVA. This test was used to examine the significance of differences between and within different samples and protocols, and a p -value ≤ 0.05 was considered statistically significant.

Results

DNA yield

The total yields of purified genomic DNA were measured spectrophotometrically (Quant-iT™ PicoGreen™ dsDNA Assay kit). Figure 2 presents DNA concentrations ($\mu\text{g g}^{-1}$ dry matter). Although the evaluated protocols were all based on direct lysis, significant differences between protocols considered two by two ($p < 0.05$) were observed. The amounts of DNA extracted with protocol III were significantly higher than those obtained with protocols I and II.

For soil samples (S-R0), the highest DNA yield was obtained with protocol III with an average of $13 \pm 0.29 \mu\text{g/g}$ dry matter. Protocols I and II resulted in significantly lower DNA concentrations at $1.58 \pm 0.69 \mu\text{g/g}$ dry matter and $0.06 \pm 0.03 \mu\text{g/g}$ dry matter, respectively.

For stem samples, based on protocol III, the DNA amounts extracted from Futura75 stems were higher (F75-R0: $1.6 \pm 0.22 \mu\text{g/g}$ dry matter; F75-R4: $1.14 \pm 0.73 \mu\text{g/g}$ dry matter) than those based on the protocol I (F75-R0: $0.05 \pm 0.00 \mu\text{g/g}$ dry matter; F75-R4: $0.05 \pm 0.01 \mu\text{g/g}$ dry matter) and protocol II (F75-R0: $0.07 \pm 0.01 \mu\text{g/g}$ dry matter; F75-R4: $0.11 \pm 0.01 \mu\text{g/g}$ dry matter) with significant

differences ($p < 0.05$). Regarding DNA extracted from the dioecious variety Finola, it can be noted that the DNA amount extracted according to protocol III was greater ($6.17 \pm 0.66 \mu\text{g/g}$ dry matter) than that extracted according to the protocol I ($0.12 \pm 0.02 \mu\text{g/g}$ dry matter) and protocol II ($0.63 \pm 0.29 \mu\text{g/g}$ dry matter) (Fig. 2).

Impact of moisture content on the amount of DNA extracted from stems

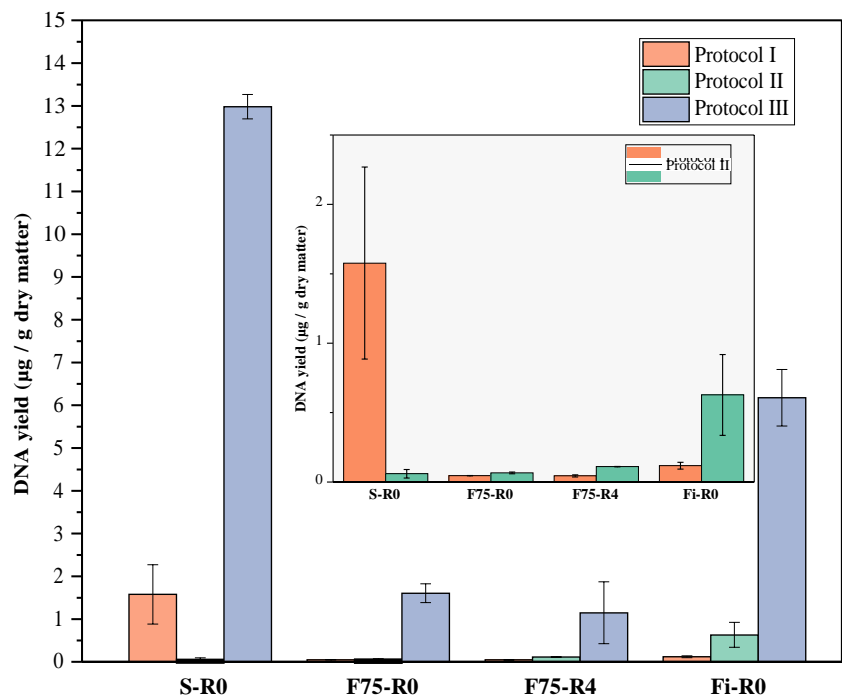
The influence of factors such as stem moisture on the amount of extracted DNA has been scarcely studied. In this study, stems provided from the same variety at the different retting degrees (unretted, 2-week, and 4-week retted stems) were chosen.

Based on protocol III, DNA amounts extracted from F75-R0 stems ($1.6 \pm 0.22 \mu\text{g/g}$ dry matter) were close to those extracted from F75-R2 stems ($2.3 \pm 0.4 \mu\text{g/g}$ dry matter) and F75-R4 stems ($1.1 \pm 0.7 \mu\text{g/g}$ dry matter) ($p > 0.05$) (Fig. 3). The different moisture contents of each stem sample seem not influencing the performance of the DNA extraction method. Nevertheless, this result should be interpreted with caution as the retting degree of the used stems differs (not retted, retted for 2 and 4 weeks).

DNA purity

Table 1 shows the purity levels ($A_{260/280}$ and $A_{260/230}$) of the DNA samples extracted from the soil (S-R0) and stem

Fig. 2 DNA yields extracted from soil (S-R0) and stem samples (F75-R0, F75-R4, and Fi-R0) for each extraction method (I, II, and III). Bars correspond to averages of three replicates. Fi-R0 samples, unretted Finola stems; F75-R0 samples, unretted Futura 75 stems; F75-R4 samples, Futura 75 stems retted for 4 weeks. Protocol I, commercial FastDNA™ Spin Kit for soil; Protocol II, Gns-GII procedure; Protocol III, custom procedure from the Genosol platform



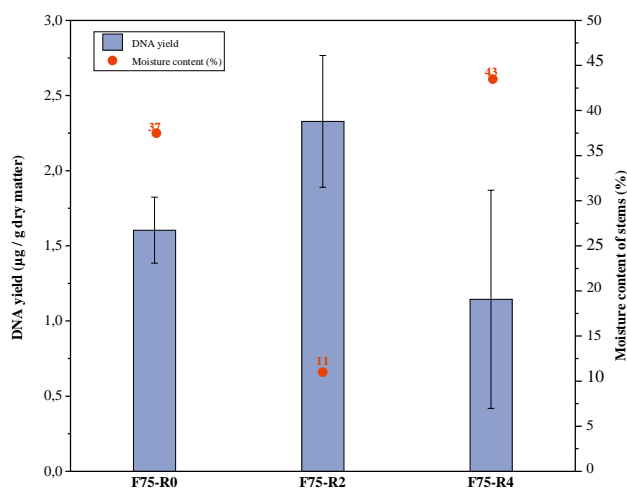


Fig. 3 Comparison of the amount of DNA recovered from Futura75 stem samples using Protocol III with different moisture content: F75-R0 (37% of humidity); F75-R2 (11% of humidity) and F75-R4 (43% of humidity). Bars correspond to averages of three replicates. R0, unretted stems; R2, retted for 2 weeks; R4, retted for 4 weeks

samples: unretted Finola stems (Fi-R0), unretted Futura 75 stems (F75-R0), and Futura 75 stems retted for 4 weeks (F75-R4).

$A_{260/280}$ ratio

DNA purity, estimated through the ratio $A_{260/280}$, was determined using NanoDrop® to verify the quality of purified extracted DNA (Table 1). The values of the $A_{260/280}$ ratio were similar ($p > 0.05$) regardless of the origin of the samples (soil or stems) and the DNA recovery procedure, except for samples F75-R0 and F75-R4. For soil samples, $A_{260/280}$ ratio values were within the acceptable range for all three protocols and vary between 1.62 and 1.90. For stem samples extracted with protocols I and II, $A_{260/280}$ ratios were slightly above 2.00 which may indicate low contamination with RNA

Table 1 $A_{260/280}$ and $A_{260/230}$ averages for the three replicates of DNA extracted from both soil (S-R0) and stem (Fi-R0, F75-R0, and F75-R4) samples using each of the extraction methods (I, II, and III). Fi-R0 samples, unretted Finola stems; F75-R0 samples, unretted Futura 75 stems; F75-R4 samples, Futura 75 stems retted for 4 weeks. Protocol I, commercial FastDNA™ Spin Kit for soil; Protocol II, Gns-GII procedure; Protocol III, custom procedure from the Genosol platform

Sample*	Protocol					
	$A_{260/280}$ ratio			$A_{260/230}$ ratio		
	I	II	III	I	II	III
S-R0	1.90 ± 0.04	1.62 ± 0.13	1.89 ± 0.01	0.10 ± 0.01	0.05 ± 0.03	1.77 ± 0.05
F75-R0	2.13 ± 0.23	2.50 ± 0.00	1.81 ± 0.03	0.01 ± 0.01	0.07 ± 0.01	1.05 ± 0.05
F75-R4	1.94 ± 0.05	2.20 ± 0.05	1.80 ± 0.03	0.03 ± 0.01	0.03 ± 0.01	0.81 ± 0.19
Fi-R0	2.34 ± 0.10	2.03 ± 0.42	1.80 ± 0.01	0.01 ± 0.00	0.08 ± 0.09	1.33 ± 0.12

* $n = 3$ replicates per 4 independent experiments

(Boiteux et al. 1999). In contrast, for stem samples extracted using protocol III, the ratios were stable at 1.80.

$A_{260/230}$ ratio

The $A_{260/230}$ ratio was also used as a secondary measure of nucleic acid purity. The $A_{260/230}$ values provide a measure of non-nucleic acid contamination (humic substances). This ratio varies considerably between the three extraction methods.

For soil samples, the $A_{260/230}$ ratio was greater (1.77 ± 0.05) when using protocol III than when using both protocols I and II (values of 0.10 or less). As for protocol III, the use of an appropriate bead mixture not only increases DNA yield but also decreases the amount of humic substances extracted.

DNA extracted from plants tends to contain impurities from polysaccharides and polyphenols that can be detected at 230 nm (Varma et al. 2007). For all samples, the $A_{260/230}$ ratio was greater with protocol III (0.81-1.33) than with both protocols I and II (values of 0.10 or less).

Gel electrophoresis

The yield, purity, and quality of extracted DNA are important parameters to check before PCR amplification. Degraded DNA may not be suitable for PCR amplification, because the formation of chimeric PCR products is greater with a fragmented template (Cullen and Hirsch 1998).

Agarose gel electrophoresis of DNA stem samples extracted with the protocol I shows that the DNA was highly degraded and/or fragmented. For protocol II, the extracted DNA samples were barely detectable in electrophoresis due to the low yield of DNA (data not shown).

In contrast, DNA samples extracted using protocol III were of good quality. No DNA degradation was observed in the soil samples (S-R0), and moderate degradation was

observed in the stem samples (F75-R0, F75-R4, and Fi-R0) (Fig. 4).

Fungal and bacterial population densities

To further evaluate the DNA extraction methods, the abundance of total bacterial and fungal communities from the studied samples (soil and stems) has been estimated by quantifying the 16S and 18S rDNA genes by qPCR. Low values of the $A_{260/230}$ ratio did not affect qPCR amplification for

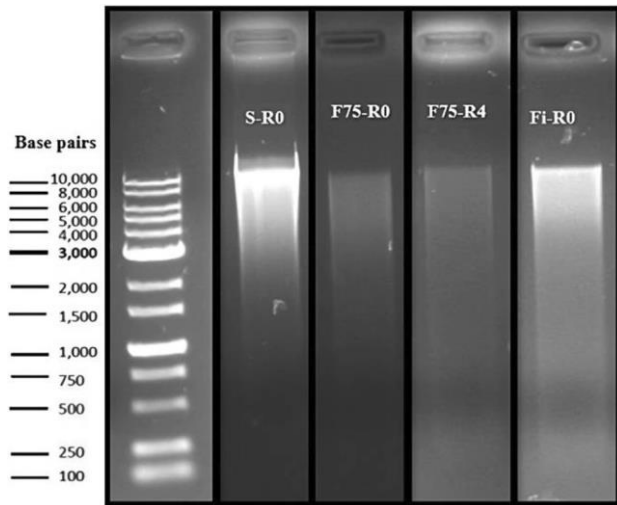


Fig. 4 Electrophoresis gel (1% agarose gel in 1× TAE buffer) of genomic DNA extracts of Protocol III (custom procedure from the Genosol platform) obtained from soil (S-R0) and stem samples (F75-R0, F75-R4, and Fi-R0). Fi-R0 samples, unretted Finola stems; F75-R0 samples, unretted Futura 75 stems; F75-R4 samples, Futura 75 stems retted for 4 weeks

both protocols I and II. Amplification was successful for all samples regardless of the DNA extraction method. However, the highest abundance values were obtained with protocol III (Fig. 5).

Fungal community levels

DNA extracted from all samples and with all three protocols was successfully amplified (Fig. 5a). For the soil samples, the copy number of DNA extracted with the protocol I and III was higher (8.9×10^7 and 9.1×10^6 copies/g dry matter, respectively) than DNA extracted with protocol II (3.3×10^4 copies/g dry matter). However, for all stem samples (F75-R0, F75-R4, and Fi-R0), copy number values were higher for the DNA template extracted with protocol III than for both protocols I and II (2.5–4 log difference).

Bacterial community levels

DNA extracted from all samples and with all three protocols was successfully amplified (Fig. 5b). For all samples (soil and stems), the copy number values were higher by using protocol III than protocols I and II. Similar abundances were observed for both protocol I and II regardless of sample type ($p > 0.05$). For protocol III, the copy number values of the stem samples range from 9.3×10^7 to 6.7×10^8 16SrDNA copies/g dry matter.

High-throughput sequencing quality

Characterization of microbial communities can be affected by the method used to obtain metagenomic DNA (Terrat et al. 2015). Therefore, it is important to test the efficacy of

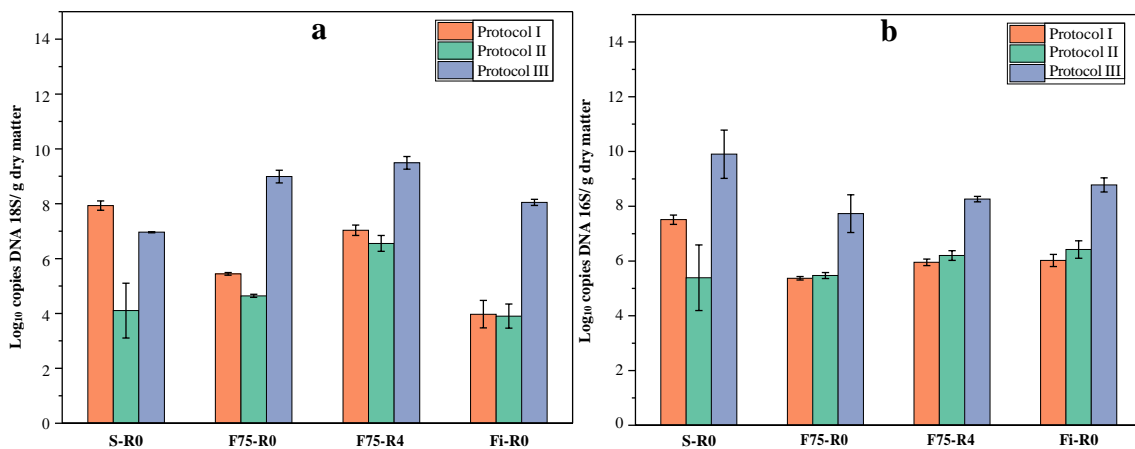


Fig. 5 Microbial abundances of soil and stem samples according to three different DNA extraction procedures (I, II, and III). Quantification of **a** 18S rDNA genes and **b** 16S rDNA genes of soil (S-R0) and stems samples (F75-R0, F75-R4, and Fi-R0). Bars correspond to averages of three replicates. Fi-R0 samples, unretted Finola stems;

F75-R0 samples, unretted Futura 75 stems; F75-R4 samples, Futura 75 stems retted for 4 weeks. Protocol I, commercial FastDNA™ Spin Kit for soil; Protocol II, Gns-GII procedure; Protocol III, custom procedure from the Genosol platform

DNA extraction procedures in terms of the representativeness of bacterial and fungal communities.

For each sample, the sequencing quality score, the total number of analyzed reads, the average size of sequences, and the percentage of classified sequences were determined. The overall quality score was 9.97 for bacterial samples and 10.2 for fungal samples. The analyzed reads were 20357 for bacterial samples and 29 452 for fungal samples.

The sequencing results for all samples exhibit an average quality control (QC) of 10, which is still a suitable quality for Oxford Technologies Nanopore sequencing. The average sequence size obtained for the fungal population was approximately 417 bp, then giving a read percentage of 98.8% (the expected sequence length was 422 bp). The average size of the sequences obtained for the bacterial population was 395 bp, then giving a read percentage of 84.4% (the expected sequence is 468 bp). Based on this quality control, the sequencing procedure was considered feasible.

Microbial community structure

The amount of DNA extracted using protocol II was not sufficient to perform the sequencing step. Therefore, sequencing was performed on the soil and stem samples extracted using both protocols I and III. Moreover, a sequencing issue was identified with sample F75-R0. Therefore, the fungal community present in this sample is not represented in this study.

Fungal community structure

The same fungal phyla were present in similar abundance in both stem and soil samples (Fig. 6a). Regardless of the protocol (I or III), a clear dominance of the *Ascomycota* was observed. For soil samples (S-R0-I and S-R0-III), the phyla were *Ascomycota* ($88.2 \pm 2.1\%$) and *Basidiomycota* ($2.1 \pm 0.3\%$). For stem samples (F75-R4-I and F75-R4-III), the phyla were also *Ascomycota* ($86.6 \pm 3.2\%$) and *Basidiomycota* ($4.5 \pm 2.3\%$).

Moreover, a similar fungal structure (genus level) was observed for the soil on the one hand and the stem on the other, regardless of the protocol used (Fig. 6b). For soil samples (S-R0-I and S-R0-III), the most represented genera were *Botrytis* ($44.9 \pm 11.2\%$), *Talaromyces* ($12.9 \pm 2.8\%$), *Brettanomyces* ($8.0 \pm 1.3\%$), *Zygorulaspota* ($6.7 \pm 0.9\%$), and *Sugiyamaella* ($6.2 \pm 1.9\%$). For retted stems (F75-R4), a close fungal structure was observed: *Botrytis* ($48.3 \pm 4.0\%$), *Talaromyces* ($13.1 \pm 3.5\%$), *Brettanomyces* ($13.1 \pm 3.5\%$), *Zygorulaspota* ($5.5 \pm 1.8\%$), and *Malassezia* ($6.2 \pm 2.8\%$).

Bacterial community structure

Community diversity was higher in bacterial samples when compared to fungal samples. For soil samples (S-R0-I and S-R0-III), the most abundant bacterial phyla were *Actinobacteria* ($46.5 \pm 6.1\%$), *Proteobacteria* ($34.2 \pm 7.1\%$), and *Bacteroidetes* ($8.3 \pm 0.8\%$) (Fig. 7a), with a co-dominance of *Actinobacteria* and *Proteobacteria*.

For unretted stem samples (F75-R0-I and F75-R0-III), the most abundant phyla were also *Proteobacteria* ($86.8 \pm 1.3\%$), *Actinobacteria* ($6.6 \pm 1.4\%$) and *Bacteroidetes* ($1.9 \pm 0.1\%$) with a dominance of *Proteobacteria*. For retted stem samples (F75-R4-I and F75-R4-III), the most abundant bacterial phyla were *Proteobacteria* ($68.0 \pm 0.9\%$), *Bacteroidetes* ($20.0 \pm 5.1\%$), and *Actinobacteria* ($8.8 \pm 5.6\%$) with a dominance of *Proteobacteria*.

The same evolution between the non-retted stems (F75-R0-I and F75-R0-III) and the stems retted for 4 weeks (F75-R4-I and F75-R4-III) was observed regardless of the protocol. An increase of *Bacteroidetes* from $1.9 \pm 0.1\%$ for F75-R0 to $20.0 \pm 5.1\%$ for F75-R4 and a decrease of *Proteobacteria* from $86.8 \pm 1.3\%$ for F75-R0 to $68.0 \pm 0.9\%$ for F75-R4 were noticed.

For soil samples (S-R0-I and S-R0-III), the same genera were present with similar relative abundances, regardless of protocol (Fig. 7b). The most represented genera were *Rubrobacter* ($40.7 \pm 6.2\%$) and *Sphingomonas* ($5.8 \pm 3.8\%$).

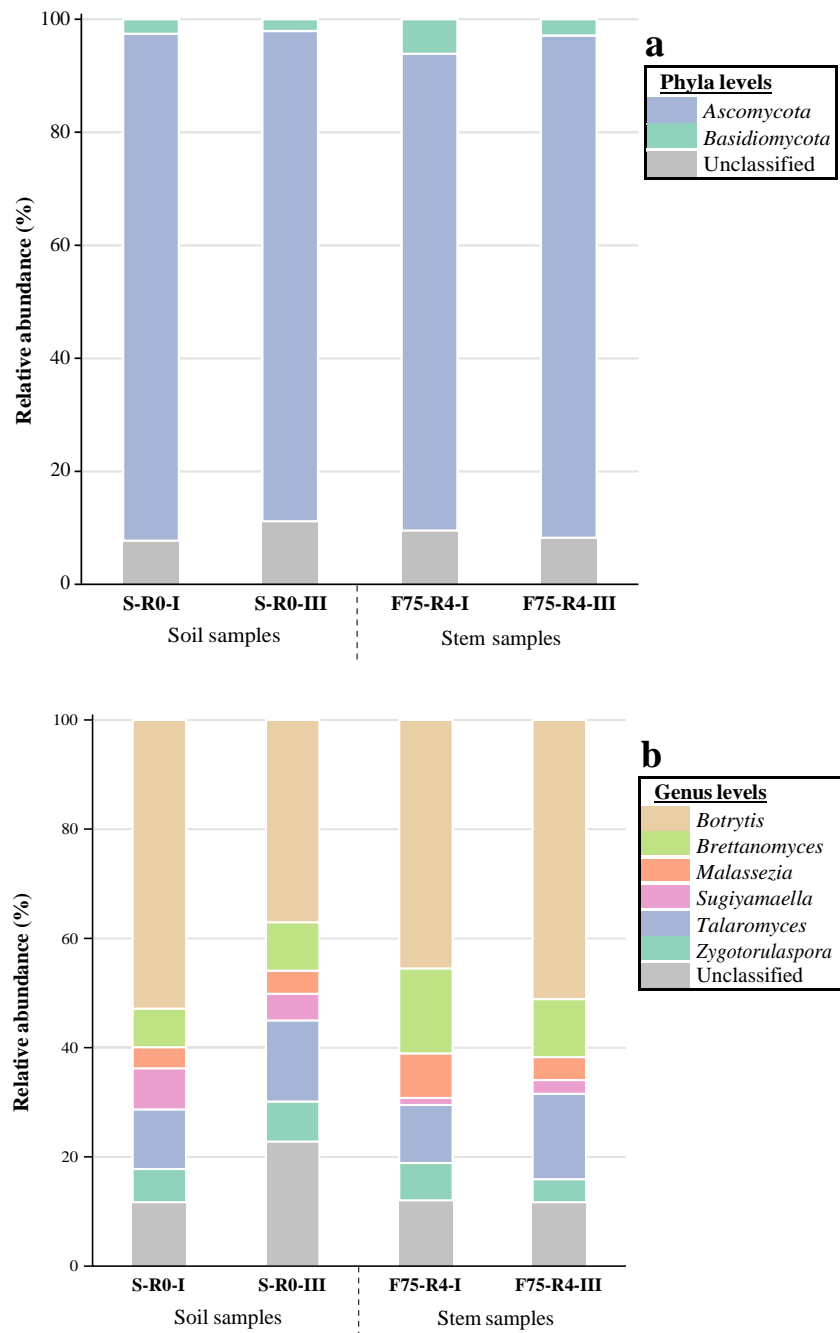
However, slight differences in the relative abundance of bacterial genera were observed in the stem samples depending on the used extraction protocol. For stem samples, the most represented genera were *Pseudomonas* ($29.4 \pm 16.1\%$), *Sphingomonas* ($24.0 \pm 11.3\%$), *Pantoea* ($16.0 \pm 2.2\%$), and *Methylobacterium* ($15.9 \pm 5.0\%$) for F75-R0 and *Sphingomonas* ($27.5 \pm 9.4\%$), *Massilia* ($13.5 \pm 8.1\%$), *Pedobacter* ($12.6 \pm 5.2\%$), *Pseudomonas* ($10.9 \pm 3.5\%$), *Pantoea* ($7.1 \pm 0.2\%$), and *Chryseobacterium* ($6.6 \pm 3.5\%$) for F75-R4.

Regardless of the protocol, the same evolution was observed between the non-retted stems (F75-R0-I and F75-R0-III) and the stems retted for 4 weeks (F75-R4-I and F75-R4-III). A decrease in *Pseudomonas*, *Pantoea*, and *Methylobacterium* from $29.4 \pm 16.1\%$, $16.0 \pm 2.2\%$, and $15.9 \pm 5.0\%$ for F75-R0 to $10.9 \pm 3.5\%$, $7.1 \pm 0.2\%$, and $2.7 \pm 2.4\%$ for F75-R4 respectively was observed. In addition, an increase of *Massilia* from $3.8 \pm 1.6\%$ for F75-R0 to $13.5 \pm 8.1\%$ for F75-R4 was observed. We also notice an appearance of new bacterial genera for F75-R4 compared to F75-R0 which were *Pedobacter* ($12.6 \pm 5.2\%$), *Chryseobacterium* ($6.6 \pm 3.5\%$), and *Hymenobacter* ($2.1 \pm 0.5\%$).

Alpha diversity of fungal and bacterial communities

Regarding fungal communities of soil samples (S-R0-I and S-R0-III), the determination of alpha diversity, as measured

Fig. 6 a Normalized fungal relative abundance at the phyla level for soil (S-R0) and stem samples (F75-R4). **b** Normalized fungal relative abundance at the genus level for soil (S-R0) and stem samples (F75-R4). The genera shown in these figures correspond to relative abundances \geq of 1%. F75-R4, Futura 75 stems retted for 4 weeks. Protocol I, commercial FastDNA™ Spin Kit for soil; Protocol III, custom procedure from the Genosol platform



with the Shannon index reveals similar results with both protocols I and III (1.87 for S-R0-I and 2.13 for S-R0-III) (Fig. 8a). For stem samples (F75-R4), similar results were observed with both protocols I and III (2.00 for F75-R4-I and 1.83 for F75-R4-III).

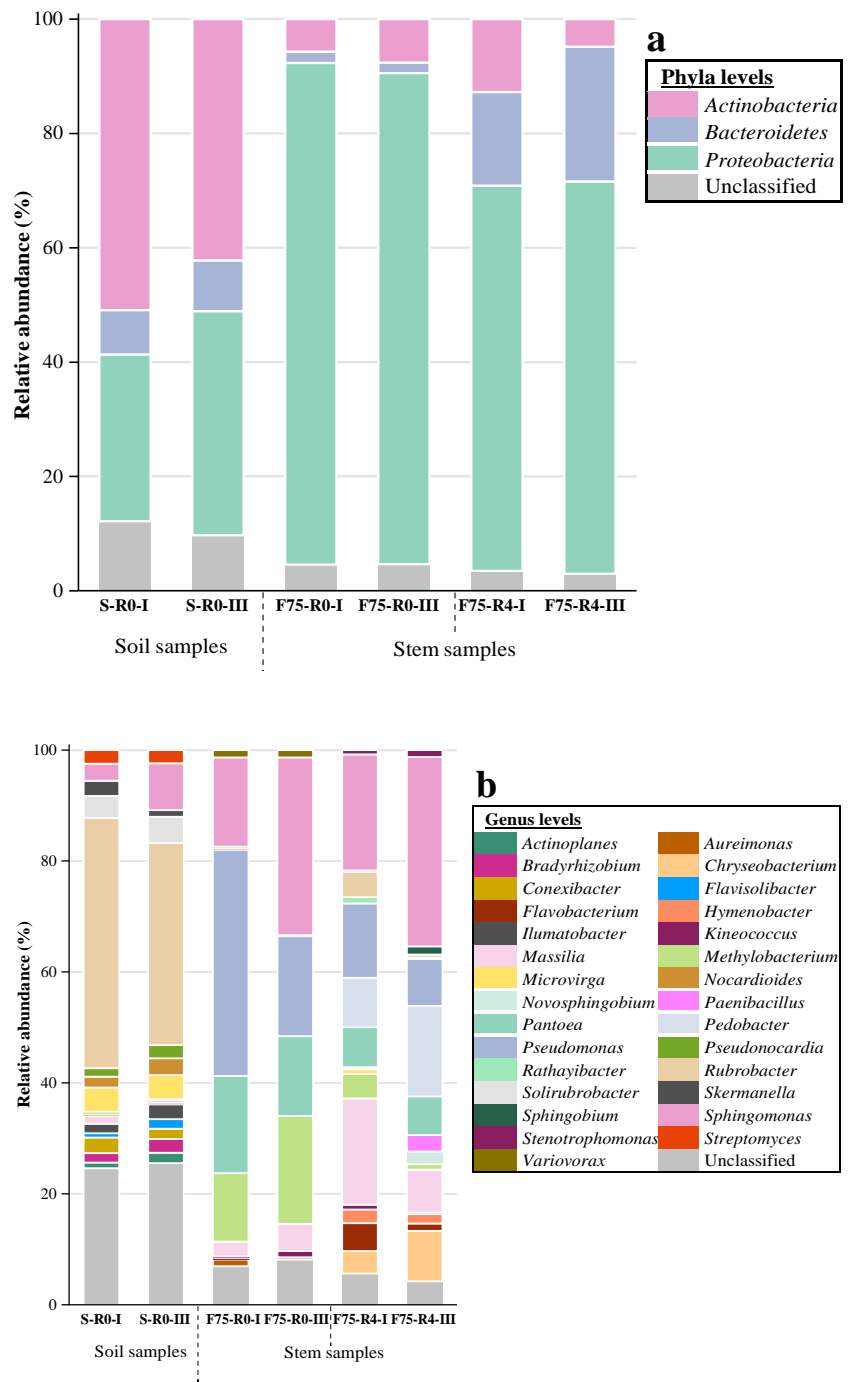
Similar evenness indices were observed regardless of the DNA extraction procedure ($0.20 \pm 0.10\%$ for S-R0 and $0.16 \pm 0.00\%$ for F75-R4) (Supplemental Table S2).

Shannon diversity index was more important for bacterial communities compared to fungal communities. Similar alpha diversity indices of bacterial communities were observed for

soil and stem samples extracted using protocols I and III. For soil samples, similar results were observed with both I and III (2.71 for S-R0-I and 3.17 for S-R0-III) (Fig. 8b). For stem samples, similar results were also observed with both protocols I and III for F75-R0 (2.40 for F75-R0-I and 2.51 for F75-R0-III) and F75-R4 samples (3.48 for F75-R4-I and 3.16 for F75-R4-III).

Slight differences in evenness indices were observed whatever the DNA extraction procedure for soil samples ($0.32 \pm 0.04\%$) and for stem samples ($0.22 \pm 0.01\%$ for F75-R0 and $0.31 \pm 0.05\%$ for F75-R4) (Supplemental Table S2).

Fig. 7 a Normalized bacterial relative abundance at the phyla level for soil (S-R0) and stem samples (F75-R0 and F75-R4). **b** Normalized bacterial relative abundance at the phyla (A) and genus (B) levels for soil (S-R0) and stem samples (F75-R0 and F75-R4). The genera shown in these figures correspond to relative abundances \geq of 1%. F75-R0 samples, unretted Futura 75 stems; F75-R4 samples, Futura 75 stems retted for 4 weeks. Protocol I, commercial FastDNA™ Spin Kit for soil; Protocol III, custom procedure from the Genosol platform



Alpha diversity calculated using the Shannon index (richness) and Pielou index (evenness) indicates that both soil and stem samples analyzed with the two different protocols (I and III) demonstrate similar fungal and bacterial diversity profiles. It can be concluded that, in this study, the DNA extraction method (I and III) seems to have no effect on microbial diversity.

Microbiota similarity

Figure 9 shows the similarity index of Bray–Curtis among samples between both protocols (I and III), according to the bacterial taxonomic data (genus level at \geq 1%). The similarity results for bacterial communities have been only computed (32 bacterial genera). Similarity results between

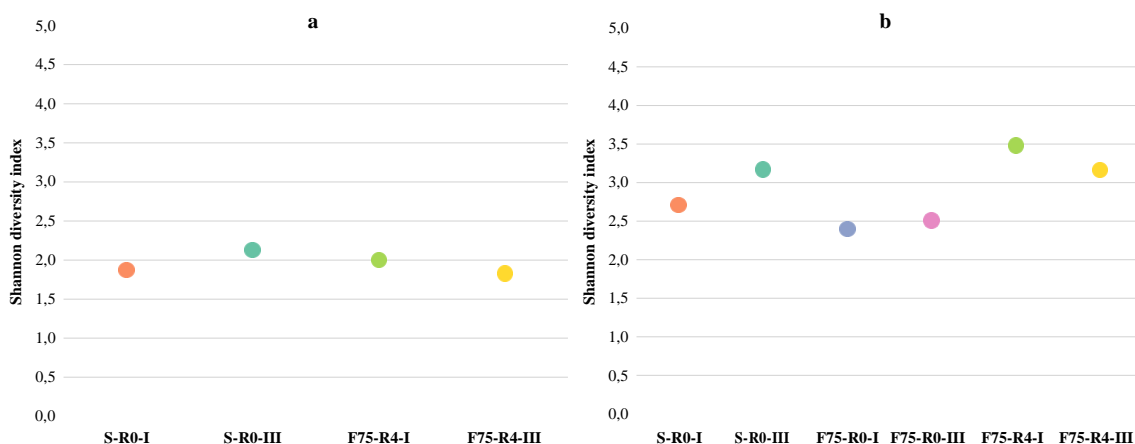


Fig. 8 Shannon diversity indices of fungal (a) and bacterial (b) communities of soil (S-R0) and stem (F75-R0 and F75-R4) samples extracted with two protocols (Protocol I and Protocol III). F75-R0

samples, unretted Futura 75 stems; F75-R4 samples, Futura 75 stems retted for 4 weeks. Protocol I, commercial FastDNA™ Spin Kit for soil; Protocol III, custom procedure from the Genosol platform

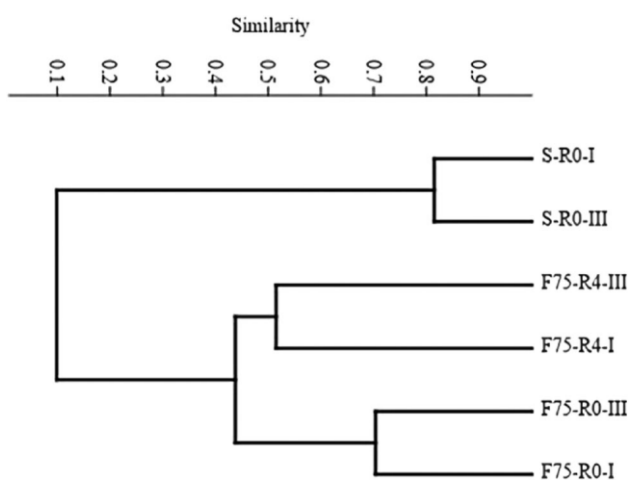


Fig. 9 Dendrogram constructed from similarity matrix (Bray-Curtis coefficient) between the bacterial communities of soil (S-R0) and stem samples (F75-R0 and F75-R4) extracted using Protocol I and Protocol III. F75-R0 samples, unretted Futura 75 stems; F75-R4 samples, Futura 75 stems retted for 4 weeks. Protocol I, commercial FastDNA™ Spin Kit for soil; Protocol III, custom procedure from the Genosol platform

fungal communities were not included because only a limited number of genera were identified (6 fungal genera). In addition, the number of classified fungal sequences was not the same for all samples.

Bray-Curtis beta-diversity of the bacterial microbiome exhibits similarity between soil (S-R0-I; S-R0-III) and stem (F75-R0-I; F75-R0-III, and F75-R4-I; F75-R4-III) samples. The maximal community similarity was noticed between soil samples (80%). However, for stem samples, the similarity between unretted stem samples F75-R0 was more important (70%) than retted stem samples F75-R4 (50%). This result

observed with the F75-R4 samples can be explained by the combination of the effects due to the retting process and the DNA extraction protocol. Retting has effectively a noticeable impact on the diversity of bacterial communities involved in this process (Djemiel et al. 2017). However, no convergence of retted stem profiles with soil ones was observed, regardless of protocols.

Discussion

DNA yield

According to the literature, DNA yield is highly variable and depends on different parameters such as the DNA extraction and purification procedure (Plassart et al. 2012; Wagner et al. 2015). The type of sample and its physicochemical properties, as well as the microbial communities, have also an impact on extracted DNA yields (Cabrol et al. 2010).

Soil samples

Some similarities are noted between all protocols (protocol I, II, and III), but differences are also observed between the recommended procedures. The composition of the lysis buffer (SDS + Tris + EDTA) is identical for all protocols. It should be noted that the cell lysis is performed with the same beating system (FastPrep-24™) and under the same conditions (speed and duration) for all protocols. However, for mechanical lysis, the main difference is in the quantity of beads. Protocol III contains twice as many beads (2 g) as protocols I and II (1 g). The volume of the lysis buffer differs from one protocol to the other.

The factors affecting the yield of extracted DNA are mainly the efficiency of the cell lysis steps (Mahmoudi et al. 2011). Consequently, the high amounts of DNA obtained with protocol III could be due to a more suitable mechanical lysis of the cells, which favored the desorption and mechanical lysis of the microorganisms and allowed to target a more microbial spectra (Cabrol et al. 2010). Moreover, several studies have shown that the use of a specific mixture of beads increased DNA yield (Gray and Herwig 1996; Carriag et al. 2007; Gobbi et al. 2019).

The amounts of DNA extracted using protocol III are consistent with values reported in the literature. Similar studies have shown that DNA yields obtained from soil samples vary between 2 and 50 $\mu\text{g/g}$ dry matter (Cullen and Hirsh 1998; Schneegurt et al. 2003; Petric et al. 2011; Plassart et al. 2012; Wüst et al. 2016; Dimitrov et al. 2017).

In addition, the observed differences in DNA yields could also be explained by the purification steps. Lower extraction yields and DNA degradation are common in the purification processes (Lim et al. 2017). Protocol II requires three purification steps for the extracted DNA: PVPP columns, Micro Bio-Spin columns, and the GeneClean® Turbo kit. To determine whether the use of the GeneClean® Turbo Kit leads to DNA losses, this kit has been used as if it was an additional purification step for protocol I. This step results in DNA losses of approximately 46% for soil samples and 73% for stem samples. It is worth noting that the purity levels ($A_{260/280}$ and $A_{260/230}$ ratios) are approximately the same before and after DNA purification (data not shown).

On the other hand, these different amounts of extracted DNA might have been caused by a higher capacity of the silica membrane used in the purification steps of protocol III, compared to the silica particles used in protocol I (Binding Matrix Suspension) (Knauth et al. 2013). Another reason for the low DNA extraction could be linked to the use of spin filters (protocol I) as the amount of eluted DNA could be reduced after purification in such tubes (Knauth et al. 2013).

The yield of extracted DNA from soil depends on soil type, pH, and organic matter, clay, and silt content, as these factors can affect either the growth of certain microbial taxa or the formation of aggregates that host microorganisms (Plassart et al. 2012). In our experiments, a clay soil was used. This is usually problematic because nucleic acids tend to adsorb to clay particles due to charge properties, which hinders DNA recovery (Högfors-Rönholm et al. 2018). Nevertheless, using protocol III, enough DNA can be extracted to perform various types of molecular biological analyses (DNA quantification using Quant-iT™ PicoGreen™ dsDNA Assay kit, gel electrophoresis, qPCR, high throughput sequencing).

To conclude, protocol III seems to be more adapted to obtaining great quantities of DNA from soil. A high-level DNA yield is required to ensure that the DNA sample is

representative of the sample gene pool (Desneux et al. 2014). Nevertheless, a larger amount of extracted DNA does not mean that a greater number of taxa or greater sequence diversity could be observed (Stach et al. 2001; Plassart et al. 2012).

Stem samples

Concerning plant molecular biology, it is difficult to obtain a genomic microbial DNA of high quality and purity from plant organs. The specific characteristics of plant stem especially fibrous ones (rigid fiber cell wall, presence of fibers, tissue quality, and biochemical composition ...) require special attention for efficient microbial DNA extraction (Varma et al. 2007). To date, only a limited number of protocols have been published for this purpose (Djemiel et al. 2017, Ribeiro et al. 2015).

Regarding the amounts of DNA extracted from stem samples using protocol III, the homogenization step in the presence of the beads may facilitate mechanical lysis of the cells (desorption of the microorganisms) and consequently penetration of the lysis buffer. The shear forces are enhanced by the quantity of beads (2 g). It has been shown that the purification steps can also have a significant effect on the DNA amount extracted from plant organs (Boiteux et al. 1999).

Influence of DNA extraction on DNA purity

DNA purity provides an indication of the presence of potential qPCR inhibitors such as humic substances, polysaccharides, proteins, and organic compounds that may be co-extracted with DNA and interfere with downstream PCR application (Le Maréchal et al. 2018). DNA samples extracted using protocol III have the highest purity for $A_{260/280}$ and $A_{260/230}$ ratios. $A_{260/280}$ ratio values between 1.8 and 2.2 reveal acceptable DNA quality for further characterization of microbial communities (density, structure).

The recommended $A_{260/230}$ ratio ranges from 2.0 to 2.2, and when this ratio is less than 1.5, it indicates the presence of a contaminant that absorbs at 230 nm. This contamination may be due to the presence of organic compounds, proteins, or chaotropic agents (EDTA, carbohydrates, phenol, guanidine).

Thus, the expected DNA quality is provided by protocol III. The purification kit combines specific buffers (SL3, SB, SW1, and SW2) and columns "NucleoSpin Inhibitor Removal Columns" to remove PCR inhibitors, which may be more effective than the procedures used for protocols I and II.

Influence of DNA extraction procedure on fungal and bacterial population densities

The results obtained for the two taxonomic groups (bacterial and fungal) indicate that the assessment of microbial abundance depends on the procedures used to extract genomic DNA. It can be concluded that protocol III is more efficient in extracting bacterial and fungal DNA from different types of hemp stems and soil samples than protocols I and II.

The $A_{260/230}$ ratio is greater for all samples (soil and stems) when using protocol III than when using both protocols, I and II (values of 0.1 or less). Protocol III seems to remove more PCR inhibitors, which may explain why this protocol is more effective than protocols I and II.

Influence of DNA extraction procedure on microbial community structure

It should be noted that the observed changes in microbial community structure in the F75-R4 samples can be due to both the retting process to which these stems were subjected and the DNA extraction protocol. The retting process has effectively an impact on the diversity and dynamics of bacterial and fungal communities involved in this process (Djemiel et al. 2017). In addition, in the case of DNA extraction from hemp stems, each sample is derived from 2 or 3 sections of two or three different stems. The biological variability associated with this procedure is definitely present in our study and may therefore affect the obtained results.

The two tested protocols (I and III) extract and amplify bacterial and fungal DNA with different efficiencies. Samples extracted using protocol I have acceptable $A_{260/280}$ purity and very low $A_{260/230}$ purity. Despite the lower purity ($A_{260/230}$) of DNA samples extracted using protocol I (0.01–0.10) compared to those extracted using protocol III (0.81–1.77), and the lower microbial density with the protocol I compared to protocol III, sequencing is successful and similar microbial structure is visualized for all samples (soil and stems) using both protocols I and III.

The capacity of presenting bacterial and fungal communities from plant samples is an important factor in choosing an appropriate DNA extraction method. Therefore, it can be concluded that both extraction methods provide a representative picture of both fungal and bacterial communities present in the soil and stem samples, regardless of their $A_{260/230}$ purity value and population density levels.

Influence of retting on fungal and bacterial community structures

Regardless of the protocol (I or III), the same fungal phyla (*Ascomycota* and *Basidiomycota*) are present in similar abundance in both stem (F75-R4) and soil (S-R0) samples

with a dominance of the *Ascomycota* phylum compared to *Basidiomycota*. These observations have already been shown in other studies concerning the field retting of flax (Djemiel et al. 2017) and hemp (Ribeiro et al. 2015; Liu et al. 2017). Several genera belonging to the phylum *Ascomycota* have shown an ability to dissociate fibers during the retting process in particular through a high level of pectinolytic activity (Henriksson et al. 1997). Some genera belonging to the phylum *Basidiomycota* are capable of secreting lignin-degrading enzymes of the fiber cell wall (Vasina et al. 2017; Yang et al. 2017).

Ascomycota is represented principally by the genus *Botrytis* for soil (S-R0) and stems samples (F75-R4). This genus has previously been identified as an agent of field retting of flax (Henriksson et al. 1997; Fila et al. 2001). *Talaromyces* is also among the most dominant genera in the soil (S-R0) and stem samples (F75-R4). This genus is capable of secreting enzymes that degrade fiber hemicelluloses (Méndez-Líter et al. 2021).

Regarding bacterial communities, the most abundant phyla are *Actinobacteria* for soil (S-R0) samples and *Proteobacteria* for stem (F75-R0 and F75-R4) samples. It has been shown that *Actinobacteria* is among the most dominant phyla in soil samples during field retting of flax (Djemiel et al. 2017). The dominance of *Proteobacteria* in stem samples has been also demonstrated in other studies concerning the field retting of flax and hemp (Ribeiro et al. 2015; Djemiel et al. 2017; Liu et al. 2017), hemp retting in a controlled environment (Law et al. 2020), and water retting of jute (Munshi and Chattoo 2008). Some studies (Djemiel et al. 2017; Liu et al. 2017; Law et al. 2020) have shown the same observations for stem samples (F75-R0 compared to F75-R4) in terms of the increase in the relative abundance of *Bacteroidetes* phylum and the decrease in *Proteobacteria*.

For stem samples genera, results underline the abundance of *Sphingomonas* and *Pseudomonas*. *Sphingomonas* genus has been previously identified during hemp, flax, and bamboo retting and is known for its ability to degrade pectin (Fu et al. 2011; Ribeiro et al. 2015; Djemiel et al. 2017). *Pseudomonas* genus has also previously been associated with the retting of various plants (flax, hemp, jute). It is known for its ability to secrete lignin and pectin-degrading enzymes and facilitates the decohesion of the fibres (Betrahet and Bhat 1958; Ribeiro et al. 2015; Zhao et al. 2016; Djemiel et al. 2017; Liu et al. 2017; Law et al. 2020). The decrease in the relative abundance of *Pseudomonas* for F75-R4 has been observed in a study of hemp retting (Liu et al. 2017). The appearance of new bacterial genera for F75-R4 compared to F75-R0 (*Pedobacter*, *Chryseobacterium*, and *Hymenobacter*) is demonstrated in other studies (Ribeiro et al. 2015; Zhao et al. 2016; Liu et al. 2017; Moawad et al. 2019; Law et al. 2020).

How to choose the most suitable DNA extraction procedure from plant stems

In conclusion, the microbiology of field retting is poorly investigated. The procedure used for genomic DNA extraction from stems can have a significant impact on molecular analyzes of microbial communities and is therefore an important consideration for microbial studies. This work does not aim to give a single general protocol for optimal DNA recovery from the biofilm colonizing vegetal stems but presents a systematic guideline to compare desorption, extraction, and purification methods, assessing their impact on quantity, quality, and representativity of DNA samples. Thus, an experimental strategy is proposed to compare three DNA extraction methods, based on criteria ranging from DNA yield, DNA purity, bacterial and fungal density levels, and diversity representation. In addition to these parameters, other factors such as processing time should be considered when selecting an appropriate extraction method (Supplemental Table S3).

Our study reveals that protocol III seems to be a suitable procedure to obtain the highest yield and quality of extracted DNA from hemp stem samples. The DNA quality is sufficient to perform downstream DNA analysis techniques such as qPCR and high throughput sequencing. The quantity of beads seems to favor DNA extraction yield. A higher DNA yield is obtained when using protocol III in the presence of 2 g of beads instead of 1 g for protocols I and II. Moreover, the purification protocol may also have a significant effect on the quality of the extracted DNA. In this study, the purification step with protocol III which combines molecular sieving (inhibitors removal) and silica membrane (DNA fixation) results in increasing the DNA quality compared to the other two protocols. Therefore, bacterial and fungal densities appear to be higher when using protocol III. Nevertheless, this procedure is time-consuming.

The amount of DNA extracted using protocol II is not sufficient to provide the structure of bacterial and fungal communities. Therefore, this limited amount of extracted DNA can be considered a limiting factor for the subsequent analysis steps (qPCR and sequencing). In return, the DNA yields of the two protocols I and III are sufficient for sequencing analysis.

DNA yield is not generally considered a selection criterion for an extraction method and no clear correlation between DNA yield and the representation of microbial diversity has been demonstrated (Xue et al. 2018). In our study, the ability to determine the evolution of bacterial and fungal communities over time from plant samples is an important factor in the choice of DNA extraction method. Both protocols I and III provide a similar representativity of bacterial and fungal communities present in soil and stem samples. All tested Alpha and Beta diversity indices,

including the Shannon diversity, Pielou evenness, and Bray–Curtis similarity indexes, revealed similar results for both protocols (I and III).

Therefore, protocols I and III are recommended for achieving a fair representation of bacterial and fungal communities from plant stems (hemp in the case of our study). They can therefore be adapted according to the objectives of the study and the type of plants.

This study shows that DNA extraction from a complex sample cannot provide an exhaustive overview of the microbial community. The critical importance to consider when evaluating biases associated with DNA extraction has been highlighted when addressing microbial ecology in the retting process. This integrated approach should be adapted to each variety of stem samples, as DNA recovery is strongly influenced by the specificity of the matrix. The extraction step is just one of several possible biases that occur in molecular analysis, from the sampling itself to the PCR amplification or metabarcoding approach.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00253-023-12582-5>.

Author contribution AB, JCB, LM, and SB designed and supervised the research. Material preparation, data collection, and experiments were performed by EBO, JR, and LM. CS and SS were responsible for the sequencing experiment and analysis. EBO performed the formal analysis and writing of the manuscript (original draft — review and editing). All authors read and approved the final manuscript.

Funding This research received a Région Occitanie Pyrénées-Méditerranée funding (ALDOCT-000877).

Data availability All data generated or analyzed during this study are included in this published article (and its supplementary information files).

Declarations

Ethics approval This article does not describe any studies with human participants or animals performed by any of the authors.

Competing interests The authors declare no competing interests.

References

- Betrabet SM, Bhat JV (1958) Pectin decomposition by species of *Pseudomonas* and their role in the retting of Malvaceous plants. *Appl Microbiol* 6:89–93. <https://doi.org/10.1128/am.6.2.89-93.1958>
- Boiteux LS, Fonseca MEN, Simon PW (1999) Effects of plant tissue and DNA purification method on randomly amplified polymorphic DNA-based genetic fingerprinting analysis in carrot. *J Am Soc Hortic Sci* 124:32–38. <https://doi.org/10.21273/jashs.124.1.32>
- Borneman J, Hartin RJ (2000) PCR primers that amplify fungal rRNA genes from environmental samples. *Appl Environ Microbiol* 66:4356–4360. <https://doi.org/10.1128/AEM.66.10.4356-4360.2000>

- Cabrol L, Malhautier L, Poly F, Lepeuple AS, Fanlo JL (2010) Assessing the bias linked to DNA recovery from biofiltration woodchips for microbial community investigation by fingerprinting. *Appl Microbiol Biotechnol* 85:779–790. <https://doi.org/10.1007/s00253-009-2253-8>
- Carrigg C, Rice O, Kavanagh S, Collins G, O’Flaherty V (2007) DNA extraction method affects microbial community profiles from soils and sediment. *Appl Microbiol Biotechnol* 77:955–964. <https://doi.org/10.1007/s00253-007-1219-y>
- Comeau D, Novinscak A, Joly DL, Filion M (2020) Spatio-temporal and cultivar-dependent variations in the *Cannabis* microbiome. *Front Microbiol* 11:1–14. <https://doi.org/10.3389/fmicb.2020.00491>
- Cullen DW, Hirsch PR (1998) Simple and rapid method for direct extraction of microbial DNA from soil for PCR. *Soil Biol Biochem* 30:983–993. [https://doi.org/10.1016/S0038-0717\(98\)00001-7](https://doi.org/10.1016/S0038-0717(98)00001-7)
- Cuscó A, Catozzi C, Viñes J, Sanchez A, Francino O (2019) Microbiota profiling with long amplicons using Nanopore sequencing: full-length 16S rRNA gene and whole run operon. *F1000Research* 7:1755. <https://doi.org/10.12688/f1000research.16817.1>
- De Fazio D, Boccarusso L, Durante M (2020) Tribological behaviour of hemp, glass, and carbon fibre composites. *Biotribology* 21:100113. <https://doi.org/10.1016/j.biotri.2019.100113>
- Desneux J, Pourcher AM (2014) Comparison of DNA extraction kits and modification of DNA elution procedure for the quantitation of subdominant bacteria from piggy effluents with real-time PCR. *Microbiologyopen* 3:437–445. <https://doi.org/10.1002/mbo3.178>
- Dimitrov MR, Veraart AJ, de Hollander M, Smidt H, van Veen JA, Kuramae EE (2017) Successive DNA extractions improve characterization of soil microbial communities. *PeerJ* 2017:1–29. <https://doi.org/10.7717/peerj.2915>
- Djemiel C, Grec S, Hawkins S (2017) Characterization of bacterial and fungal community dynamics by high-throughput sequencing (HTS) metabarcoding during flax dew-retting. *Front Microbiol* 8:2052. <https://doi.org/10.3389/fmicb.2017.02052>
- Fila G, Manici LM, Caputo F (2001) In vitro evaluation of dew-retting of flax by fungi from southern Europe. *Ann Appl Biol* 138:343–351. <https://doi.org/10.1111/j.1744-7348.2001.tb00119.x>
- Fu J, Mueller H, de Castro JV, Yu C, Cavaco-Paulo A, Guebitz GM, Nyanhongo GS (2011) Changes in the bacterial community structure and diversity during bamboo retting. *Biotechnol J* 6:1262–1271. <https://doi.org/10.1002/biot.201100105>
- Gobbi A, Santini RG, Filippi E, Ellegaard-Jensen L, Jacobsen CS, Hansen LH (2019) Quantitative and qualitative evaluation of the impact of the G2 enhancer, bead sizes and lysing tubes on the bacterial community composition during DNA extraction from recalcitrant soil core samples based on community sequencing and qPCR. *PLoS ONE* 14:1–17. <https://doi.org/10.1371/journal.pone.0200979>
- Gray JP, Herwig RP (1996) Phylogenetic analysis of the bacterial communities in marine sediments. *Appl Environ Microbiol* 62:4049–4059. <https://doi.org/10.1128/aem.62.11.4049-4059.1996>
- Guerra V, Beule L, Lehtsaar E, Liao HL, Karlovsky P (2020) Improved protocol for DNA extraction from subsoils using phosphate lysis buffer. *Microorganisms* 8:1–16. <https://doi.org/10.3390/microorganisms8040532>
- Hammer O, Harper DAT, Ryan PD (2001) PAST: paleontological statistics software package for education and data analysis. *Paleontol Electron* 4:9 http://palaeo-electronica.org/2001_1/past/issue1_01.htm
- Hemkemeyer M, Dohrmann AB, Christensen BT, Tebbe CC (2018) Bacterial preferences for specific soil particle size fractions revealed by community analyses. *Front Microbiol* 9:1–13. <https://doi.org/10.3389/fmicb.2018.00149>
- Henriksson G, Akin DE, Hanlin RT, Rodriguez C, Archibald DD, Rigsby LL, Eriksson KL (1997) Identification and retting efficiencies of fungi isolated from dew-retted flax in the United States and Europe. *Appl Environ Microbiol* 63:3950–3956. <https://doi.org/10.1128/aem.63.10.3950-3956.1997>
- Högfors-Rönholm E, Christel S, Engblom S, Dopson M (2018) Indirect DNA extraction method suitable for acidic soil with high clay content. *MethodsX* 5:136–140. <https://doi.org/10.1016/j.mex.2018.02.005>
- Holben WE, Jansson JK, Chelm BK, Tiedje JM (1988) DNA probe method for the detection of specific microorganisms in the soil bacterial community. *Appl Environ Microbiol* 54:703–711. <https://doi.org/10.1128/aem.54.3.703-711.1988>
- Knauth S, Schmidt H, Tippkötter R (2013) Comparison of commercial kits for the extraction of DNA from paddy soils. *Lett Appl Microbiol* 56:222–228. <https://doi.org/10.1111/lam.12038>
- Law AD, McNees CR, Moe LA (2020) The microbiology of hemp retting in a controlled environment: steering the hemp microbiome towards more consistent fiber production. *Agronomy* 10:1–10. <https://doi.org/10.3390/agronomy10040492>
- Lazarevic V, Gaia N, Girard M, François P, Schrenzel J (2013) Comparison of DNA extraction methods in analysis of salivary bacterial communities. *PLoS One* 8:e67699. <https://doi.org/10.1371/journal.pone.0067699>
- Le Maréchal C, Fourou S, Ballan V, Rouxel S, Souillard R, Chemaly M (2018) Detection of *Clostridium botulinum* group III in environmental samples from farms by real-time PCR using four commercial DNA extraction kits. *BMC Res Notes* 11:1–6. <https://doi.org/10.1186/s13104-018-3549-5>
- Lelievre, M (2020) Mesure de la biomasse moléculaire microbienne du sol. Partie 1: Extraction d’ADN brut du sol. https://www2.dijon.inrae.fr/plateforme_genosol/
- Lim HJ, Choi JH, Son A (2017) Necessity of purification during bacterial DNA extraction with environmental soils. *Environ Health Toxicol* 32:e2017013. <https://doi.org/10.5620/eh.t.2017013>
- Liu M, Ale MT, Kołaczowski B, Fernando D, Daniel G, Meyer AS, Thygesen A (2017) Comparison of traditional field retting and *Phlebia radiata* Cel 26 retting of hemp fibres for fibre-reinforced composites. *AMB Express* 7:58. <https://doi.org/10.1186/s13568-017-0355-8>
- Mahmoudi N, Slater GF, Fulthorpe RR (2011) Comparison of commercial DNA extraction kits for isolation and purification of bacterial and eukaryotic DNA from PAH-contaminated soils. *Can J Microbiol* 57:623–628. <https://doi.org/10.1139/w11-049>
- Mazian B, Bergeret A, Benezet JC, Malhautier L (2020) Impact of field retting and accelerated retting performed in a lab-scale pilot unit on the properties of hemp fibres/polypropylene biocomposites. *Ind Crops Prod* 143:111912. <https://doi.org/10.1016/j.indcrop.2019.111912>
- Méndez-Líter JA, de Eugenio LI, Nieto-Domínguez M, Prieto A, Martínez MJ (2021) Hemicellulases from *Penicillium* and *Talaromyces* for lignocellulosic biomass valorization: A review. *Bioresour Technol* 324:124623. <https://doi.org/10.1016/j.biortech.2020.124623>
- Moawad H, Abd El-Rahim WM, Hashem MM, Gebreil GM, Sabbor A, Zakaria M (2019) Retting and degumming of flax using biotechnology eco-friendly approach: preface. *Egypt J Chem* 62:2033–2045. <https://doi.org/10.21608/ejchem.2019.9571.1641>
- Munshi TK, Chattoo BB (2008) Bacterial population structure of the jute-retting environment. *Microb Ecol* 56:270–282. <https://doi.org/10.1007/s00248-007-9345-8>
- Petric I, Philippot L, Abbate C, Bispo A, Chesnot T, Hallin S, Laval K, Lebeau T, Lemanceau P, Leyval C, Lindström K, Pandard P, Romero E, Sarr A, Schloter M, Simonet P, Smalla K, Wilke BM, Martin-Laurent F (2011) Inter-laboratory evaluation of the ISO standard 11063 “Soil quality - method to directly extract DNA

- from soil samples.” *J Microbiol Methods* 84:454–460. <https://doi.org/10.1016/j.mimet.2011.01.016>
- Plassart P, Terrat S, Thomson B, Griffiths R, Dequiedt S, Lelievre M, Regnier T, Nowak V, Bailey M, Lemanceau P, Bispo A, Chabbi A, Maron PA, Mougel C, Ranjard L (2012) Evaluation of the ISO standard 11063 DNA extraction procedure for assessing soil microbial abundance and community structure. *PLoS ONE* 7:e44279. <https://doi.org/10.1371/journal.pone.0044279>
- Ribeiro A, Pochart P, Day A, Mennuni S, Bono P, Baret JL, Spadoni JL, Mangin I (2015) Microbial diversity observed during hemp retting. *Appl Microbiol Biotechnol* 99:4471–4484. <https://doi.org/10.1007/s00253-014-6356-5>
- Robe P, Nalin R, Capellano C, Vogel TM, Simonet P (2003) Extraction of DNA from soil. *Eur J Soil Biol* 39:183–190. [https://doi.org/10.1016/S1164-5563\(03\)00033-5](https://doi.org/10.1016/S1164-5563(03)00033-5)
- Schneegurt MA, Dore SY, Kulpa CF (2003) Direct Extraction of DNA from soils for studies in microbial ecology. *Curr Issues Mol Biol* 5:1–8. <https://doi.org/10.21775/cimb.005.001>
- Sloggett C, Goonasekera N, Afgan E (2013) BioBlend: Automating pipeline analyses within Galaxy and CloudMan. *Bioinformatics* 29:1685–1686. <https://doi.org/10.1093/bioinformatics/btt199>
- Stach JEM, Bathe S, Clapp JP, Burns RG (2001) PCR-SSCP comparison of 16S rDNA sequence diversity in soil DNA obtained using different isolation and purification methods. *FEMS Microbiol Ecol* 36:139–151. [https://doi.org/10.1016/S0168-6496\(01\)00130-1](https://doi.org/10.1016/S0168-6496(01)00130-1)
- Terrat S, Plassart P, Bourgeois E, Ferreira S, Dequiedt S, Adele-Dit-Derenseville N, Lemanceau P, Bispo A, Chabbi A, Maron PA, Ranjard L (2015) Meta-barcoded evaluation of the ISO standard 11063 DNA extraction procedure to characterize soil bacterial and fungal community diversity and composition. *Microb Biotechnol* 8:131–142. <https://doi.org/10.1111/1751-7915.12162>
- Varma A, Padh H, Shrivastava N (2007) Plant genomic DNA isolation: an art or a science. *Biotechnol J* 2:386–392. <https://doi.org/10.1002/biot.200600195>
- Vasina DV, Moiseenko KV, Fedorova TV, Tyazhelova TV (2017) Lignin-degrading peroxidases in white-rot fungus *Trametes hirsuta* 072. Absolute expression quantification of full multigene family. *PLoS One* 12:e0173813. <https://doi.org/10.1371/journal.pone.0173813>
- Venkatarajan S, Athijayamani A (2020) An overview on natural cellulose fiber reinforced polymer composites. *Mater Today Proc* 37:3620–3624. <https://doi.org/10.1016/j.matpr.2020.09.773>
- Wagner AO, Praeg N, Reitschuler C, Illmer P (2015) Effect of DNA extraction procedure, repeated extraction and ethidium monoazide (EMA)/propidium monoazide (PMA) treatment on overall DNA yield and impact on microbial fingerprints for bacteria, fungi and archaea in a reference soil. *Appl Soil Ecol* 93:56–64. <https://doi.org/10.1016/j.apsoil.2015.04.005>
- Wüst PK, Nacke H, Kaiser K, Marhan S, Sikorski J, Kandeler E, Daniel R, Overmann J (2016) Estimates of soil bacterial ribosome content and diversity are significantly affected by the nucleic acid extraction method employed. *Appl Environ Microbiol* 82:2595–2607. <https://doi.org/10.1128/AEM.00019-16>
- Xue M, Wu L, He Y, Liang H, Wen C (2018) Biases during DNA extraction affect characterization of the microbiota associated with larvae of the Pacific white shrimp, *Litopenaeus vannamei*. *PeerJ* 6:e5257. <https://doi.org/10.7717/peerj.5257>
- Yang J, Li W, Ng BT, Deng X, Lin J, Ye X (2017) Laccases: Production, expression regulation, and applications in pharmaceutical biodegradation. *Front Microbiol* 8:832. <https://doi.org/10.3389/fmicb.2017.00832>
- Yu Y, Lee C, Kim J, Hwang S (2005) Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. *Biotechnol Bioeng* 89:670–679. <https://doi.org/10.1002/bit.20347>
- Zhao D, Liu P, Pan C, Du R, Ping W, Ge J (2016) Bacterial succession and metabolite changes during flax (*Linum usitatissimum* L.) retting with *Bacillus cereus* HDYM-02. *Sci Rep* 6:1–9. <https://doi.org/10.1038/srep31812>
- Zimniewska M (2022) Hemp fibre properties and processing target textile: A review. *Materials* 15:1901. <https://doi.org/10.3390/ma15051901>