Degree project

UNIVERSITY OF SKÖVDE

DETECTION OF *SCLEROTINIA SCLEROTIORUM* IN OILSEED RAPE USING OXFORD NANOPORE SEQUENCING AND QPCR.

Master Degree Project in Bioscience

Second Cycle 30 credits

Spring term 2023

Student: Ana Guadalupe Herrera Hernandez
a22anahe@student.his.se

Supervisor: Maria Algerin
maria.algerin@his.se

Examiner: Magnus Fagerlind
magnus.fagerlind@his.se

Abstract

Sclerotinia sclerotiorum is a notorious phytopathogenic fungus and is the causal agent of the disease Sclerotinia stem rot (SSR) of rapeseed (*Brassica napus*). SSR is one of the main diseases affecting the yield and oil quality of rapeseed crops worldwide. This disease is very hard to predict and control due to all the different factors that are involved in the development of the disease. Successful disease management depends on accurate identification and early detection of plant pathogens. qPCR is a fast, specific, reproducible, and reliable technique for plant pathogen diagnostics. However, one limitation of qPCR is that it is unsuitable to identify and study unknown species, other than those intended, making the detection of unknown pathogens very difficult. An alternative solution is to apply single molecule sequencing, which can provide information at species and strain level. In this study, a total sample of 15 rapeseed leaves coming from three different fields in Sweden with known incidence of SSR disease were analyzed using qPCR and other 15 leaves, coming from the same fields, were analyzed using Oxford Nanopore sequencing to attempt to identify pathogens, *S. sclerotiorum* being the main target. *S. sclerotiorum* was not identified with none of the previous mentioned techniques in any of the samples. Perhaps, *S. sclerotiorum* was not present on the samples at the time of the collection, due to the unfavorable weather conditions for the release of the spores. However, some issues were present during the development of the qPCR assays that also could have affected the results. Regarding Oxford Nanopore sequencing, other fungal species were identified instead.

Popular scientific summary

Sclerotinia stem rot, caused by the fungi *Sclerotinia sclerotiorum*, is one of the most relevant diseases affecting rapeseed crops due to the serious economic losses that causes every year. This disease is very hard to predict and control due to all the different factors that are involved in the development of the disease. Moreover, the use of fungicides need to be in optimal time to help preventing the disease, since once the disease is manifested, the fungicides do not have any curative effect, negatively affecting the environment and the local fauna. Early and precise detection of plant pathogens are essential for effective disease management. Many pathogens are difficult to identify based on morphological characteristics and it is also time consuming. Quantitative polymerase chain reaction (qPCR) assays is an effective tool for the fast diagnosis of plant diseases and many samples can be analyzed at the same time, identifying the pathogen at the same time as quantifying it. This study aimed to use this technique to evaluate if there is a difference between processing samples individually, combined, or pooled. In the present study, individual samples were considered DNA extracted from individual leaves separately. Composite samples were considered DNA extracted from one cut made to five individual leaves, making five cuts. The pooled samples were made combining five DNA samples from the individual samples. Another method used to identify pathogens in this study was Oxford Nanopore Sequencing, this is a more recent technology, addressing one of the qPCR limitations which is the inability to identify and study unknown species, other than the targets. Oxford Nanopore Sequencing can detect multiple pathogens simultaneously in the same sample. In this study, rapeseed leave samples coming from three different fields in Sweden, with known disease prevalence, were analyzed. The pathogen *S. sclerotiorum* was not identified by qPCR neither by Oxford Nanopore Sequencing in any of the samples. The way this pathogen spreads between the fields is through the release of spores (microscopic biological particles) and they need favorable conditions to be released. One of the reasons could have been that the release of the spores did not occur, due to the unfavorable weather conditions, such as the low precipitation and the cold temperatures prior to the collection. However, some issues were present during the development of this study that could also explain why the pathogen was not identified. One of the issues that took place during the qPCR assays were the presence of inhibitors. Inhibitors are any factor that prevents the amplification of nucleic acids through the polymerase chain reaction, and inhibitors are a quite common cause of amplification failure. Another problem was the possible contamination of the lab equipment or the reagents used during the assays. Lastly, Oxford Nanopore Sequencing identified multiple fungal species, but the results were not consistent when using two different reference databases. One of the explanations could be that the reference database used for analysis will affect the sensitivity and specificity of the analysis, which in turn can impact the accuracy and reliability of the results.

Abbreviations

Bp	Base pairs
Cq	Quantitation cycle
CTAB	Cetyl Trimethylammonium Bromide
DNA	Deoxyribonucleic acid
DSI	Disease severity index
Fg	Femtogram
GOI	Gene of interest
ITS	Internal transcribed spacer
kb	Kilo bases
ng	Nanogram
NCBI	National Centre for Biotechnology Information
NTC	Non Template Control
nrDNA	Nuclear DNA
nrRNA	Nuclear ribosomal RNA
ONT	Oxford nanopore Technology
PCR	Polymerase chain reaction
qPCR	Quantitative Polymerase Chain Reaction
R2	Regression Factor Value
RNA	Ribonucleic acid
SSR	Sclerotinia stem rot
6-FAM	6-carboxy-fluorescein

Table of content

Introduction

Plant diseases caused by pathogenic microorganisms such as bacteria, fungi, viruses and phytoplasmas are among the main factors limiting crop production and cause substantial economic losses (Chalupowicz et al., 2018). They destroy up to 30% of crop products through disease and spoilage processes. Control of these fungi is essential for improving food security (Avery et al., 2019; Fones et al., 2020).

Among these fungi, *Sclerotinia sclerotiorum* (*S. slerotiorum*) stands out for being a notorious phytopathogenic Ascomycota fungus with a remarkably broad host range of, encompassing over 600 plant species worldwide. It is the causal agent of the fungal disease Sclerotinia stem rot (SSR) of rapeseed (*Brassica napus*) and it is one of the principal diseases affecting the yield and oil quality of rapeseed crops (Ding et al., 2021; Mbengue et al., 2016).

S. sclerotiorum produce dense dark resting structures called sclerotia and these structures help fungi to survive during unfavorable conditions such as freezing, desiccation, microbial attack, or the absence of a host (Smith et al., 2015). Sclerotia germinate to produce apothecia (cup-shaped, spore-bearing structures) during favorable conditions and release ascospores that can disseminate over several kilometers infecting the petals, which then settle into the crop canopy, primarily on leaves and branches or leaf axils, and subsequently start leaf and stem infection, under the right circumstances. SSR symptoms typically emerge on leaves, stems, and branches, but they can also appear on pods and roots. SSR symptoms on leaves include round, greyish, water-soaked lesions that are frequently surrounded by a yellow halo, whereas lesions on stems, branches, roots, and pods cause lodging, premature shattering, and yield loss in plants that are highly infected. Sclerotia eventually develop inside of the affected stems and roots, and occasionally even on the outside (Kamal et al., 2015; Khangura & van Burgel, 2021).

Different factors have effect on SSR infection, like number of years since last oilseed rape crop was grown, disease incidence in last Sclerotinia host crop, crop density and rain in the last two weeks before flowering (Twengström et al., 1998). One key factor to consider, for risk assessment, is the disease incidence and can easily be estimated at late reproductive stages by scouting fields and taking counts of diseased plants in several, 4 at least, representative parts of the field. By dividing the number of diseased plants by the total number of plants an estimate of disease incidence for the field can be determined and it is represented in percentage. Tracking disease incidence across years also will help determine the potential load that may be present in a particular field (Peltier et al., 2012).

Currently, the SSR is managed mainly through the use of fungicides and crop rotations because there are no commercially available resistant cultivars (del Río et al., 2007; Derbyshire & Denton-Giles, 2016). However, the effectiveness of crop rotations is limited due to the pathogen's ability to spread airborne inoculum and the capacity of the sclerotia produced by pathogen to remain viable in soil for several years (Kamal et al., 2015; Kutcher & Wolf, 2006).

Chemical control alone has been found to be comparatively less effective because of the mismatch in spray timing and ascospore releases, and many fungicides are gradually losing their efficacy due to the increasing development of resistant strains (Kamal et al., 2015). Timing of fungicide spray is critical for the successful management of SSR. Fungicides need to be applied before the appearance of any symptoms, as all fungicides recommended for SSR management have no curative activity (Khangura & van Burgel, 2021).

A reliable and fast method is required to accurately detect early disease stages, in order to reduce the application of chemicals and its negative impact on the environment, and the fields where fungicide application would be beneficial (Wang et al., 2015; Yin et al., 2009). Successful disease management depends on accurate identification and early detection of plant pathogens and PCR-

based assays have developed into effective tools for the rapid diagnosis of plant diseases (Ma & Michailides, 2007). qPCR is a fast, specific, reproducible, and reliable technique for plant pathogen diagnostics and provides a rapid and accurate assessment of the target pathogen (Almquist & Wallenhammar, 2014).

In order to identify *S. Sclerotiorum* by qPCR the use of specific primers is necessary. According to Freeman et al. (2002), PCR assays using consensus fungal primers (ITS4/ITS5) and *S. Sclerotiorum* specific primers (SSFWD/SSREV) were done using DNA from a range of different fungal species including pathogens of rapeseed, and other air-dispersed fungi. DNA from all the species tested was amplified using the consensus fungal primers, but primers SSFWD/SSREV only amplified DNA from *S. slerotiorum*. The size of the SSFWD/SSREV PCR products from the *S. sclerotiorum* isolates was around 278 base pairs and results demonstrated SSFWD/SSREV amplified *S. sclerotiorum* DNA specifically.

Nucleic acids can be amplified, detected, and quantified simultaneously using qPCR. Quantification can be reached by absolute or relative quantification. The most popular technique is relative quantification. It is cheap and simple to perform, however it depends on using one or more reference genes to normalize the quantities of the genes of interest (GOIs). Empirical analysis is used to identify the optimal reference gene selection and number. An appropriate reference gene must have stable expression across experimental, have similar amplification efficiency and abundance to the GOIs. In reality, this is uncommon since reference genes frequently introduce bias into an experiment, causing inaccurate results interpretation (Boulter et al., 2016).

The absolute quantification method relies on a standard curve constructed from known concentrations of standards to measure the actual copy numbers of a particular target (Dhanasekaran et al., 2010). Absolute quantification does not depend on reference genes and, in terms of copy number or concentration, absolute quantification establishes the precise quantity of the target DNA, compared to relative quantification, which calculates the ratio between the amounts of target and reference genes. Therefore, absolute quantification is considered to indicate a more reliable measurement of the amount of target DNA (Boulter et al., 2016).

One of the parameters used to indicate a reliable standard curve is linear range. A good standard curve should have a linear range that covers the expected range of the sample concentrations. This implies that the Cq value(quantitation cycle) should increase linearly as the concentration of the target DNA increases. Another parameter is the correlation coefficient, R2. An R2 value of 0.98 or greater is typically regarded as acceptable, indicating a good linear correlation between the Cq values and the log concentration of the standard samples. How successfully the PCR process amplifies the target DNA template is known as efficiency and to ensure precise quantification, the PCR efficiency of the amplification reaction should be close to 100% (between 90 and 110%). A high PCR efficiency shows that the amplification process is effective and reproducible, and it suggests that the target DNA quantification in the test samples will probably be accurate and reliable. The slope of the standard curve produced by a series of known template concentrations is used to calculate the PCR efficiency. A slope between -3.1 and -3.6 is acceptable, and an efficiency of 100 percent with a slope of -3.3 is ideal. Also, the standard curve should be produced using at least five standard points and the accuracy and precision of the quantification can be increased by using additional standard points (Ramakers et al., 2003).

According to Dhanasekaran et al. (2010) PCR product standards have good stability in concentrated form but may lose integrity when serially diluted and aliquoted for standard construction. The chance of degradation is likely to be increased by repeated freeze–thaw and general handling of diluted standards. But plasmid standards appear to survive such handling better. Dhanasekaran et al. (2010) also determined that serial dilutions of cloned plasmid are the more robust and suitable standard for long-term study purposes, even if plasmid construction and cloning requires extensive work and is expensive. Though PCR product standards are easier to

make and less expensive, their data suggest that care should be taken to use only freshly prepared serial dilutions.

However, one limitation of PCR-based methods is that they are unsuitable to identify and study unknown species, other than those intended, making the detection of multiple pathogens simultaneously in plants with unidentified pathogens very difficult (Bollmann-Giolai et al., 2022; Loit et al., 2019). An alternative solution is to apply single molecule sequencing technologies (PacBio/Oxford Nanopore), which can provide information at species and strain level (Bollmann-Giolai et al., 2022). The development of single-molecule sequencing technologies, often called third generation sequencing (TGS), has advantages over other sequencing methods, such as, enabling the sequencing of nucleotide molecules (DNA or RNA) without the need for PCR amplification of the template (in some cases). They also allow the real-time analysis of the produced data, TGS technologies directly target single nucleotide molecules, enabling real-time sequencing, where reads are available for analysis as soon as they have passed through the sequencing device (Figure 1). Moreover, they require only a small amount of DNA with low cost for a single run. Furthermore, another advantage is the increase in read length, from dozens of bases to dozens of thousands of bases per read, TGS platforms of Pacific Biosciences and Oxford Nanopore Technologies allow average sequence lengths of 20,000 bases (Athanasopoulou et al., 2021; Chalupowicz et al., 2018; Loit et al., 2019; Lu et al., 2016).

The first nanopore-based sequencer was released in 2014 by Oxford Nanopore Technologies (ONT). Some of the benefits of Oxford nanopore sequencing devices are, firstly, since nucleotide detection in Oxford nanopore sequencing does not require imaging equipment, the system is reduced in size to a portable level fitting in the palm of a hand. The cost of the device is also much lower compared to other sequencers. Oxford nanopore sequencing devices (MinION) can be powered through the USB port of laptop computers, so sequencing can be conducted anywhere. Without an image analysis step, real-time base calling is also possible during sequencing, enabling quick detection of target DNA for the screening of pathogens from clinical samples, for example (Athanasopoulou et al., 2021; Kono & Arakawa, 2019).

The first step is library preparation, which is required for many applications that use MinION. In this stage, double-stranded DNA (dsDNA) should be used so that sequencing of both strands can be done. Nanopore technology relies on detecting changes in the electric current, caused by the disorder of nanopore proteins when DNA or RNA strands pass through them. Each nucleotide's characteristic changes in the ionic current give each base its own individual signature. The process of converting the electrical signal into a nucleic acid sequence is called basecalling and create FASTQ files, these files can be used in downstream analyses that require a bioinformatic tools (Athanasopoulou et al., 2021; Lu et al., 2016).

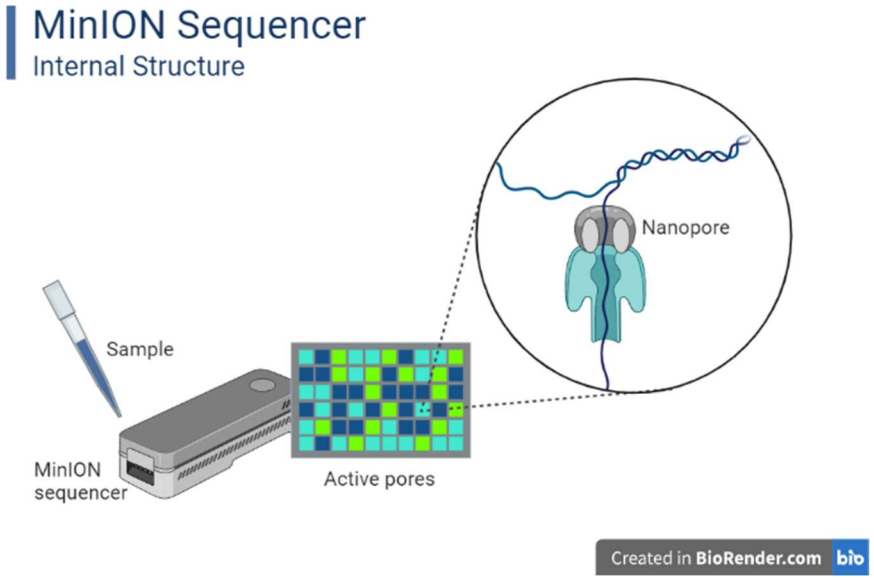

Figure1. Image showing the MinION device. Created with BioRender.com.

Many study fields, including the health sciences and agriculture, depend on the identification of fungi at the species level to define the most appropriate treatments, understand epidemics, and understand transmission mechanisms. The strategy based on the sequencing of standardized genomic fragments (DNA barcoding) is widely used nowadays. DNA barcoding uses sequence variation within a short and standardized section of the genome, known as a "barcode", to identify species with accuracy. In order to identify taxonomic connections, this method is based on the analysis of variability within a standard DNA barcode region (Badotti et al., 2017; Toju et al., 2012). In cataloging multiple fungal species and thereby assemble reference databases of fungal diversity, “DNA barcoding” based on the nucleotide sequence information of a target gene region can be highly effective. DNA barcoding is a method of identifying unknown samples through known classifications (Toju et al., 2012).

Molecular identification through DNA barcoding of fungi has become an essential part of fungal research. Interest in sequenced-based analysis of environmental samples (environmental barcoding) has escalated in the past years as it allows to study abundance and species richness of fungi at a high rate and more reliably than conventional biotic surveys, such as fingerprinting methods based on banding patterns obtained from restriction site polymorphisms or denaturing gradient gel electrophoresis profiles (Martin & Rygiewicz, 2005; O’Brien et al., 2005). The internal transcribed spacer (ITS) of nuclear DNA (nrDNA) is the preferred DNA barcoding marker, for fungi, both for the identification of single taxa and mixed environmental templates (environmental DNA barcoding) (Martin & Rygiewicz, 2005).

For a wide number of fungal lineages analyzed, the ITS region of nuclear ribosomal RNA (nrRNA) showed the highest probability of correct identification and the most clearly defined barcode gap (Badotti et al., 2017). Schoch et al. (2012) discovered that, for a very broad group of examined fungi, the ITS region is among the markers with the highest probability of accurate identifications, compared to the three markers studied in this experiment, the Small SubUnit-coding sequence, the Large SubUnit-coding sequence and the largest subunit of RNA polymerase II. Since then, the ITS region has been accepted as the standard barcode marker for fungi.

The ITS region comprise the ITS1 and ITS2 regions, divided by the 5.8S gene which is considered part of the ITS region. Traditional Sanger sequencing methods have been frequently used to target the whole ITS region and typically ranges between 450 and 700 bp (Martin & Rygiewicz, 2005). Using high-throughput sequencing, thousands of sequences can be analyzed from a single

environmental sample, allowing a thoroughly analysis of the fungal diversity. A variety of primers can be used for amplifying the entire or specific portions of the ITS region (Bellemain et al., 2010).

DNA sequence data, resulting from Oxford Nanopore Sequencing, are usually used to analyze fungi and fungal communities through barcoding and metabarcoding efforts. Kraken2 is a bioinformatics tool used for taxonomic classification of DNA sequencing reads. It uses a precompiled database of reference genomes and their taxonomic classifications to assign taxonomic labels to input sequences. The Kraken2 database includes sequences from a variety of sources, like, RefSeq (completely assembled and annotated reference genomes of archaea, bacteria, and viruses from the NCBI RefSeq database), GenBank (a database of genetic sequences maintained by NCBI), among others. In addition to these databases, Kraken2 also includes the ability to add custom databases to the classification process, allowing users to tailor the tool to their specific needs (Marcelino et al., 2020).

Other important databases are UNITE and the International Nucleotide Sequence Databases Collaboration (INSDC). The nuclear ribosomal internal transcribed spacer (ITS) region is the primary genetic marker for fungi, and more than 1 000 000 full-length, Sanger-derived fungal ITS sequences are available for reference in INSDC. UNITE is a database for the molecular identification of fungi and it was launched in 2003 as a curated copy of the public fungal ITS sequences. It also targets the ITS region and offers around 1 000 000 public fungal ITS sequences for reference (Lu & Salzberg, 2020; Nilsson et al., 2019). It depends on the database and the type of data being updated how frequently databases are updated (Austin-Tse et al., 2022). It is the responsibility of the researcher to use the most recent version of a database to guarantee the precision and reliability of their findings.

Loit et al. (2019) reports the ITS1catta-ITS4ngsUni primer pair amplified mostly fungal DNA (99.9% of identified reads). The ITS1catta primer covers nearly all Ascomycota and Basidiomycota as well as selected groups of zygo-mycetes and early diverging lineages but discriminates against plants and most other eukaryote groups (including fungal taxa Mortierellomycota and Tulasnellaceae) (Loit et al., 2019).

Almquist and Wallenhammar (2014) used qPCR for the detection of *S. sclerotiorum*, using leaves as samples, but this study did not compare if a difference exits in the results between individual, composite and pooled samples. Knowing if there is a significant difference in the results between these samples could save time and resources in the detection of this pathogen.

Aims

- Discover if *S. sclerotiorum* or other plant pathogens can be detected by Oxford Nanopore Sequencing using the device MinION from Oxford Nanopore Technologies, using rapeseed leaves collected from three naturally infected fields with SSR as sample.

- Discover if *S. sclerotiorum* can be detected amplifying the sequence of 278 bp from the 5.8s ribosomal RNA gene by qPCR , and if it is, then compare composite, pooled and individual leaf samples.

Objectives

- Extract DNA from rapeseed leaves collected from naturally infected fields with different stem rot incidence.
- Use PCR to amplify the ITS region of the DNA isolated from the samples, using the primer pair ITS1catta-ITS4ngsUni.
- Use Oxford Nanopore Sequencing to obtain FASTQ files (reads) from the samples.
- Use the databases Kraken2 and UNITE to identified fungal pathogens.

- Perform an absolute qPCR and discover if *S. sclerotiorum* can be identified using three different sets of leaves samples (individual, combined and pooled) and quantify the DNA.
- Use a statistical test to compare and find out if there is a significant difference on DNA quantities between the individual, combined and pooled samples.

Materials and methods

Samples

A total of 30 rapeseed leaves were collected during May 2022 from three different fields in Skaraborg Sweden, Dala, Grevbäck and Hovby (10 leaves from each field). Dala field had a sclerotinia stem rot incidence of 6%, Grevbäck 8% and Hovby 25%. The leaves were kept in a freezer at -20°C since then. Half of the leaves from each field (15 leaves, five from each field) were used as samples to carry out Oxford Nanopore Sequencing and the other half (15 leaves, five from each field) to perform qPCR assays.

The DNA samples used to perform Oxford Nanopore Sequencing were extracted from each leave individually, then the five DNA samples proceeding from the same field were pooled together.

One of the objectives of this study was to find out if there was a significant difference on DNA quantities between the individual, combined and pooled samples through qPCR. Each field had its own individual, combined and pooled samples.

Five leaves were used to carry out DNA extractions individually, from round cuts done to each one of the leaves of around 5 cm in diameter. To create the combined samples, one cut from each one of the five leaves used for individual extractions, made with the same tool, were combined before the DNA extraction process. After completing the DNA extraction procedure, purity and concentration were measured from each elution, and 6 µl from each one of the five individual samples were put together in the same tube to generate the pooled samples, and once again, purity and concentration were measured.

DNA extraction from leaves

DNA extraction was performed using a modified version of the E.Z.N.A SP Plant DNA Kit extraction method (Omega Bio-Tek) as described by Almquist & Wallenhammar (2014). DNA purity was measured using the spectrophotometer Nanodrop (Thermo Fisher Scientific), by determining the absorbance at two different wavelengths, the ratio of the absorbance at these two wavelengths is generally indicated in two ratios 260/280 and 260/230. DNA concentration was measured using fluorometer Qubit dsDNA HS assay kit (Thermo Fisher Scientific). Samples were kept in a freezer at -20°C until further use.

PCR and gel electrophoresis

Several PCR reactions were performed using different annealing temperatures to optimize the annealing temperature. Phusion™ Hot Start II High-Fidelity DNA polymerase kit (Thermo Scientific) was used for PCR. Three DNA samples from the three different fields (Dala, Grevbäck and Hovby), were used to perform these reactions. Each PCR was run with the settings shown in Table 1, only the annealing temperatures varied.

Table 1. PCR settings.

Step	Temperature	Time
Initial denaturation	98 °C	30 s
Denaturation (40 cycles)	98 °C	10 s
Annealing	55 °C, 55.2 °C, 55.5 °C, 55.7 °C, 55.9 °C, 60 °C, 60.3 °C, 60.5 °C, 60.8 °C, 61 °C, 61.2 °C, 61.5 °C, 61.7 °C, 61.9 °C, 62 °C, 62.2 °C, 62. 4 °C, 62.7 °C, 62.9 °C, 63 °C.	30 s
Elongation	72 °C	60 s
Final extension	72 °C	10 m

The reaction mixture used for PCR optimization was 20µl, following the Phusion™ Hot Start II High-Fidelity DNA polymerase kit protocol (Thermo Scientific), shown in Table 2. The primer pair ITS1Catta (5-ACCWGCGGARGGATCATTA-3) and ITS4ngsUni (5-CCTSCSCTTANTDATATGC−3) (SIGMA) Was used to perform all PCR reactions targeting the ITS region. After an appropriate annealing temperature was chosen, the three samples were amplified using a reaction mixture of 50 µl, as shown in Table 2. All the reactions were performed with a PTC-200 Peltier thermal cycler (Bio-Rad).

The resulting amplicons were run on agarose gel, 10 µl of PCR product with 2 µl 6X purple gel loading dye (New England Biolabs). 100 bp molecular weight DNA ladder (NEB) was loaded in the first and last well, in 1 % agarose (1X TAE buffer) stained with 1X GelRed® (Biotium). The gel was run at 90 volts for 90 minutes.

Table 2. PCR reactions components.

Component	Concentration	Volume (µl)
H_2O	-	8.4
Phusion HF Buffer	5X	4
dNTPs	10 mM	0.4
Forward primer ITS1Catta	0.5 µM	1
Reverse primer ITS4ngsUni	0.5 µM	1
Phusion Hot start II polymerase	0.02 U/µl	0.2
Template DNA (3µl)	Dala 2.5 ng/ Grevbäck 4.8 ng/ Hovby 4.7 ng	5

PCR Cleaning

The samples were cleaned using QIAquick® PCR purification kit (Qiagen) and following the protocol provided by the manufacturer, to remove nucleotides, primers and other impurities originated from the PCR process. Concentration was measured using fluorometer-Qubit dsDNA HS assay kit (Thermofisher) and spectrophotometer-Nanodrop (Thermofisher) to measure purity. This kit was used because it was available in the laboratory.

Oxford Nanopore Sequencing

Library preparation

Three samples (one from each field) were used for Oxford Nanopore Sequencing using the MinION device. The Native barcoding expansion kit 1-12 (EXP-NBD104), ligation sequencing kit (SQK-LSK109) and protocol from Oxford Nanopore Technologies were used during this procedure. The first step was the library preparation of each sample where 200 fmol of each sample were used. All the instructions from the Oxford Nanopore Technologies protocol were followed. This technique allows for simultaneous target-enrichment and barcode-multiplexing of up to 12 libraries, that can be loaded in the same sequencing run (Karamitros & Magiorkinis, 2018).

Sequencing and basecalling

All the samples were sequenced in the same run with a Spot-ON Flow Cell R9 Version (Oxford Nanopore Technologies) in the laboratory at room temperature, connecting the MinION device (Oxford Nanopore Technologies) to a laptop where the MinKNOW software was installed. The Qscore selected was 10. Guppy command line software was set on high accuracy basecalling and placed the reads into passed and failed files. Sequencing was stopped after 96 hours (due to the easter holidays).

Sequence analysis

All the sequences in FASTQ files obtained, were analyzed using Kraken2 taxonomic classification system, for this, two reference databases were used. The first one was the Kraken2 database containing all fungal sequences from NCBI's Refseq (complete fungal genomes), NCBI's Refseq database is updated daily when new information is received from an external source (Pruitt et al., 2020). Since kraken 2 allows to add databases to the classification process, allowing users to tailor the tool, the second database was from UNITE which contains all fungal ITS sequences of the UNITE and INSDC databases, using the last update October 16th 2022 (Abarenkov et al., 2022) . After analyzing and comparing all the sequences obtained with the databases, the bioinformatic tool Pavian was used for further taxonomic analysis and the generation of Sankey diagrams to visualize the taxonomic classification of each sample.

Plasmid digestion and purification

The other technique used to detect *S. sclerotiorum* in this study was qPCR and part of the objectives in this experiment was to quantify the *S. sclerotiorum* DNA present in the samples from the fields. In order to quantify DNA in a qPCR assay, the use of a standard curve is necessary. To create the standard curve used in this study, the following steps were performed.

Due to lack of time, it was necessary to use a previous isolated pCR4 - TOPO® vector, which was previously sequenced and confirmed to contain the target fragment (the sequence of 278 bp from the 5.8s ribosomal RNA gene). This sample was digested with endonuclease NotI to linearize the plasmid, before creating dilution series to generate the standard curve.

First, a restriction digestion reaction was conducted to linearize the plasmid containing the 5.8s ribosomal RNA gene, following the New England Biolabs protocol. The reaction was composed of restriction enzyme NotI (10 units), mixed with Buffer 3.1 (1x, NEB), plasmid DNA (2.80 µg) and nuclease free water to a final reaction volume of 50 µl. The reaction was incubated at 37° C for 60 minutes, followed by inactivation of the enzyme by heating at 65° C for 20 minutes. The product of the digestion reaction was loaded in 1 % agarose (1X TAE buffer) stained with 1X GelRed® (Biotium), 1kb bp molecular weight DNA ladder (NEB) was also loaded and undigested sample was loaded as control to confirm linearization of the product. The gel was run at 90V for 90 minutes. Cleaning of linearized plasmid was done using Qiaquick PCR Purification Kit (Qiagen) according to the manufacturer's instructions. Again, this kit was used because it was available in the laboratory. Concentration and purity were measured, as described previously.

qPCR Assay

The gene to be amplified with qPCR was the 278 bp fragment of ribosomal RNA gene from *S. sclerotiorum.* The kit used was TaqMan® Universal PCR Master Mix (Life Technologies) and the primer pair used for amplification was SSFWD (5´- GCTGCTCTTCGGGGCCTTGTATGC-3´) and SSREV (5´- CAGCTTGGTATTGAGTCCATGTCA-3´) (Sigma), also shown in Appendix 1. The amplification was detected using a hydrolysis MGB probe (5´- CGCCAGAGAATATCAA-3´) labelled with 6-carboxy-fluorescein (6-FAM) (Life Technologies). qPCR was performed on an AriaMx Real-time PCR system (Agilent) on MicroAmp Optical 96-Well Reaction Plates (Life Technologies) and all samples were run in triplicates, every essay also included triplicates of a no template control (NTC) and triplicates of the 10-fold dilution series to create the standard curve. The composition of the reactions mixture was composed of 1X TaqMan Universal PCR Mastermix (Applied

Biosystems), 0.2 µM of each primer, 0.2 µM of the probe, 5 µl of template DNA and nuclease free water up to a total volume of 25 µl, as described by Almquist & Wallenhammar (2014). The thermal cycling conditions were an initial denaturation for 10 min at 95°C, followed by a touchdown PCR consisting of a total of 54 cycles: 14 cycles of 95°C for 15 s followed by annealing for 60 s with temperature decreasing by 0.5°C every cycle from 72–65°C, 40 cycles of 95°C for 15 s followed by annealing for 60 s at 65°C, same as described by Almquist & Wallenhammar (2014). The software used to visualize the amplification plots and the Cq values was AriaMx Version 1.8 (Agilent).

Absolute quantification

The amount of DNA was quantified using a standard curve. The standard curve was created using a 10-fold dilution series and contained known amounts of a linearized plasmid (0.002 fg to 2000 fg), carrying the 278 bp target sequence from *S. sclerotiorum.* The 10-fold dilution series were analyzed in triplicates and Cq values were plotted against the log amount of plasmid DNA to create a standard curve. During each qPCR essay, the 10-fold serial dilutions were included in the same reaction plate to create its own standard curve and quantify the amount of DNA from the unknown samples, each standard curve contained seven measuring points. The total number of DNA copies present in each dilution are reported in Appendix 3, table 9.

Calculation of number of DNA copies

The correlation between the actual number of copies and concentration of plasmid DNA containing the target gene used for creating the standard curve was calculated with the following equation (Short & Zehr, 2005):

Number of total DNA copies = [DNA Conc (g/µl) / (660 g/mol x total base pairs) x 6.023 x 10^{23}] x Volume of sample.

Where DNA concentration is expressed in grams per µl, 660 g/mol is the average weight of one mol base pair, total base pairs is the plasmid DNA base pairs plus the insert DNA base pairs, and the volume of sample is expressed in µl used for each reaction.

Results

DNA extraction from leaves

Results of DNA concentration and purity measurements with Nanodrop Spectrophotometer (Thermo Fisher Scientific) and Qubit Fluorometer (Thermo Fisher Scientific), after extraction and cleaning, are shown in Appendix 2.

PCR and electrophoresis

Three DNA samples from the three different fields (Dala, Grevbäck, Hovby), were used for PCR and primers were evaluated to find the most optimal annealing temperature. Many different annealing temperatures, ranging from 55°C to 63°C, were tested and compared using the primer pair ITS1Catta and ITS4ngsUni. The expected amplicon sizes were 400-900 bp, the decision on the annealing temperature was based on the visualized bands after gel electrophoresis was run. The most suitable annealing temperature was chosen at 55.2° C. After deciding on the proper annealing temperature, a PCR with an annealing temperature of 55.2 °C was performed with primer pair ITS1Catta and ITS4ngsUni and gel electrophoresis was run (Figure 2). The size of the product obtained was 700 bp, which was in the range of 400-900 bp and indicated a good amplification of the ITS region from the DNA isolated from the field samples. Although the gel showed other bands at lower sizes, due to variability of the ITS region among species, the 700 bp band was the most prominent band.

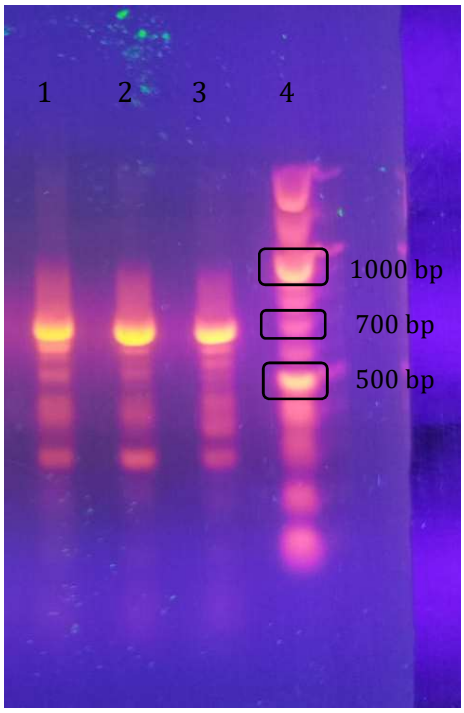

Figure 2. Gel image illustrating PCR product run on agarose gel 1%, after amplification of the three DNA samples obtained from different fields using primer pair ITS1Catta and ITS4ngsUni at the annealing temperature of 55.2 °C. Well number 4 represent 100 bp DNA ladder (NEB). Well number 1 represents the sample of Greväck field. Well number 2 represents the sample of Hovby field. Well 3 number represents sample of Dala field.

PCR cleaning

The samples were cleaned, after, purity and concentration were measured. The results of these measurements are shown in Table 3 . These samples were used for sequencing.

Table 3. DNA quantification and purity results.

Sample ID	DNA concentration ng/µl	Absorbance 260/230	Absorbance 260/280
Dala	46.4	2.44	1.86
Greväck	38.8	2.13	1.84
Hovby	46.0	2.97	1.81

Oxford Nanopore Sequencing

To analyze the sequences obtained from Oxford Nanopore Sequencing, two databases were used. The first one was the Kraken2 database containing all fungal sequences from NCBI's Refseq (complete fungal genomes). The second database was from UNITE containing all fungal ITS sequences of the UNITE and INSDC databases. Using the Kraken2 database, the percentage of classified data in the three samples were quite similar, between 61-65% (Table 4). Using the UNITE database the percentage of classified data was higher for all the samples, between 96-97% (Table 4). Both databases were known to contain the *S. sclerotiorum* sequences.

Table 4. Comparison of classification between the databases.

Sample ID	Percentage of reads (%)	Kraken2 database	UNITE database
Dala	Classified	65.2	97.1
	Unclassified	43.8	2.9
Grevbäck	Classified	66.3	96.6
	Unclassified	33.7	3.4
Hovby	Classified	61.7	97.8
	Unclassified	38.3	2.2

After this step, Pavian tool was used for further taxonomic analysis and classification, and the generation of Sankey diagrams to visualize the taxonomic classification of each sample resulted from each database. When using the first reference database (Kraken2), only fungi were classified (Figure 3) as expected since the database only contained fungal sequences from NCBI's Refseq. However, when analyzing the sequences using the UNITE database the results were very different, see Figure 4. The fungi species identified with the two databases were not the same even though the same sequences were used for both. Not only the fungi species were different, but the UNITE database classified most of the sequences (around 70%) under the Viridiplantae kingdom, being Brassica napus (rapeseed plant) the biggest portion. Also, a very small portion (around 6%) of the sequences was classified under the kingdom Metazoa. These results were not expected at all since the UNITE database was expected to contain only fungal ITS sequences of the UNITE and INSDC databases. The sample from the field Dala (Figure 3), the sample from the field Grevbäck (Figure 5) and the sample from Hovby field (Figure 6) presented very similar patterns when using the Kraken2 database. The three fields also presented very similar patterns (Figure 4) when using UNITE database.

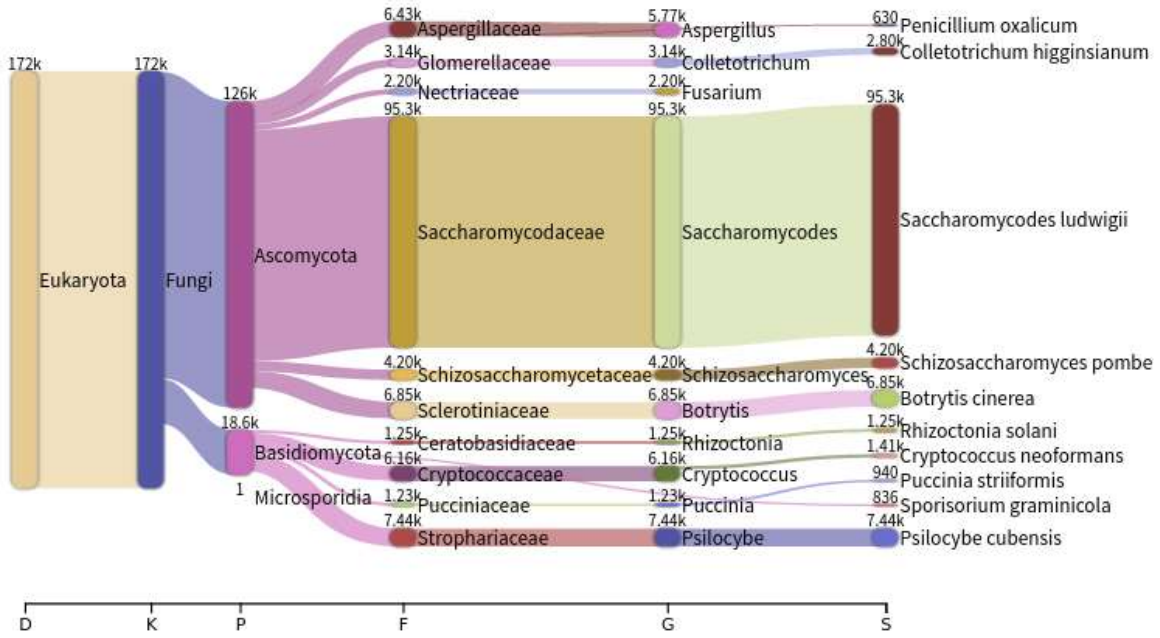

Figure 3. Detailed taxonomic identification of sample Dala using Kraken2 database. Numbers on the top of the tree and each branch represent the number of reads for each group. Taxonomic groups are represented by letters D-domain, K-kingdom, P-phylum, F-family, G-genus and S-specie.

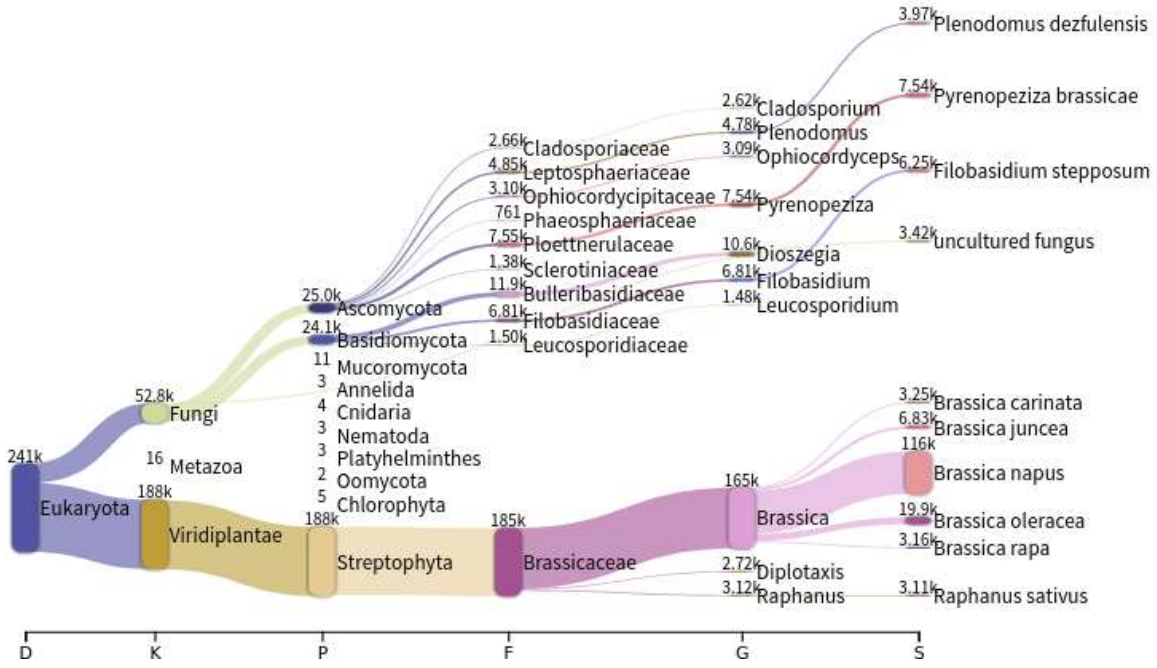

Figure 4. Detailed eukaryotic taxonomic identification of Dala sample using UNITE database. Numbers on the top of the tree and each branch represent the number of reads for each group. Taxonomic groups are represented by letters D-domain, K-kingdom, P-phylum, F-family, G-genus and S-specie.

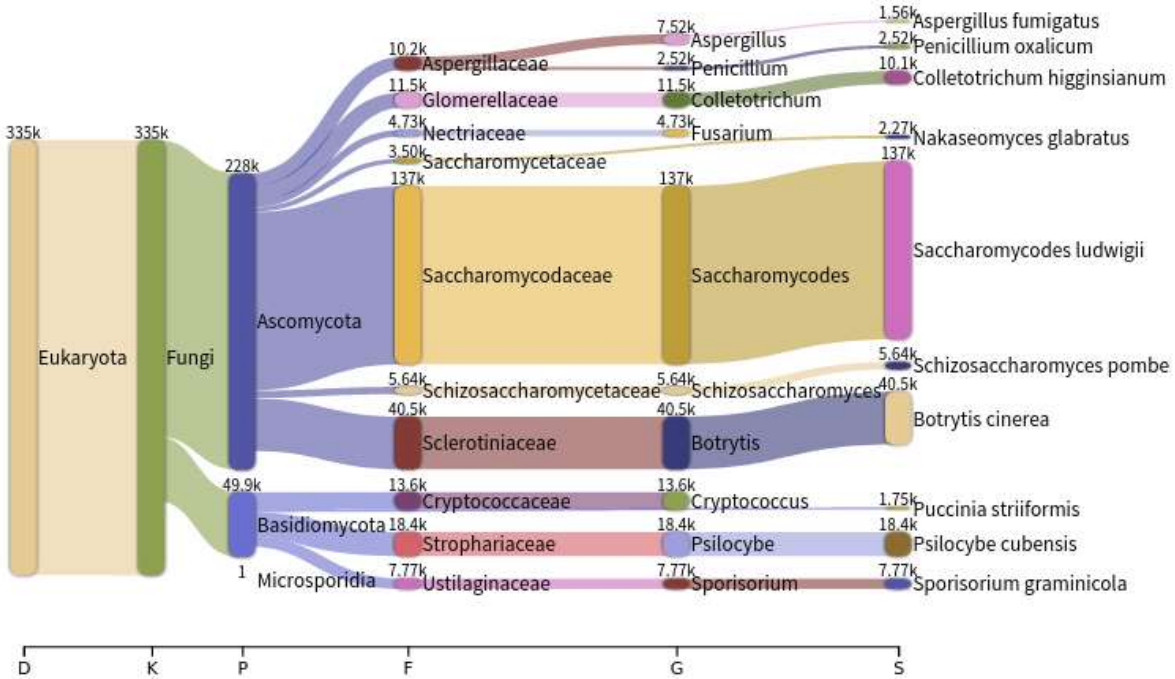

Figure 5. Detailed taxonomic identification of sample Greväck using Kraken2 database. Numbers on the top of the tree and each branch represent the number of reads for each group. Taxonomic groups are represented by letters D-domain, K-kingdom, P-phylum, F-family, G-genus and S-specie.

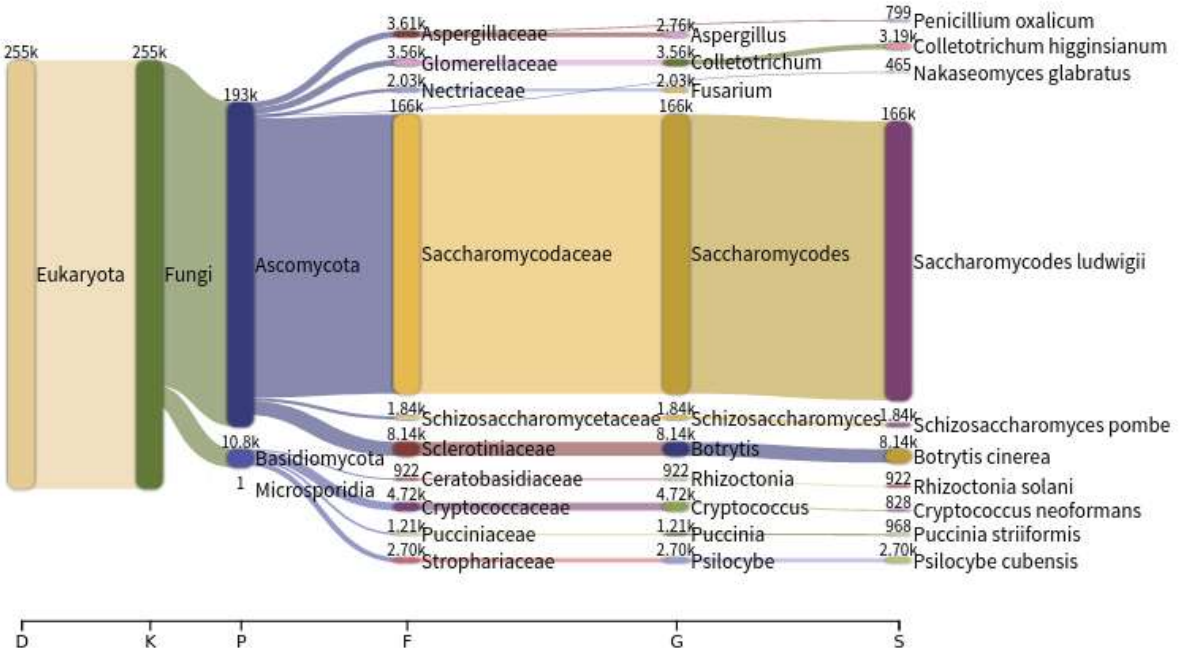

Figure 6. Detailed taxonomic identification of sample Hovby using Kraken2 database. Numbers on the top of the tree and each branch represent the number of reads for each group. Taxonomic groups are represented by letters D-domain, K-kingdom, P-phylum, F-family, G-genus and S-specie.

For all the samples the most abundant phyla and species identified, using Kraken2 database, are shown in Table 5.

Table 5. Taxonomic identification of each sample using Kraken2 database.

Sample	No. of reads obtained	No. of most abundant phyla (reads)	No. of most abundant genus/species (reads)
Dala	172,032	Ascomycota (126,113)	*Saccharomycodes ludwigii* (95,324) *Botrytis cinerea* (6,852) *Schizosaccharomyces pombe* (4,243)
		Basidiomycota (18,608)	*Psilocybe cubensis* (7,444) *Cryptococcus neoformans* (1,413) *Rhizoctonia solani* (1,252)
Greväck	335,101	Ascomycota (228,275)	*Saccharomycodes ludwigii* (137,421) *Botrytis cinerea* (40,5143) *Colletotrichum higginsianum* (10,134)
		Basidiomycota (49,931)	*Psilocybe cubensis* (18,421) *Sporisorium graminicola* (7,771) Puccinia striiformis (1,752)
Hovby	254,701	Ascomycota (193,004)	*Saccharomycodes ludwigii* (166,421) *Botrytis cinerea* (8,142) *Colletotrichum higginsianum* (3,195)
		Basidiomycota (10,812)	*Psilocybe cubensis* (2,708) *Puccinia striiformis* (968) *Rhizoctonia solani* (922)

For all the samples the most abundant fungi phyla and species identified, using the UNITE database, are shown in Table 6.

Table 6. Taxonomic identification of fungi in each sample using UNITE database.

Sample	No. of reads obtained	No. of most abundant phyla (reads)	No. of most abundant genus/species (reads)
Dala	241,169	Ascomycota (25,327)	*Pyrenopeziza brassicae* (7,542) *Plenodomus dezfulensis* (3,971)
		Basidiomycota (24,133)	*Filobasidium stepposum* (6,253)
Greväck	455,941	Ascomycota (85,521)	*Pyrenopeziza brassicae* (46,623) *Ophiocordyceps xueferengensis* (10,743)
		Basidiomycota (37,741)	*Leucosporidium drummii* (5,174)
Hovby	384,015	Ascomycota (32,024)	*Pyrenopeziza brassicae* (11,622) *Ophiocordyceps xueferengensis* (11,231)
		Basidiomycota (16,220). Only genus identified in this sample, no species identified	*Pyrenopeziza* (11,604) *Holtermanniella* (5,091)

It was not possible to find *S. sclerotiorum* in none of the samples when using the Kraken2 and UNITE databases to analyze the sequences obtained after Nanopore Sequencing. However, many other different species of fungi where identified.

Plasmid digestion and purification

Electrophoresis of digested plasmid DNA showed a clear band between 4-4.5 kb that corresponded with the total length of pCR4 - TOPO® vector containing target insert for digested samples (4234 bp). Undigested samples showed bands of different sizes corresponding to different forms of undigested plasmid like supercoiled, nicked and linear (data not shown). After confirming the digestion on the gel, the digested sample was cleaned, then purity and concentration were measured using Nanodrop 2000 spectrophotometer (Thermo scientific) and Qubit Fluorometer (Thermo Fisher Scientific), results are shown in Appendix 2, Table 3.

qPCR Assay

Some problems occurred during qPCR assays and affected the reactions performance. There was no amplification of *S. sclerotiorum* DNA from any of the field samples (Dala, Greväck and Hovby), in none of the diverse types of samples (individual, combined and pooled), which goes in accordance with the sequencing results. However, the plasmid DNA dilutions were amplified as expected and it was possible to create a standard curve with almost optimal parameters (Figure 7). The R^2 value was 0.969, the efficiency was 110.84 and the value of the slope was -3.087. Also, amplification was detected in the No Template Control (NTC) samples. Contamination was detected in the NTC's during all the different qPCR assays and the Cq values among the replicates showed very little variation (Cq 34-36).

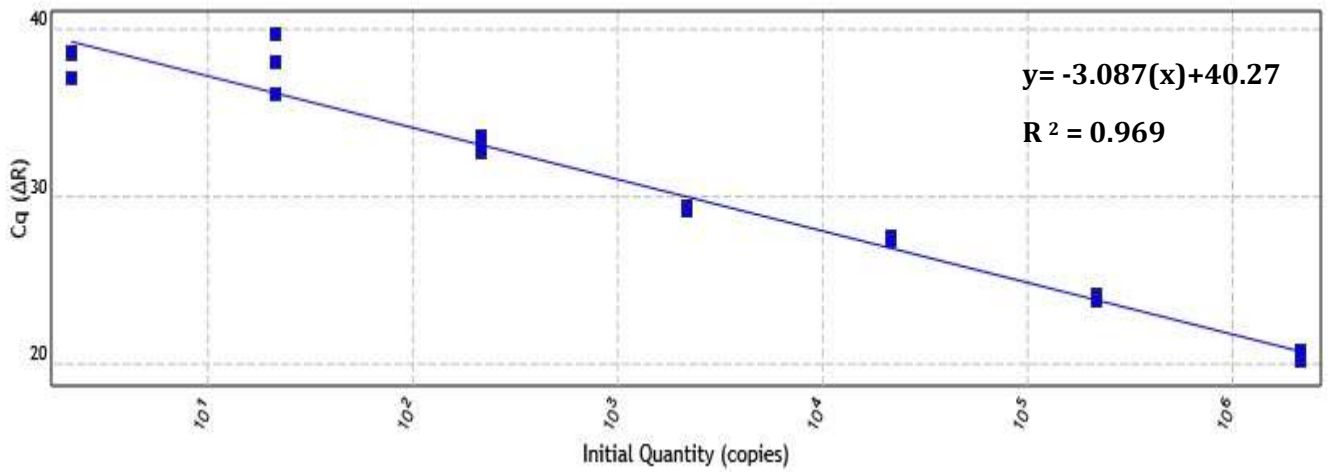

Figure 7. Quantification curves from 10-fold plasmid DNA serial dilutions, with *S. sclerotiorum* target, generated for quantification. Linear regression was generated by plotting threshold cycle values (Cq), obtained in qPCR assay, and the logarithm of total DNA copies in each plasmid DNA dilution. Correlation value (R2) was 0.969. Results are shown in triplicates.

After the previous results, it was decided to spike one of the triplicates of each one of the different types of samples from the different fields with a known number of DNA copies, 1000 copies from one of the standards. Again, in this qPCR assay there was no amplification detected in any of the samples, not even the spiked samples, after adjusting the threshold in the amplification plots. This led to the conclusion that there was presence of inhibitors in the DNA samples from the fields, since the dilutions from the plasmid DNA were amplified without problems. In the spiked qPCR essay the standard curve presented also good parameters. The R^2 value was 0.976, the efficiency was 109.82 and the value of the slope was -3.107. Amplification was detected in negative controls again. Some of the spiked samples showed a Cq value of more than 34 cycles, which didn't correspond to any point on the standard curve (Figure 8). Regardless of the lack of proper amplification, it was decided to quantify and determine the number of copies present in the spiked samples (Appendix 2, Table 8), the highest number of copies detected was in the spiked individual sample number 1 from the field Greväck having 1.10 copies.

These results reinforced the theory of inhibitors present in the DNA samples from the fields. Due to the lack of more volume of DNA samples, it was not possible to continue with the qPCR assays. More time and resources were needed to do new DNA extractions from new leave samples.

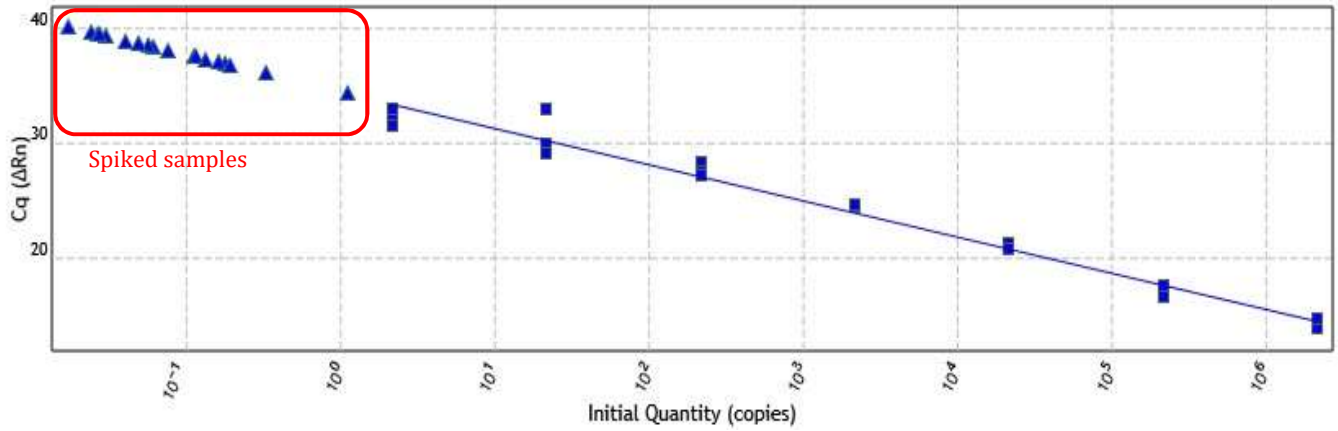

Figure 8. Quantification curves from 10-fold plasmid DNA serial dilutions, with *S. sclerotiorum* target, generated for quantification. This figure shows how the spiked samples had a Cq value of more than 34, which didn't correspond to any point on the standard curve. Results are shown in triplicates.

Discussion

The results from this experiment were not as expected, the pathogen *S. sclerotiorum* was not identified in any of the samples collected from fields with known incidence of SSR disease by any of the methods used, qPCR and Oxford Nanopore Sequencing. According to Ficke et al. (2018), first the sclerotia needs to germinate in the soil. This requires a high soil moisture content and a certain temperature range and once the sclerotia have germinated and the ascospores have been released, they can attach to tissue of its crops. But only under high humidity (minimum 80–86%) or in the presence of free water, and at least 15 °C, and nutritious tissue, the ascospores will be able to germinate, penetrate and start the infection. *S. sclerotiorum* can be identified through qPCR in this stage (Almquist & Wallenhammar, 2014). From the infected tissue, the pathogen can spread to the stem causing the characteristic symptoms of SSR in rapeseed crops and the symptoms can be detected after the flowering period (Ficke et al., 2018).

Rapeseed is most susceptible to SSR during flowering (July-August) and targeting fungicide application at the right bloom stage is essential for the maximum effectiveness of fungicides. Fungicide timing treatments that were applied between 10 and 50% bloom were more successful than those applied sooner or later (Khangura & van Burgel, 2021).

The weeks prior to the collection of the samples from the field were not ideal for the disease to develop. The humidity level in the south region of Sweden, where the fields are located, was 69% on average, being May reported as the driest month of the year 2022 and sclerotia needs 80–86% humidity. With only 0.9 mm of precipitation, the least amount of rain falls was during April 2022 (AccuWeather, 2022), sclerotia needs the presence of free water (rain). Due to these unfavorable conditions, perhaps the pathogen was not present at the moment of the samples collection and that was the reason why the pathogen was not identified by any of the techniques used in this study. However, the symptoms of the disease were present in the fields after flowering during the same year the samples were collected, which indicates that maybe the spores were released after the samples collection.

Other issues to address were the problems faced during the performance of qPCR essays. The first issue was the presence of inhibitors in the DNA samples extracted from the leaves collected from the fields. In this study, A260/A230 ratios were very low, lower than 1 in all the samples (Appendix 2), the desired A260/A230 ratios are 2.0-2.2 and are considered pure and free of different inhibitors (Demeke & Jenkins, 2009). A260/230 ratios below 1.8 could indicate the presence of different inhibitors such as polysaccharides, humic acid and salts (Olson & Morrow, 2012). These samples were used even though A260/230 ratios were very low, this decision was made based on the results obtained by Ning et al. (2009), where A260/A230 ratios even below 0.5 did not affect qPCR amplification in the case of DNA extracted from soil samples. Another factor considered in this decision was that, during purity measurement using Nanodrop, an alert suggesting that A260/A230 could be inaccurate due to the very low concentration of the samples, therefore it was difficult to know if the A260/A230 ratios were reliable.

Efficient in vitro DNA polymerization during qPCR, demands high DNA polymerase activity as well as favorable interactions between nucleic acids (target denaturation and primer annealing), meaning that both biochemical and biophysical processes are involved. Any compound affecting any of the critical reagents or the sub reactions in the polymerization process acts as an inhibitor (Sidstedt et al., 2020). Inhibitory compounds can cause a shift, or increase, in the quantification cycle (Cq), that is, the number of temperature cycles at which the target nucleic acid has been amplified enough to reach a defined threshold. Inhibitory substances primarily disrupt the amplification of target nucleic acids by interfering with the extraction of nucleic acid, through the

degradation or capture of nucleic acid and by inhibiting qPCR amplification (e.g., interfering with polymerases) (Gibson et al., 2012).

Also, the inhibitory mode of action of some of the compounds may be linked with denaturation of DNA or the ability of the polymerase enzyme to bind to magnesium ions. Inhibitors may kinetically modify PCR amplification by chelating Mg^{2+}, a cofactor for all DNA polymerases, or by binding to target DNA or the DNA polymerase. Inhibitors may originate from either the plant tissue or the reagents used for DNA isolation (Demeke & Jenkins, 2009). Wang et al. (2017) also reports that inhibitory components such as humic substances are co-extracted along with nucleic acids during DNA extractions. Plants carry many substances, such as polysaccharides, polyphenols, pectin and xylan, which may be co-extracted and thereafter hamper qPCR assays (Ma & Michailides, 2007; Minas et al., 2011; Schrader et al., 2012).

To effectively detect pathogens in infected plant tissues, in soil, air, or water samples, suitable DNA extraction procedures are necessary in order to eliminate PCR inhibitors released from these samples. DNA must be extracted from infected plant tissue to effectively detect small amount of fungal DNA in these samples. Commercial DNA extraction kits are able to remove most PCR inhibitors efficaciously, but not all inhibitors in some cases (Ma & Michailides, 2007). The consequences of this include decreased detection accuracy (when estimating copy numbers of target DNA fragments) or increased numbers of false negatives (when estimating presence/absence). Even a slight inhibition during qPCR has been demonstrated to cause a substantial shift in quantified gene copy numbers in environmental samples (McKee et al., 2015).

Fluorescence measurement is the main means to detect and quantify nucleic acids in DNA in qPCR analysis, for these fluorophores attached to primers or nucleotides are essential in and any substance that disturbs the function of the fluorophore will impair the analysis. Besides the actual amplicon generation, the detection of amplicons through fluorescence measurements must function properly. Standard PCR analysis has been proven to be less affected by inhibitors than qPCR. For qPCR quantification cycle (Cq) values are linked to a standard curve and any inhibition effect skewing the Cq values will directly affect quantification. Possibly, the partitioning of the samples into many minutes reactions, during standard PCR, play a role in the elevated resistance to inhibitors due to less interaction between inhibitor molecules and the molecules involved in the polymerization process (Sidstedt et al., 2020).

There are multiple mechanisms to eliminate inhibitors and clean DNA. One of the simplest ways is by dilution, in this case inhibitors are not eliminated but diluting the DNA sample can help to reduce the concentration of inhibitors, decreasing their effect. However, this may also decrease the sensitivity of the qPCR reaction (Gibson et al., 2012; Schneider et al., 2009). Moreover, inhibition levels are very dependent on the concentration of the target gene in the DNA extracts. Stronger inhibition could happen when quantifying genes with low copy numbers compared to those with high copy numbers. For genes with low copy numbers, dilution may further lead to increases in gene susceptibility to inhibitors (Wang et al., 2017). It was not possible to dilute the samples in this study due to the very low concentrations of the extracted samples (as expected). Although, there are other methods that could have been perform with more time and resources, the most common are explained bellow.

A second DNA purification step after DNA isolation is another method to eliminate inhibitors. Using commercial kits can remove impurities and inhibitors from the sample. These kits use different methods, such as column-based or bead-based, to purify DNA. The cetyltrimethylammonium bromide (CTAB) method has been widely used for extraction of DNA from leaves, seeds/grains and processed food/feed samples and CTAB-extracted DNA needs further purification to be used for real-time PCR (Demeke & Jenkins, 2009). According to the experiment done by Demeke et al. (2009), the A 260/A 230 ratio of CTAB-extracted DNA extracted from soybean was low (average of 1.29), indicative of contamination with compounds such as

polysaccharides which may inhibit qPCR and further purification of the soybean DNA with either the Zymo kit (Zymo Research) or Qtip 100 (Qiagen) provided clean DNA that was suitable for real-time PCR. Therefore, the low 260/A 230 obtained in the present experiment could indicate contamination with polysaccharides. CTAB was used in combination with the E.Z.N.A kit for the DNA extraction from the leaves, so a second DNA purification kit was necessary, based on the 260/A 230 ratios obtained in this experiment. Demeke et al. (2009) reports that purification of CTAB-extracted rapeseed DNA with the Zymo kit produced clean DNA that was suitable for qPCR.

Another method is ethanol precipitation, is a simple and cost-effective method to remove inhibitors from DNA (Gibson et al., 2012; Ma & Michailides, 2007). Other method is gel filtration, it involves passing the DNA sample through a column packed with a gel matrix that separates DNA from contaminants and inhibitors based on size (Gibson et al., 2012). Schrader et al. (2012) reports other methods to clean DNA, such as column chromatography to remove polysaccharides from many sample matrices. Cetrimonium bromide effectively eliminates polysaccharides and denatured proteins from the preparation by forming an insoluble combination with them. A repeated extraction using silica columns could also remove inhibitors. In addition, Sipahioglu et al. (2006) reached the removal of phenols from leaves by drying at 65°C for 2 days and conserving them under hermetic conditions at 4°C. According to Samarakoon et al. (2013) polysaccharides and phenolic compounds are common inhibitors found in plant tissues and possibly these were the inhibitors affecting the samples.

The reason why no further DNA cleaning was performed, was that the presence of inhibitors problem was taking in consideration too late into the experiment and at this point there was no more DNA left from the original samples to clean and performed another qPCR assay. Another DNA extraction kit and more time would have been needed to do new DNA extractions from new leaves coming from the same fields.

Another issue during this study was the problem with DNA contamination. Amplification was detected in the NTC triplicates in all of the qPCR assays. NTC contamination is a common problem that can lead to inaccurate and unreliable results (Nishikawa et al., 2015). Cross-contamination can occur during the preparation of samples or when pipetting the template DNA into the reaction mix (Nishikawa et al., 2015). Another source of contamination could be aerosols, contamination can occur when opening tubes, this can release tiny droplets into the air (Corless et al., 2000). Contaminated reagents, for example in this case, qPCR mastermix, primers or probes contaminated with DNA can result in NTC contamination. Using sterile reagents and checking them for contamination can help prevent this (Corless et al., 2000; Nishikawa et al., 2015). Contaminated lab equipment can also be a source of NTC contamination (Corless et al., 2000; Nishikawa et al., 2015). Contamination in any of the previous procedures could have taken place at any moment in the laboratory during preparation of qPCR assays.

First it was suspected that the source of contamination was cross contamination or aerosols. To try to avoid this problem, it was decided to change labs, pipettes and all the lab equipment involve in the preparation of the qPCR assay, as well as the use of only filtered pipette tips. It is worth to mention that another measure that could have taken place was the use of a sterilized laminar airflow bench during the preparation of the qPCR assay, to further address the potential problem of contamination by aerosols. However, the NTC contamination remained, which led to suspect that one or more reagents were contaminated. The Cq values among NTC replicates during the qPCR assays showed very little variation, this reinforced the theory of the contaminated reagents. After arriving to this conclusions, and with more time and resources, the following steps would have been to test the reagents, discard contaminated reagents and use new ones.

The pathogen was not identified by Oxford Nanopore Sequencing either, again, possibly the pathogen was not present in the samples used for sequencing. The percentage of classified data was expected to be very similar between the two databases used as reference, since the same

sequences were used in both. Many fungal species, other than *S. sclerotiorum*, were identified and these species showed a big variation between databases. Sequencing was stopped after 96 hours since, due to the easter holidays, the access to the laboratory was restricted. Besides *S. sclerotiorum*, there are other important fungal pathogens species affecting rapeseed crops in Sweden. *Pyrenopeziza brassicae*, *Alternaria brassicae, Leptosphaeria maculans* and *Fusarium spp*. are some of the most relevant species. Among the species obtained, when using the reference database UNITE, the only pathogen identified out of the previous mentioned species was *Pyrenopeziza brassicae*, being identified in all the samples. So, it can be concluded that Oxford Nanopore Sequencing is suitable to identify *Pyrenopeziza brassicae,* being the causal agent of light leaf spot disease, which is an economically important fungal pathogen of oilseed rape and other *Brassica* species (Karandeni Dewage et al., 2021).

The species identified when analyzing the sequences with the two different reference databases were very inconsistent. When using the first reference database (Kraken2), only fungi were classified as expected since the database only contained fungal sequences from NCBI's Refseq (complete fungal genomes). However, when analyzing the sequences using the UNITE database the results were very different. The fungi species identified with two databases were not the same even though the same query sequences were used for both. Not only the fungi species were different, the UNITE database classified most of the sequences under the Viridiplantae kingdom, being Brassica napus (rapeseed plant) the biggest portion. Also, a very small portion of the sequences was classified under the kingdom Metazoa. These results were not expected at all since the UNITE database only contained fungal ITS sequences of the UNITE and INSDC databases (Abarenkov et al., 2022).

Cheng et al. (2015), reported that even though ITS has the highest discriminatory power of the candidate barcodes currently in use and being the most widely used in plant molecular systematics, ITS amplification and sequencing could occasionally be affected from non-specificity and poor PCR and sequencing success due to issues related with the primers. As fungi in many cases are symbiotic with plants in the environment, it is easy to obtain nontarget amplicons of ITS from fungi when amplifying plant ITS fragments (Cheng et al., 2015).

The biggest drawbacks to date, for nanopore sequencing, have been a lower throughput of sequence data and a higher error rate compared to other sequencing methods, like Ilumina. The results of analyzing nanopore sequences can vary depending on the reference database used for comparison. This is because the reference database used for analysis will affect the sensitivity and specificity of the analysis, which in turn can impact the accuracy and reliability of the results (Petersen et al., 2019). The first one was the Kraken2 database containing all fungal sequences from NCBI's Refseq (complete fungal genomes) and the second database was UNITE containing all fungal ITS sequences of the UNITE and INSDC databases. Also, according to Ciuffreda et al. (2021) reference databases can vary in terms of the genomes they represent. Some reference databases may include a more comprehensive set of genomes, while others may only include a limited number of genomes. This can lead to differences in the coverage and accuracy of the results obtained. In this study the most updated available versions were used.

The way different bioinformatics pipelines influence data analysis is very complex. However, some of the other possible reasons why this big variation in the results was present when using different reference databases are suggested next. Another source of variation could be the differences in quality, as mentioned by Ciuffreda et al. (2021) reference databases can also vary in terms of their quality. In this study, even though the two databases used are considered to be very high quality, they are still different databases which already introduce a source of variation.

The Genetic Variation could be another factor. Different reference databases may also have differences in genetic variation, such as single nucleotide polymorphism (SNP), which is a genomic variant at a single base position in the DNA, small InDels and/or more complex structural variants such as large deletions, duplications, inversions and translocations, all can impact the accuracy of

the analysis, some reference databases may be more comprehensive in terms of representing genetic variation, while others may be less so (Magi et al., 2017). Additionally, some genomes may be better assembled than others, resulting in more accurate results (Phannareth et al., 2021). Again, this study used two different databases which introduces a variation factor.

Because this approach relies on sequences from previously characterized organisms for comparison, the quality of the results is only as good as the existing reference data set. As such, the lack of sequence data from identical or closely related taxa can lower the probability of an accurate identification (Coissac et al., 2016).

Regarding the analysis of metagenomic data, improvements are expected in taxonomic analysis and sequence comparison software to achieve better resolution of closely related strains and higher classification accuracy, along with the development of efficient indexing techniques for metagenomics databases (Ciuffreda et al., 2021).

Conclusion

The choice of reference database can have a significant impact on the results obtained from nanopore sequencing analysis. It is important to carefully consider the properties of different reference databases and choose the one that best matches the specific goals of the analysis. Also, further analysis and validation of the accuracy of nanopore sequencing reads could be achieved by the inclusion of parallel sequencing methods with higher accuracy rate like Illumina.

As a consequence of the poor results obtained during the qPCR assays, it was not possible to evaluate if there is a significant difference between individual, combined and pooled samples. With more time and resources, perhaps, it would have been possible to obtained results leading to better and more accurate conclusions.

Ethical aspects and impact on the society

There are no ethical aspects to board on this study, *Sclerotinia sclerotiorum* is a common plant fungal pathogen with a broad host range and not known conflicts of interest were identified during this study. All the biological material used in this experiment was collected from different fields in Sweden with known incidence of the disease. No personal information was attached to any of the samples in this study. No personal bias was introduced during the performance of this experiment and all the conclusions were made based on available scientific articles.

However, an aspect worth to mention is that due to the limited preventive and managing measures of this pathogen, many farmers usually decide to apply fungicides when perhaps it is not necessary. This use of unnecessary chemicals have a negative impact on the environment and the local flora and fauna, so it should be avoided when possible. Another concern are the long term effects these fungicides have over human health. Many articles describe the toxic and the negative physiological effects as consequence of the ingestion of these products. It is a collective responsibility to make an effort to move towards more sustainable agriculture practices. Improving the detection of the pathogen in the crops using different techniques, as well as to make them more accessible and cheaper, could help to make better and more precise decisions about the disease management, to avoid economic losses and increase the production of rapeseed oil. This could also lower the cost of the production of rapeseed oil, since unnecessary use of fungicides could be avoided, and consequently the price of the oil would also reduce in the market, which could also reduce the costs of products that include rapeseed oil in their manufacturing and perhaps make them more accessible to larger portions of the population. Although third generation sequencing technologies have been available for a few years now and are very promising with multiple advantages over other techniques, like the reduction in costs, is still not widely used in many biological fields. As these technologies get used more frequent and become

more popular among the scientific community, the tools to analyze this kind of data will also improve and it will become more accurate.

The results from this experiment could give a better perspective on how to handle certain obstacles arising during the performance of similar studies and how to improve the management of time and resources.

Future perspectives

To discover if a significant difference exists between individual, combined and pooled samples, more qPCR assays will have to be performed, handling the inhibitors issue since the beginning of the experiment, doing a second DNA purification step after DNA isolation. If there is a significant difference between these kinds of samples, specially between individual and combines samples, time and resources could be saved and a larger number of samples could be analyzed at the same time. Furthermore, if Oxford Nanopore Sequencing is proven to be a reliable method to identify *S. sclerotiorum* from field samples, it could be a very good diagnostic alternative to qPCR, since this method needs less equipment and could reduce some costs. Also, the ability to spend minimal time to prepare a library, load it into the device, and produce large amounts of sequence data, is a great opportunity for diagnostic screening. In future studies, using Nanopore Sequencing to detect *S. sclerotiorum*, the use of a positive control could be very useful in order to discover if this technique is suitable for the detection of the pathogen. However, the principal concern using this technology is that the current Oxford Nanopore Sequencing has relatively higher error rates compared to other sequencing methods, overcoming these challenges will require further breakthroughs in nanopore technology and bioinformatics software, as well as the use of standardized databases and validated thresholds for reporting detected pathogens, will be of crucial importance.

Acknowledgements

I would like to thank to my thesis supervisor Maria Algerin for her help and guidance throughout this project. Also, my lab supervisor Johan Nordén, thank you. A distinctive thank you to my co-supervisor John Baxter for his patience and all the help provided throughout my studies at Skövde University, his support and kindness go above and beyond. Thank you to my examiner Magnus Fagerlind and his constructive criticism. Finally, most of all I want to thank you my family, their love and support in every step of my way are absolutely invaluable.

References

Abarenkov, Kessy; Zirk, Allan; Piirmann, Timo; Pöhönen, Raivo; Ivanov, Filipp; Nilsson, R.Henrik; Kõljalg, Urmas (2022): Full UNITE+INSD dataset for Fungi. UNITE Community. 10.15156/BIO/2483925.

Almquist, C., & Wallenhammar, A.-C. . (2014). Monitoring of plant and airborne inoculum ofSclerotinia sclerotiorumin spring oilseed rape using real-time PCR. *Plant Pathology*, *64*(1), 109–118. https://doi.org/10.1111/ppa.12230.

Athanasopoulou, K., Boti, M. A., Adamopoulos, P. G., Skourou, P. C., & Scorilas, A. (2021). Third-Generation Sequencing: The Spearhead towards the Radical Transformation of Modern Genomics. *Life*, *12*(1), 30. https://doi.org/10.3390/life12010030

Austin-Tse, C. A., Jobanputra, V., Perry, D. L., Bick, D., Taft, R. J., Venner, E., Gibbs, R. A., Young, T., Barnett, S., Belmont, J. W., Boczek, N., Chowdhury, S., Ellsworth, K. A., Guha, S., Kulkarni, S., Marcou, C., Meng, L., Murdock, D. R., Rehman, A. U., & Spiteri, E. (2022). Best practices for the interpretation and reporting of clinical whole genome sequencing. *Npj Genomic Medicine*, *7*(1), 1–13. https://doi.org/10.1038/s41525-022-00295-z

Avery, S. V., Singleton, I., Magan, N., & Goldman, G. H. (2019). The fungal threat to global food security. *Fungal Biology*, *123*(8), 555–557. https://doi.org/10.1016/j.funbio.2019.03.006

Badotti, F., de Oliveira, F. S., Garcia, C. F., Vaz, A. B. M., Fonseca, P. L. C., Nahum, L. A., Oliveira, G., & Góes-Neto, A. (2017). Effectiveness of ITS and sub-regions as DNA barcode markers for the identification of Basidiomycota (Fungi). *BMC Microbiology*, *17*(1). https://doi.org/10.1186/s12866-017-0958-x

Bellemain, E., Carlsen, T., Brochmann, C., Coissac, E., Taberlet, P., & Kauserud, H. (2010). ITS as an environmental DNA barcode for fungi: an in silico approach reveals potential PCR biases. *BMC Microbiology*, *10*(1), 189. https://doi.org/10.1186/1471-2180-10-189

Bollmann-Giolai, A., Malone, J. G., & Arora, S. (2022). Diversity, detection and exploitation: linking soil fungi and plant disease. *Current Opinion in Microbiology*, *70*, 102199. https://doi.org/10.1016/j.mib.2022.102199

Boulter, N., Suarez, F. G., Schibeci, S., Sunderland, T., Tolhurst, O., Hunter, T., Hodge, G., Handelsman, D., Simanainen, U., Hendriks, E., & Duggan, K. (2016). A simple, accurate and universal method for quantification of PCR. *BMC Biotechnology*, *16*(1). https://doi.org/10.1186/s12896-016-0256-y

Chalupowicz, L., Dombrovsky, A., Gaba, V., Luria, N., Reuven, M., Beerman, A., Lachman, O., Dror, O., Nissan, G., & Manulis-Sasson, S. (2018). Diagnosis of plant diseases using the Nanopore sequencing platform. *Plant Pathology*, *68*(2), 229–238. https://doi.org/10.1111/ppa.12957

Cheng, T., Xu, C., Lei, L., Li, C., Zhang, Y., & Zhou, S. (2015). Barcoding the kingdom Plantae: new PCR primers forITSregions of plants with improved universality and specificity. *Molecular Ecology Resources*, *16*(1), 138–149. https://doi.org/10.1111/1755-0998.12438

Ciuffreda, L., Rodríguez-Pérez, H., & Flores, C. (2021). Nanopore sequencing and its application to the study of microbial communities. *Computational and Structural Biotechnology Journal*, *19*, 1497–1511. https://doi.org/10.1016/j.csbj.2021.02.020

Coissac, E., Hollingsworth, P. M., Lavergne, S., & Taberlet, P. (2016). From barcodes to genomes: extending the concept of DNA barcoding. *Molecular Ecology*, *25*(7), 1423–1428. https://doi.org/10.1111/mec.13549

Corless, C. E., Guiver, M., Borrow, R., Edwards-Jones, V., Kaczmarski, E. B., & Fox, A. J. (2000). Contamination and Sensitivity Issues with a Real-Time Universal 16S rRNA PCR. *Journal of Clinical Microbiology*, *38*(5), 1747–1752. https://doi.org/10.1128/jcm.38.5.1747-1752.2000

Crossley, B. M., Bai, J., Glaser, A., Maes, R., Porter, E., Killian, M. L., Clement, T., & Toohey-Kurth, K. (2020). Guidelines for Sanger sequencing and molecular assay monitoring. *Journal of Veterinary Diagnostic Investigation*, *32*(6), 767–775. https://doi.org/10.1177/1040638720905833

del Río, L. E., Bradley, C. A., Henson, R. A., Endres, G. J., Hanson, B. K., McKay, K., Halvorson, M., Porter, P. M., Le Gare, D. G., & Lamey, H. A. (2007). Impact of Sclerotinia Stem Rot on Yield of Canola. *Plant Disease*, *91*(2), 191–194. https://doi.org/10.1094/pdis-91-2-0191

Demeke, T., & Jenkins, G. R. (2009). Influence of DNA extraction methods, PCR inhibitors and quantification methods on real-time PCR assay of biotechnology-derived traits. *Analytical*

and Bioanalytical Chemistry, *396*(6), 1977–1990. https://doi.org/10.1007/s00216-009-3150-9

Demeke, T., Ratnayaka, I., & Phan, A. (2009). Effects of DNA Extraction and Purification Methods on Real-Time Quantitative PCR Analysis of Roundup Ready Soybean. *Journal of AOAC INTERNATIONAL*, *92*(4), 1136–1144. https://doi.org/10.1093/jaoac/92.4.1136

Dentinger, B. T. M., Didukh, M. Y., & Moncalvo, J.-M. (2011). Comparing COI and ITS as DNA Barcode Markers for Mushrooms and Allies (Agaricomycotina). *PLoS ONE*, *6*(9), e25081. https://doi.org/10.1371/journal.pone.0025081

Derbyshire, M. C., & Denton-Giles, M. (2016). The control of sclerotinia stem rot on oilseed rape (Brassica napus): current practices and future opportunities. *Plant Pathology*, *65*(6), 859–877. https://doi.org/10.1111/ppa.12517

Dhanasekaran, S., Doherty, T. M., & Kenneth, J. (2010). Comparison of different standards for real-time PCR-based absolute quantification. *Journal of Immunological Methods*, *354*(1-2), 34–39. https://doi.org/10.1016/j.jim.2010.01.004

Ding, L.-N., Li, T., Guo, X.-J., Li, M., Liu, X.-Y., Cao, J., & Tan, X.-L. (2021). Sclerotinia Stem Rot Resistance in Rapeseed: Recent Progress and Future Prospects. *Journal of Agricultural and Food Chemistry*, *69*(10), 2965–2978. https://doi.org/10.1021/acs.jafc.0c07351

Feng, Y., Hu, Y., Fang, P., Zuo, X., Wang, J., Li, J., Qian, W., & Mei, J. (2021). Silicon Alleviates the Disease Severity of Sclerotinia Stem Rot in Rapeseed. *Frontiers in Plant Science*, *12*. https://doi.org/10.3389/fpls.2021.721436

Ficke, A., Grieu, C., Brurberg, M. B., & Brodal, G. (2018). The role of precipitation, and petal and leaf infections in Sclerotinia stem rot of spring oilseed Brassica crops in Norway. *European Journal of Plant Pathology*, *152*(4), 885–900. https://doi.org/10.1007/s10658-018-1569-6

Fones, H. N., Bebber, D. P., Chaloner, T. M., Kay, W. T., Steinberg, G., & Gurr, S. J. (2020). Threats to global food security from emerging fungal and oomycete crop pathogens. *Nature Food*, *1*(6), 332–342. https://doi.org/10.1038/s43016-020-0075-0

Freeman, J., Ward, E., Calderon, C., & McCartney, A. (2002). A Polymerase Chain Reaction (PCR) Assay for the Detection of Inoculum of Sclerotinia sclerotiorum . *European Journal of Plant Pathology*, *108*(9), 877–886. https://doi.org/10.1023/a:1021216720024

Gibson, K. E., Schwab, K. J., Spencer, S. K., & Borchardt, M. A. (2012). Measuring and mitigating inhibition during quantitative real time PCR analysis of viral nucleic acid extracts from large-volume environmental water samples. *Water Research*, *46*(13), 4281–4291. https://doi.org/10.1016/j.watres.2012.04.030

Kamal, M. M., Lindbeck, K. D., Savocchia, S., & Ash, G. J. (2015). Biological control of sclerotinia stem rot of canola using antagonistic bacteria. *Plant Pathology*, *64*(6), 1375–1384. https://doi.org/10.1111/ppa.12369

Kang, S., Mansfield, M. A., Park, B., Geiser, D. M., Ivors, K. L., Coffey, M. D., Grünwald, N. J., Martin, F. N., Lévesque, C. A., & Blair, J. E. (2010). The Promise and Pitfalls of Sequence-Based Identification of Plant-Pathogenic Fungi and Oomycetes. *Phytopathology®*, *100*(8), 732–737. https://doi.org/10.1094/phyto-100-8-0732

Karamitros, T., & Magiorkinis, G. (2018). Multiplexed Targeted Sequencing for Oxford Nanopore MinION: A Detailed Library Preparation Procedure. *Methods in Molecular Biology (Clifton, N.J.)*, *1712*, 43–51. https://doi.org/10.1007/978-1-4939-7514-3_4

Karandeni Dewage, C. S., Qi, A., Stotz, H. U., Huang, Y., & Fitt, B. D. L. (2021). Interactions in the Brassica napus – Pyrenopeziza brassicae pathosystem and sources of resistance to P. brassicae (light leaf spot). *Plant Pathology*, *70*(9), 2104–2114. https://doi.org/10.1111/ppa.13455

Khangura, R., & van Burgel, A. J. (2021). Foliar fungicides and their optimum timing reduce sclerotinia stem rot incidence, improve yield and profitability in canola (Brassica napus L.). *Indian Phytopathology*, *74*(2), 549–558. https://doi.org/10.1007/s42360-021-00321-7

Kono, N., & Arakawa, K. (2019). Nanopore sequencing: Review of potential applications in functional genomics. *Development, Growth & Differentiation*, *61*(5), 316–326. https://doi.org/10.1111/dgd.12608

Kutcher, H. R., & Wolf, T. M. (2006). Low-drift fungicide application technology for sclerotinia stem rot control in canola. *Crop Protection*, *25*(7), 640–646. https://doi.org/10.1016/j.cropro.2005.09.003

Loit, K., Adamson, K., Bahram, M., Puusepp, R., Anslan, S., Kiiker, R., Drenkhan, R., & Tedersoo, L. (2019). Relative Performance of MinION (Oxford Nanopore Technologies) versus Sequel (Pacific Biosciences) Third-Generation Sequencing Instruments in Identification of Agricultural and Forest Fungal Pathogens. *Applied and Environmental Microbiology*, *85*(21). https://doi.org/10.1128/aem.01368-19

Lu, H., Giordano, F., & Ning, Z. (2016). Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics, Proteomics & Bioinformatics*, *14*(5), 265–279. https://doi.org/10.1016/j.gpb.2016.05.004

Lu, J., & Salzberg, S. L. (2020). Ultrafast and accurate 16S rRNA microbial community analysis using Kraken 2. *Microbiome*, *8*(1). https://doi.org/10.1186/s40168-020-00900-2

Ma, Z., & Michailides, T. J. (2007). Approaches for eliminating PCR inhibitors and designing PCR primers for the detection of phytopathogenic fungi. *Crop Protection*, *26*(2), 145–161. https://doi.org/10.1016/j.cropro.2006.04.014

Magi, A., Semeraro, R., Mingrino, A., Giusti, B., & D'Aurizio, R. (2017). Nanopore sequencing data analysis: state of the art, applications and challenges. *Briefings in Bioinformatics*. https://doi.org/10.1093/bib/bbx062

Marcelino, V. R., Clausen, P. T. L. C., Buchmann, J. P., Wille, M., Iredell, J. R., Meyer, W., Lund, O., Sorrell, T. C., & Holmes, E. C. (2020). CCMetagen: comprehensive and accurate identification of eukaryotes and prokaryotes in metagenomic data. *Genome Biology*, *21*(1). https://doi.org/10.1186/s13059-020-02014-2

Martin, K. J., & Rygiewicz, P. T. (2005). Fungal-specific PCR primers developed for analysis of the ITS region of environmental DNA extracts. *BMC Microbiology*, *5*(1), 28. https://doi.org/10.1186/1471-2180-5-28

Mbengue, M., Navaud, O., Peyraud, R., Barascud, M., Badet, T., Vincent, R., Barbacci, A., & Raffaele, S. (2016). Emerging Trends in Molecular Interactions between Plants and the Broad Host

Range Fungal Pathogens Botrytis cinerea and Sclerotinia sclerotiorum. *Frontiers in Plant Science*, *7*. https://doi.org/10.3389/fpls.2016.00422

McKee, A. M., Spear, S. F., & Pierson, T. W. (2015). The effect of dilution and the use of a post-extraction nucleic acid purification column on the accuracy, precision, and inhibition of environmental DNA samples. *Biological Conservation*, *183*, 70–76. https://doi.org/10.1016/j.biocon.2014.11.031

Minas, K., McEwan, N. R., Newbold, C. J., & Scott, K. P. (2011). Optimization of a high-throughput CTAB-based protocol for the extraction of qPCR-grade DNA from rumen fluid, plant and bacterial pure cultures. *FEMS Microbiology Letters*, *325*(2), 162–169. https://doi.org/10.1111/j.1574-6968.2011.02424.x

Nilsson, R. H., Larsson, K.-H., Taylor, A. F., Bengtsson-Palme, J., Jeppesen, T. S., Schigel, D., Kennedy, P., Picard, K., Glöckner, F. O., Tedersoo, L., Saar, I., Kõljalg, U., & Abarenkov, K. (2019). The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Research*, *47*(Database issue), D259–D264. https://doi.org/10.1093/nar/gky1022

Nilsson, R. H., Tedersoo, L., Ryberg, M., Kristiansson, E., Hartmann, M., Unterseher, M., Porter, T. M., Bengtsson-Palme, J., Walker, D. M., de Sousa, F., Gamper, H. A., Larsson, E., Larsson, K.-H., Kõljalg, U., Edgar, R. C., & Abarenkov, K. (2015). A Comprehensive, Automatically Updated Fungal ITS Sequence Dataset for Reference-Based Chimera Control in Environmental Sequencing Efforts. *Microbes and Environments*, *30*(2), 145–150. https://doi.org/10.1264/jsme2.me14121

Ning, J., Liebich, J., Kästner, M., Zhou, J., Schäffer, A., & Burauel, P. (2009). Different influences of DNA purity indices and quantity on PCR-based DGGE and functional gene microarray in soil microbial community study. *Applied Microbiology and Biotechnology*, *82*(5), 983–993. https://doi.org/10.1007/s00253-009-1912-0

Nishikawa, Y., Hosokawa, M., Maruyama, T., Yamagishi, K., Mori, T., & Takeyama, H. (2015). Monodisperse Picoliter Droplets for Low-Bias and Contamination-Free Reactions in Single-Cell Whole Genome Amplification. *PLOS ONE*, *10*(9), e0138733–e0138733. https://doi.org/10.1371/journal.pone.0138733

O'Brien, H. E., Parrent, J. L., Jackson, J. A., Moncalvo, J.-M. ., & Vilgalys, R. (2005). Fungal Community Analysis by Large-Scale Sequencing of Environmental Samples. *Applied and Environmental Microbiology*, *71*(9), 5544–5550. https://doi.org/10.1128/aem.71.9.5544-5550.2005

O'Sullivan, C. A., Belt, K., & Thatcher, L. F. (2021). Tackling Control of a Cosmopolitan Phytopathogen: Sclerotinia. *Frontiers in Plant Science*, *12*. https://doi.org/10.3389/fpls.2021.707509

Olson, N. D., & Morrow, J. B. (2012). DNA extract characterization process for microbial detection methods development and validation. *BMC Research Notes*, *5*(1). https://doi.org/10.1186/1756-0500-5-668

Peltier, A. J., Bradley, C. A., Chilvers, M. I., Malvick, D. K., Mueller, D. S., Wise, K. A., & Esker, P. D. (2012). Biology, Yield loss and Control of Sclerotinia Stem Rot of Soybean. *Journal of Integrated Pest Management*, *3*(2), 1–7. https://doi.org/10.1603/ipm11033

Petersen, L. M., Martin, I. W., Moschetti, W. E., Kershaw, C. M., & Tsongalis, G. J. (2019). Third-Generation Sequencing in the Clinical Laboratory: Exploring the Advantages and

Challenges of Nanopore Sequencing. *Journal of Clinical Microbiology*, *58*(1). https://doi.org/10.1128/jcm.01315-19

Phannareth, T., Nunziata, S. O., Stulberg, M. J., Galvez, M. E., & Rivera, Y. (2021). Comparison of Nanopore Sequencing Protocols and Real-Time Analysis for Phytopathogen Diagnostics. *Plant Health Progress*, *22*(1), 31–36. https://doi.org/10.1094/php-02-20-0013-rs

Pruitt, K., Murphy, T., Brown, G., & Murphy, M. (2020, January 31). *RefSeq Frequently Asked Questions (FAQ)*. Www.ncbi.nlm.nih.gov; National Center for Biotechnology Information (US).https://www.ncbi.nlm.nih.gov/books/NBK50679/#:~:text=Both%20Gene%20and%20RefSeq%20records

Ramakers, C., Ruijter, J. M., Deprez, R. H. Lekanne., & Moorman, A. F. M. (2003). Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience Letters*, *339*(1), 62–66. https://doi.org/10.1016/s0304-3940(02)01423-4

Samarakoon, T., Wang, S. Y., & Alford, M. H. (2013). Enhancing PCR Amplification of DNA from Recalcitrant Plant Specimens Using a Trehalose-Based Additive. *Applications in Plant Sciences*, *1*(1), 1200236. https://doi.org/10.3732/apps.1200236

Schneider, S., Enkerli, J., & Widmer, F. (2009). A generally applicable assay for the quantification of inhibitory effects on PCR. *Journal of Microbiological Methods*, *78*(3), 351–353. https://doi.org/10.1016/j.mimet.2009.06.010

Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., Chen, W., Bolchacova, E., Voigt, K., Crous, P. W., Miller, A. N., Wingfield, M. J., Aime, M. C., An, K.-D. ., Bai, F.-Y. ., Barreto, R. W., Begerow, D., Bergeron, M.-J. ., Blackwell, M., & Boekhout, T. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences*, *109*(16), 6241–6246. https://doi.org/10.1073/pnas.1117018109

Schrader, C., Schielke, A., Ellerbroek, L., & Johne, R. (2012). PCR inhibitors - occurrence, properties and removal. *Journal of Applied Microbiology*, *113*(5), 1014–1026. https://doi.org/10.1111/j.1365-2672.2012.05384.x

Seena, S., Pascoal, C., Marvanová, L., & Cássio, F. (2010). DNA barcoding of fungi: a case study using ITS sequences for identifying aquatic hyphomycete species. *Fungal Diversity*, *44*(1), 77–87. https://doi.org/10.1007/s13225-010-0056-y

Short, S. A., & Zehr, J. P. (2005). Quantitative Analysis of nifH Genes and Transcripts from Aquatic Environments. *Methods in Enzymology*, 380–394. https://doi.org/10.1016/s0076-6879(05)97023-7

Sidstedt, M., Rådström, P., & Hedman, J. (2020). PCR inhibition in qPCR, dPCR and MPS—mechanisms and solutions. *Analytical and Bioanalytical Chemistry*, *412*(9), 2009–2023. https://doi.org/10.1007/s00216-020-02490-2

Sipahioglu, H. M., Usta, M., & Ocak, M. (2006). Use of dried high-phenolic laden host leaves for virus and viroid preservation and detection by PCR methods. *Journal of Virological Methods*, *137*(1), 120–124. https://doi.org/10.1016/j.jviromet.2006.06.009

Smith, M. E., Henkel, T. W., & Rollins, J. A. (2015). How many fungi make sclerotia? *Fungal Ecology*, *13*, 211–220. https://doi.org/10.1016/j.funeco.2014.08.010

Toju, H., Tanabe, A. S., Yamamoto, S., & Sato, H. (2012). High-Coverage ITS Primers for the DNA-Based Identification of Ascomycetes and Basidiomycetes in Environmental Samples. *PLoS ONE*, *7*(7), e40863. https://doi.org/10.1371/journal.pone.0040863

Twengström, E., Sigvald, R., Svensson, C., & Yuen, J. (1998). Forecasting Sclerotinia stem rot in spring sown oilseed rape. *Crop Protection*, *17*(5), 405–411. https://doi.org/10.1016/s0261-2194(98)00035-0

Wang, H., Qi, J., Xiao, D., Wang, Z., & Tian, K. (2017). A re-evaluation of dilution for eliminating PCR inhibition in soil DNA samples. *Soil Biology and Biochemistry*, *106*, 109–118. https://doi.org/10.1016/j.soilbio.2016.12.011

Wang, Y., Duan, Y.-B., & Zhou, M.-G. (2015). Baseline sensitivity and efficacy of fluazinam in controlling Sclerotinia stem rot of rapeseed. *European Journal of Plant Pathology*, *144*(2), 337–343. https://doi.org/10.1007/s10658-015-0771-z

Wood, D. E., & Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biology*, *15*(3), R46. https://doi.org/10.1186/gb-2014-15-3-r46

Yin, Y., Ding, L., Liu, X., Yang, J., & Ma, Z. (2009). Detection of*Sclerotinia sclerotiorum in Planta*by a Real-time PCR Assay. *Journal of Phytopathology*, *157*(7-8), 465–469. https://doi.org/10.1111/j.1439-0434.2009.01543.x

Appendix

Appendix 1. Primers sequences used for qPCR and Oxford Nanopore Sequencing.

Primer	Sequence 5'-3'
ITS1CATTA	ACCWGCGGARGGATCATTA
ITS4ngsUni	CCTSCSCTTANTDATATGC
SSFWD	GCTGCTCTTCGGGGCCTTGTATGC
SSREV	CAGCTTGGTATTGAGTCCATGTCA

Appendix 2. Results of DNA concentration and purity measurements with Nanodrop Spectrophotometer (Thermo Fisher Scientific) and Qubit Fluorometer (Thermo Fisher Scientific) from samples used for qPCR and Nanopore Sequencing.

Table 1. DNA quantification and purity results of Grevbäck field samples for qPCR.

Sample	**DNA concentration ng/µl**	**Absorbance 260/230**	**Absorbance 260/280**
G1	0.436	0.68	2.35
G2	0.132	0.42	1.87
G3	0.300	0.57	2.19
G4	0.114	0.38	2.17
G5	0.304	0.36	2.03
GCombined	0.344	0.49	1.81
GPooled	0.322	0.44	1.89

Table 2. DNA quantification and purity results of Hovby field samples for qPCR.

Sample	**DNA concentration ng/µl**	**Absorbance 260/230**	**Absorbance 260/280**
G1	0.680	0.54	2.33
G2	1.16	0.46	1.99
G3	0.850	1.12	2.04

G4	0.466	0.57	1.87
G5	0.688	0.59	1.77
Combined	0.912	0.48	1.92
Pooled	0.880	0.70	1.96

Table 3. Concentration and Purity results of isolated digested plasmid after cleaning.

Sample	DNA concentration ng/µl	Absorbance 260/230	Absorbance 260/280
Digested plasmid	2.44	0.05	2.01

Table 4. DNA quantification and purity results of Dala field for Sequencing.

Sample	DNA concentration ng/µl	Absorbance 260/230	Absorbance 260/280
D1	0.264	0.55	2.60
D2	2.44	1.26	2.01
D3	0.232	0.84	1.96
D4	0.770	0.62	1.87
D5	0.556	0.50	1.82
DPooled	0.828	0.70	1.92

Table 5. DNA quantification and purity results of Grevbäck field for Sequencing.

Sample	DNA concentration ng/µl	Absorbance 260/230	Absorbance 260/280
G1	1.26	0.76	2.44
G2	0.818	0.52	2.0
G3	2.86	0.68	2.18
G4	1.37	0.63	1.87
G5	1.07	0.52	1.90
GPooled	1.62	0.65	1.99

Table 6. DNA quantification and purity results of Hovby field for Sequencing.

Sample	DNA concentration ng/µl	Absorbance 260/230	Absorbance 260/280
H1	0.306	0.62	2.26
H2	1.83	0.50	2.04
H3	1.21	0.54	1.84
H4	1.41	0.68	1.90
H5	1.15	0.61	1.99
HPooled	1.56	0.63	2.03

Table 7. DNA quantification and purity results of PCR product after cleaning, for Nanopore Sequencing.

Sample	DNA concentration ng/µl	Absorbance 260/230	Absorbance 260/280
Dala	46.4	2.44	1.86
Greväck	46.0	2.37	1.81
Hovby	38.8	2.13	1.84

Appendix 3. Results of qPCR assay of spiked unknown samples and standard curve.

Table 8. Spiked unknown samples.

Sample	Number of DNA copies	Mean Cq values
Dala individual 1	0.287	35.93
Dala individual 3	0.023	39.51
Dala individual 4	0.014	39.99
Dala individual 5	0.163	36.79
Dala Combined	0.055	38.24
Dala Pooled	0.024	39.38

Grevbäck individual 1	1.104	34.27
Grevbäck individual 3	0.043	38.79
Grevbäck individual 4	0.128	37.93
Grevbäck individual 5	0.127	37.17
Grevbäck Combined	0.064	38.36
Hovby individual 1	0.154	37.41
Hovby individual 2	0.302	38.68
Hovby individual 3	0.031	39.20
Hovby individual 4	0.128	36.92
Hovby individual 5	0.045	38.54
Hovby Combined	0.025	39.35
Hovby Pooled	0.103	37.37

Table 9. Standard curve used to quantify the spiked unknown samples.

Sample	**Number of DNA copies in 5 µl**	**Mean Cq values**
Dilution 1	2154993. 48	20.18
Dilution 2	215499. 34	23.08
Dilution 3	21549. 93	27.76
Dilution 4	2154.99	30.24
Dilution 5	215.49	33.65
Dilution 6	21.54	36.62
Dilution 7	2.15	38.19