Relative performance of Oxford Nanopore MinION vs. Pacific Biosciences Sequel third-generation sequencing platforms in identification of agricultural and forest pathogens

Kaire Loit[a#], Kalev Adamson[b#], Mohammad Bahram[c#], Rasmus Puusepp[d#], Sten Anslan[e], Riinu Kiiker[a], Rein Drenkhan[b], Leho Tedersoo[d,f*]

[a]Institute of Agricultural and Environmental Sciences, Estonian University of Life Sciences, Fr.R. Kreutzwaldi, 5, 51006 Tartu, Estonia

[b]Institute of Forestry and Rural Engineering, Estonian University of Life Sciences, Fr.R. Kreutzwaldi, 5, 51006 Tartu, Estonia

[c]Department of Ecology, Swedish University of Agricultural Sciences, Ulls väg 16, 75651 Uppsala, Sweden

[d]Institute of Ecology and Earth Sciences, University of Tartu, 14a Ravila, 50411 Tartu, Estonia

[e]Zoological Institute, Technische Universität Braunschweig, Mendelssohnstrasse 4, 38106 Braunschweig, Germany

[f]Natural History Museum, University of Tartu, 14a Ravila, 50411 Tartu, Estonia

#Equal contribution

*Correspondence: Leho Tedersoo, email leho.tedersoo@ut.ee, tel. +372 5665986

Running head: Third-generation sequencing-based pathogen diagnostics

ABSTRACT Culture-based molecular characterization methods have revolutionized detection of pathogens, yet these methods are either slow or imprecise. The second-generation sequencing tools have much improved precision and sensitivity of detection, but the analysis processes are costly and take several days. Of third-generation techniques, the portable Oxford Nanopore MinION device has received much attention because of its small size and possibility of rapid analysis at reasonable cost. Here, we compare the relative performance of two third-generation sequencing instruments, MinION and Pacific Biosciences Sequel in identification and diagnostics of pathogens from conifer needles and potato leaves and tubers. We demonstrate that Sequel is efficient in metabarcoding of complex samples, whereas MinION is not suited for this purpose due to the high error rate and multiple biases. However, we find that MinION can be utilized for rapid and accurate identification of dominant pathogenic organisms from plant tissues following both amplicon-based and metagenomics-based approaches. Using the PCR-free approach with shortened extraction and incubation times, we performed the entire MinION workflow from sample preparation through DNA extraction, sequencing, bioinformatics and interpretation in two and half hours. We advocate the use of MinION for rapid diagnostics of pathogens, but care needs to be taken to control or account for all potential technical biases.

IMPORTANCE We develop new and rapid protocols for MinION-based third-generation diagnostics of plant pathogens that greatly improves the speed and precision of diagnostics. Due to high error rate and technical biases in MinION, PacBio Sequel platform is more useful for amplicon-based metabarcoding from complex biological samples.

KEYWORDS PacBio Sequel, Oxford Nanopore MinION, molecular diagnostics, metabarcoding, metagenomics, plant pathogens, potato (*Solanum tuberosum*)

INTRODUCTION

Fungal pathogens and pests cause enormous losses in agriculture and forestry. Rapid and precise identification of pathogens enables efficient countermeasures and reduces the costs of biocides and losses to disease (Comtet et al., 2015). Direct morphology-based and culture-based diagnoses are often too imprecise or slow. Molecular methods such as specific primers, probes and sequence analysis are more accurate and can be rapidly applied to infected tissues (Kashyap et al., 2017). Although specific probes or primers combined with PCR/qPCR can be rapidly applied to tissue samples and environmental material, these methods lack the capacity to detect species or strains other than those intended, or worse, yield false positive signals (Grosdidier et al., 2017). Although high-quality sequencing reads are highly precise, Sanger sequencing of PCR products takes 1-3 days, depending on access to sequencing laboratory, and it may fail when DNA of several species or polymorphic alleles are amplified (Hyde et al., 2013).

Second- and third-generation high-throughput sequencing (HTS) platforms read hundreds of thousands to billions of DNA molecules, recovering the targeted taxa when present at very low proportions (Bik et al., 2012; Nilsson et al., 2019). However, library preparation and running of HTS instruments typically takes several days and there are queues of weeks to months in commercial service providers. Furthermore, a single run costs >500 EUR, which renders it unfeasible for rapid identification of pathogens (Tedersoo et al., 2019). In spite of millions of output reads, the second-generation SOLiD, Roche 454, Illumina and Ion Torrent platforms suffer from short sequence length that is suboptimal for accurate identification of microorganisms because of low taxonomic resolution of short marker gene fragments (100-500 bp; Mosher et al., 2014; Schloss et al., 2016). Third-generation sequencing platforms Pacific Biosciences (PacBio; RSII and Sequel instruments; www.pacb.com) and Oxford Nanopore (MinION, GridION and PromethION instruments; https://nanoporetech.com/) enable average sequence length of >20,000 bases, but this comes at 5%-20% error rate (Weirather et al., 2017; Jain et al., 2018; Tedersoo et al., 2018, 2019). In PacBio instruments, the built-in circular consensus sequencing generates multiple copies of the same fragment with highly accurate consensus (Eid et al., 2009; Rhoads and Au, 2015). Therefore,

long consensus molecules have been readily used in de novo assembly of complex genomes (Giordano et al., 2017) and DNA barcoding (Hebert et al., 2018). PacBio-based metabarcoding analyses provide greater resolution compared with second-generation HTS tools in bacteria (Singer et al., 2016; Wagner et al., 2016; Schloss et al., 2016) and fungi (Tedersoo et al., 2018), including plant pathogens (Walder et al., 2017).

Compared with other HTS platforms represented by large and quite expensive machines, the Oxford Nanopore MinION device is of the size of a cell phone that costs ca. 900 EUR, making it affordable to governmental institutions, research laboratories and small companies (Mikheyev and Tin, 2014; Lu et al., 2016). Its small size and low power consumption enable carrying the device, basic analysis toolkit, batteries and computer virtually anywhere, as demonstrated by in situ runs in a tropical rain forest (Quick et al., 2016), Antarctic desert (Johnson et al., 2017) and space station (Castro-Wallace et al., 2017). MinION has a capacity to produce >1 million sequences per day, with maximum read length approaching 1,000,000 bases (Jain et al., 2018). Because of low sequence quality, MinION has been mostly used in whole-genome sequencing analyses to resolve long repeats and bridge contigs or re-sequencing genomes (Quick et al., 2014; Jain et al., 2018). The error rate of reads can be reduced from 10-15% to 1-5% by sequencing of the complementary strand ($1D^2$ method) or preparing tandem repeat molecules (concatemers), but these solutions are laborious, of low sequencing depth, and hence not broadly used (Cornelis et al., 2019; Li et al., 2016; Calus et al., 2018; Volden et al., 2018). MinION has been used to generate long DNA barcodes from consensus sequences (Pomerantz et al., 2018) and to detect specific human pathogens that are easily distinguishable and well represented in reference sequence databases (Kilianski et al., 2015; Ashikawa et al., 2018). Although multiple reports claim achieving species-level taxonomic resolution (Benitez-Paez et al., 2016; Benitez-Paez and Sans, 2017; Kerkhof et al., 2017), the high error rate renders nanopore sequencing poorly suited for exploratory metabarcoding analyses of real communities. The metagenomic approach has gained popularity for identification of human pathogens to skip the PCR step and avoid associated biases (Quick et al., 2016; Schmidt et al., 2017; Votintseva et al., 2017). Recently, Bronzato Badial et al. (2018) demonstrated that plant pathogenic bacteria and viruses can be detected using MinION, whereas Hu et al. (2019) extended this to fungal pathogens of cereals in a preprint.

The main objective of this study is to develop protocols for metabarcoding-based and metagenomics-based detection of fungal plant pathogens using third-generation sequencing tools. In particular, we aim to 1) test the relative biases and shortfalls of MinION-based and Sequel-based identification of pathogens and evaluate the perspectives of these methods in pathology and ecology; and 2) optimize MinION protocols for ultra-rapid pathogen identification. We performed several HTS runs using MinION and Sequel instruments and validated the results by comparing these to Sanger sequencing, species-specific priming PCR and morphology-based assessment where relevant. We tested the third-generation HTS methods in two plant pathosystems, conifer needles and potato (*Solanum tuberosum*) leaves and tubers.

RESULTS

Technical features of MinION and Sequel runs Compared with Sequel, MinION had several-fold greater initial sequencing depth, which depended on the loaded DNA content and sequencing time (Table 1). The high sequence number of MinION was reduced several-fold during the quality filtering and demultiplexing, reaching the level comparable with Sequel. Among samples, variation in sequencing depth was slightly greater in MinION (CV, 67.8%-93.4%) compared with Sequel (62.8%-64.5%). Pearson correlation coefficient of sequencing depth of samples in MinION and Sequel ranged from 0.585 in the potato data sets to 0.853 in the needle data sets, suggesting a substantial library preparation or sequencing bias in the potato amplicon pool but not in the needle sample pool.

For the MinION data sets, chimeras were detected using the reference-based method but not *de novo* method. Putatively chimeric molecules contributed 1.5%-1.8% to the mapped reads, but nearly half of these were false positives based on manual checking (cf. Hyde et al., 2013). Interestingly, nearly half of the true chimeras included parents from different samples, indicating some chimera formation during the library preparation or sequencing process in addition to PCR. Further manual inspection of demultiplexed sequences revealed that 5-8% of these are self-chimeric, i.e. 1.5-fold to 6-fold repeats of itself. In the Sequel data sets, chimeras accounted for 1.9-3.7% of reads (including 1.5-2.4% detected *de novo*), with no self-chimeric reads remaining.

Index switching rate was much greater in MinION (3.6% of reads in ONT2) than Sequel (0.14%, ONT2S). Based on positive control samples, we estimated that the error rate in Sequel is around 0.1% (corresponding to polymerase errors), but around 11-16% (depending on species) for the 1D method and 11% for the $1D^2$ method of MinION. Based on alignments of hundreds of positive control and other dominant sequence types, we noticed that errors were non-randomly distributed, i.e. occasionally there were no errors across 4-5 bases of the alignment, whereas homopolymeric sites were infested with large amounts of combined indels and substitutions (Fig. 1). Because of these non-random errors, we were able to construct consensus at 98.5-99.5% accuracy (only deletions remaining) with 100 or more reads.

All MinION runs on R9.4 flowcells from one batch (flowcells #1 and #2) were contaminated by *Coniothyrium* sp. (INSD accession JX320132), but this taxon was not observed in negative control samples, another batches (flowcell #3 and runs not reported here) or runs using R9.5 flow cells, or PacBio Sequel. At least partly because of this, the dominant fungal taxa recovered in samples differed in the MinION and Sequel runs (Tables 2, 3).

Metabarcoding analyses of MinION and Sequel The MinION ONT1 run included diseased and asymptomatic needle samples and pure cultures of pathogens. Of 792,748 passed reads, 189,150 (23.9%) were demultiplexed and 183,343 (23.1%) were mapped to reference sequence databases based on the quality criteria (e-value <e-40 and sequence similarity >75%). The ITS1catta forward primer amplified mostly Fungi (99.9% of identified reads). Best hits were distributed across 2483 fungal OTUs, with the well-known conifer pathogens yielding hits to 1-2 different accessions. On average, needle samples hosted 203.4±130.5 (mean±SD) OTUs. Best hits to the contaminant *Coniothyrium* sp. contributed 26.3% of all sequences on average. Of expected taxa, *Hormonema macrosporum* (6.2%), *Lophodermium conigenum* (5.0%) and *Didymella lentis* (4.3%) yielded the greatest number of hits (Fig. 2a) and all these taxa occurred in 94%-100% of needle samples.

The ONT1S Sequel run revealed 121,965 demultiplexed reads that were clustered into 535 OTUs, all above the quality threshold. Needle samples harboured on average 51.5±41.6 OTUs, nearly four times less than in the MinION data set. Altogether 99.9% reads were ascribed to Fungi, with *Lophodermium pinastri* (17.8%),

Dothiostroma pini (8.9%) and *Sydowia* sp. (7.0%) dominating across the entire data set (Fig. 2b). These dominant taxa occurred in 43-65% of samples.

The ONT2 MinION run recovered 255,137 passed sequences, of which 16.2% were demultiplexed and 14.4% mapped to reference database reads. Based on the distribution of pine-specific pathogens in potato samples, we estimated that 13.4% of the reads were carried over from the previous ONT1 run, in which we used the same primer and tag combinations. In the ONT2 run, these pine-specific species had a proportionally similar relative abundance when comparing across the same index combinations. The ITS1catta forward primer amplified mostly Fungi (74.2% of identified reads; Fig. 2c). Reads corresponding to potato (9 OTUs) and *Coniothyrium* sp. accounted for 26.1% and 13.2% of sequences. Of putative potato pathogens and endophytes, the ITS1catta forward primer revealed *Boeremia lycopersici* (7.5% of reads), *Mycosphaerella tassiana* (4.6%) and *Peyronellaea* sp. (4.4%) as dominants. The average richness was 81.7±43.3 OTUs per sample. The oomycete ITS1Oo primer comprised only 1.6% of all reads that were dominated by Oomycota (47.7%), other Stramenopila (19.2%), Fungi (23.6%) and Viridiplantae (9.5%). In each sample, 0-3 oomycete taxa were found and all of these occurred only once or twice (Table 4). The majority of samples produced no amplicon with the ITS1Oo primer and these samples contained no Oomycetes based on the HTS analysis.

The ONT2S Sequel run revealed 75,573 demultiplexed reads that were all matched to reference sequences and separated into 308 OTUs. On average, 39.6±20.3 OTUs were recovered per sample. In the ITS1catta amplicons, Fungi, Viridiplantae, Alveolata and Rhizaria contributed to 51.0%, 48.4%, 0.5% and 0.1% of reads, respectively. All plant reads were distributed across 25 OTUs that were all assigned to potato. Six of the OTUs probably represent naturally high variation among ITS sequences of potato (based on INSD entries), whereas others represent pseudogenes or non-functional copies. These were rare to common (up to 3% of all variants) and sometimes exceeded the abundance of regular variants in individual samples. Of Fungi, the largest number of reads belonged to *Boeremia* sp. (8.0%), Hysteriaceae sp. (7.4%) and *Cladosporium herbarum* (3.3%; Fig 2.d). The ITS1Oo primer accounted for 1.4% of sequences that were mostly assigned to Oomycota (62.9%), other Stramenopila (33.9%), Viridiplantae (3.0%) and Alveolata (0.2%). This data subset yielded 0-2 OTUs of Oomycota or other Stramenopila per sample (Table 4).

Sequel and MinION recover the same dominant fungal species (excluding the contaminant) in 60% and 63% of cases in the needle and potato samples, respectively. These values increased to 78% and 83%, respectively, when considering overlap in the three best matching taxa. Inspection of the discordant samples revealed that contamination from the previous run blurred the results of the potato samples and MinION produced one to two orders of magnitude less high-quality reads matching to multiple species such as *Lophodermium pinastri*, *Vishniacozyma victoriae*, *Cystobasidium* sp. and *Dendryphion* sp. as compared with Sequel. These species had a relatively high number of homopolymers (>3-mers) per ITS sequence length compared with dominant but equally shared taxa ($F_{1,8}$=5.79; P=0.088). The Stramenopile data subsets were in a stronger agreement in Sequel and MinION apart from the lack of *Peronospora variabilis* amongst MinION reads and hence its unsuccessful diagnosis from three potato leaf samples.

The ONT2a and ONT2b MinION runs were designed to test whether long indexes relieve the massive index switching. The ONT2a run revealed a tag switch rate of 3.8%, whereas the ONT2b run failed for unknown reasons. The potato (14 OTUs) contributed to 18.7% of reads, whereas the contaminant *Coniothyrium* species accounted for 16.7% of reads, prevailing in half of the eight potato samples. Of other fungal species, Tremellales sp. (11.5%), *Filobasidium stepposum* (7.0%) and *Mycosphaerella tassiana* (5.6%) dominated. These species were less common in these eight samples in the ONT2 run, (3.8%, 2.7% and 4.6%, respectively). Nonetheless, the same best fungal hits prevailed in 75% of the samples in the ONT2a and ONT2 runs.

The ONT2f run was intended to test suitability of the $1D^2$ method. This run recovered only 3241 $1D^2$ reads. Only 29.7% of reads fell within 10% of the expected read length of ca. 3200 bases and the median read length was 954 bases. As the positive control revealed no reads, the tag switch rate could not be calculated. Of all sequences, potato (17 OTUs) accounted for 54.19% of sequences. Of Fungi (39.8%), *Taphrina populina* (6.0%), *Parastagonospora* sp. (3.8%) and *Glarea lozoyensis* (3.0%) dominated. These species were somewhat less common in the ONT2 library (0.1%, <0.1% and <0.1%, respectively). The same species were among the dominants in only 25% of samples as based on the ONT2 and ONT2f runs. It remains unknown whether these biases are related to sequencing of long amplicons or the $1D^2$ method.

Metabarcoding vs. metagenomics approach The ONT2g run representing a metagenome of a single diseased potato tuber sample (KL036) yielded 66,133 and 400,355 'passed' and 'failed' sequences, respectively. The 5000 randomly selected sequences from each bin included 1325 'passed' reads and 1 'failed' read that met our quality standards. Altogether 37.4% of the 'passed' reads represented ITS sequences carried over from a previous run. After removal of these reads, the metagenomics data set was dominated by plant and bacterial reads. Best hits to *Lycopersicon esculentum* (tomato, 29.0% of reads) and seven species of *Solanum* (altogether 22.6%) collectively represented the potato. Of Bacteria, hits to *Agrobacterium tumefaciens* (10.5%), *Variovorax paradoxus* (9.7%) and *Sphingopyxis alaskensis* (3.8%) dominated. Fungal hits were less common; these to *Rhizoctonia solani* (1.6%) and *Boeremia exigua* (1.0%) prevailed. Of these taxa, *A. tumefaciens* and *V. paradoxus* are probably present given their best matches of 93% and 92% and average matches of 87% and 85% similarity, respectively, to database sequences. Conversely, *S. alaskensis, B. exigua* and *R. solani* are probably absent, because of their best hits reached 84%, 88% and 86%, and all hits averaged 79%, 80% and 80% similarity to reference sequences, respectively.

The ONT2h run represented a long amplicon of the same sample, recovering 342,923 'passed' reads and 423,688 'failed' reads. Of the randomly selected 5000 sequences, 1876 'passed' reads and 1068 'failed' reads met the quality threshold. The positive control used in the next to previous run accounted for 0.2% of all sequences, mostly in the 'failed' bin. Out of 18 most commonly hit species, the proportion of 11 differed significantly ($P<0.001$) among the 'passed' and 'failed' bins, indicating that reads of certain taxa are much more likely to be recorded as failed. Of the 'passed' sequences, matches to *Lignincola laevis* (Pleosporales, 64.3%), *Verticillium biguttatum* (Hypocreales, 5.0%) and *Thanatephorus cucumeris* (Cantharellales, 3.0%) dominated. In the fail bin, *Verticillium biguttatum* (19.9%), *L. laevis* (15.7%) and *Plectosphaerella cucumerina* (Pleosporales, 7.8%) prevailed, followed by *T. cucumeris* (6.0%). Of the dominant taxa recovered, probably only *V. biguttatum, T. cucumeris* and *P. cucumerina* are identified to the species level given their high maximum (>90%) and mean (>85%) blast similarity. Taxa relatively more abundant in the 'failed' bin tended to possess more and longer homopolymers than those in the 'passed' bin.

In the ONT2g metabarcoding and ONT2h metagenomics data sets derived from the same sample, none of the fungal taxa were shared. Although *R. solani* is regarded as a synonym of *T. cucumeris* (or vice versa), the isolate named as *R. solani* with available genome is probably heterospecific with the *T. cucumeris* isolate that was best matched in the amplicon data set. The *R. solani-T. cucumeris* complex has high variability in the rRNA marker genes and its taxonomy is far from settled (Veldre et al., 2013). Other fungal species common in the metabarcoding data set were absent from the metagenomics data set probably because their genome is unavailable. Several of these ascomycetes may have best matched to *Boeremia exigua* that has a genome sequence available. *B. exigua* was represented by a single read, potentially resulting from carry-over from the first run. This situation highlights limitations of the metagenomics approach when insufficient reference is available.

Express identification The ONT2i run intended to minimize time from sampling to diagnosis based on a single infected potato tuber sample and metagenomics approach. Using forceps, we mounted ca 20 mg of infected tissue into 2 ml Eppendorf tube containing 100 µl lysis buffer from the Phire Plant Direct PCR Kit. Based on previous optimisation for speed, we reduced the step of lysis to 15 min that included tissue disruption using bead beating (5 min at 30 Hz), brief centrifugation at 5000 g, incubation at 30 °C for 5 min and final centrifugation at 11,000 g for 1 min (the recommended protocol includes lysis without tissue disruption at room temperature >2 h). The DNA was concentrated from lysate using the FavorPrep kit following the manufacturer's instructions except centrifugation steps for 1 min and final elution using 50 µl water (altogether 25 min). Qubit measurement revealed DNA concentration of 4.1 ng μl^{-1} (5 min). Library preparation followed the G004 protocol (50 min). The MinION run was interrupted at 1200 reads (50 min) and the 436 'passed' fastq reads were analysed in PipeCraft that generated a list of 10 best hitting taxa in <5 min and revealed *T. cucumeris* as a prevalent pathogen. The parallel WIMP analysis failed because of server maintenance at the time of analysis. The entire procedure took 2 hours and 30 minutes. Notably, the sequencing process was suboptimal because of the low amount of DNA used, which resulted in <20% pores effectively used at termination of this run. Some extra time was required to finalize the results for a written report. The metagenomics reads were dominated by hits to tomato (72.7%), followed by various bacteria (6.4%), and *T. cucumeris* (5.5%), and *P. cucumerina* (3.0%) that are both known pathogens of

potato. Subsequent Sanger sequencing from tuber samples with black scurf symptoms of the same diseased potato revealed *T. cucumeris* (all four subsamples) and *Pyronemataceae* sp. (50% of subsamples).

DISCUSSION

Use of third-generation sequencing instruments for DNA metabarcoding Using the same amplicon pools and additional morphology-based, Sanger sequencing-based diagnosis or species-specific priming PCR, we had a unique opportunity to evaluate the relative performance and biases of MinION compared with Sequel. Nanopore sequencing revealed somewhat greater among-sample variability in sequencing depth, which may be related to library preparation, sequencing, data processing or a combination of these.

MinION suffered from a unique issue with sequence carry-over from a previous run as also noted by Cusco et al. (2018) for 6% of reads. In our study, a washing procedure with the supplier's Wash Kite still yielded 13-37% sequence carry-over from a previous run. Furthermore, we could recover traces of a positive control used in the next to previous run at 0.2% relative abundance. It is theoretically possible that such carry-over contamination occurs on re-usable flowcells or chips of other HTS platforms as we have commonly seen it in the end of untrimmed Sanger reads.

In our analyses, MinION had an issue of contamination with a fungus matching *Coniothyrium* sp. that was not observed in Sequel run and in none of our previous data sets. INSD records indicate that this taxon is common in temperate USA. This contamination occurred in two R9.4 flowcells (#1 and #2) supplied with the MinION instrument, but not in another R9.4 batch (flowcell #3 and others not reported here) or R9.5 batch. The flowcells #1 and #2 were used over 6 months, considering several independent laboratory contamination events unlikely. Therefore, we suspect that this contamination may be related to the supplier.

Chimeric reads were common in both Sequel and MinION data. UCHIME effectively detected chimeric molecules from the Sequel data, but it performed poorly on MinION data. The error-infested reads were probably too different from each other to be recognized as chimeric. MinION data also included a substantial proportion of chimeric molecules with parents from different samples, representing a unique hybrid issue of index switching and chimera formation. A large proportion of long MinION reads represented self-chimeras that were not

277 recognized by the chimera filtering software. This issue was common on PacBio RSII instrument (Tedersoo et al.,
278 2018), but it was not observed in the current Sequel runs. Since MinION reads are typically mapped to reference,
279 we estimate that the abundant chimeric molecules create virtually no bias, except those with switched tags.

280 Index switches during library preparation or sequencing make a strong and perhaps predominant
281 contribution to sample 'contamination' (Schnell et al., 2015). The observed index switching rate of 3.6-3.8% in
282 MinION compares poorly with that of Sequel (<0.2% in; this study) and various Illumina instruments (0.1-10%;
283 Costello et al. 2018). The double indexes performed equally poorly, suggesting that index switches are
284 attributable to processes in library preparation or sequencing rather than sequencing errors. At least partly, high
285 rates of index switching spilled the dominant taxa in the deeply sequenced MinION data sets across nearly all
286 samples and resulted in 2-fold to 4-fold greater richness per sample. Certainly, the high error rate and inaccurate
287 mapping-based method of OTU construction contribute to this difference. The MinION-derived error-infested
288 metagenomics and amplicon sequences may be easily mapped to various closely related species, thus elevating
289 richness artificially. Conversely, clustering at 98% sequence similarity may be too conservative, because many
290 pathogenic taxa differ from each other by only a few bases in the ITS region (e.g. needle pathogens *Dothistroma*
291 *pini* and *D. septosporum*, see Barnes et al., 2016), and therefore several species with distinct ecology and
292 pathology may be lumped into a single taxon (Kõljalg et al., 2013). In spite of substantial disparity in the taxon-
293 rich MinION and Sequel fungal data subsets, these two instruments revealed high-level overlap in the species-
294 poor oomycete data subset.

295 The average error rate of MinION reads was 10-15%, depending on the proportion of homopolymers in
296 the marker gene region of particular species. Species possessing homopolymer-rich ITS markers were up to two
297 orders of magnitude less common than in the Sequel run. This was also supported by relative prevalence of
298 homopolymer-rich taxa in the 'failed' bin. We showed that this may substantially bias estimates of dominant
299 fungal and perhaps oomycete taxa in specific samples and overall.

300 We observed discrimination against longer reads when sequencing potato amplicons using the $1D^2$
301 approach (ONT2f), which confirms a previous report (Cusco et al., 2018). Preferential recovery of shorter reads

seems to be inherent to both PCR and all sequencing instruments including Sequel (Tedersoo et al., 2018, 2019; Nilsson et al., 2019).

Using the same amplicon pools, Sequel and MinION revealed contrasting results in metabarcoding of conifer needle and potato samples. The results of Sequel were generally consistent with morphology-based and species-specific priming PCR diagnosis (needle samples) and Sanger sequencing results (potato samples), but failed to differentiate the closely related needle pathogens *D. pini* and *D. septosporum.* Apart from the species of *Coniothyrium* contaminating the MinION data sets, the two platforms revealed different taxa (by names and INSD accessions) prevailing in the same samples. In some cases, these taxa are considered synonyms (e.g. *M. tassiana* and *C. herbarum*), but mostly these belong to closely related sister taxa that share the UNITE Species Hypothesis at 2% level as confirmed by manual comparisons of best-matching reads. In nearly 20% of occasions, inconsistencies between MinION and Sequel were attributable to the poor recovery of taxa with abundant homopolymers in ITS sequences by MinION. Possibilities to solve this include lowering of the initial phred score or including the 'failed' reads in analyses.

Rapid molecular diagnostics We tested both metagenomics- and amplicon-based approaches of nanopore sequencing for rapid identification of pathogens. Most taxa recovered in the metagenome run were rare in the amplicon data set and vice versa. Although we detected severe biases in MinION amplicon sequencing, we believe that amplicon-based analyses are more accurate and that reference bias accounts for much of the difference; i.e., in metagenomics analyses, species with available reference genomes have much greater chance of accumulating hits compared with species with no available genomes. In our analyses, this is illustrated by misidentification of potato as a tomato. Mapping reads of an ascomycete pathogen to genomes of several others, as in our study, is likely to remain cryptic. A solution may be sufficiently deep metagenomics sequencing to secure coverage of mitochondrial or ITS marker genes that are in multiple copies per genome. Because genomes of most bacterial and fungal human pathogens have available genome sequences, nanopore metagenomics-based identification may be better suited to medical samples, but this situation is likely to improve very soon in plant pathology.

We demonstrate that accurate molecular identification from sample collection through DNA extraction, concentration, library preparation, sequencing, bioinformatics and taxonomic interpretation may take as little as two and half hours using MinION sequencing in the metagenome mode. This is less than the four hours reported for identification of bacterial human pathogens from urine, which requires no specific DNA extraction (Schmidt et al., 2017), and one working day as commonly reported in multiple studies (e.g. Quick et al., 2016; Votintseva et al., 2017; Charalampous et al., 2018). The protocols for plant and animal tissues can be potentially optimised to 90 minutes in cases where DNA is easily extractable in high quantity and purity; library preparation and sequencing process can be limited to 10-20 minutes each, when using rapid library kits and halting efficient sequencing runs after a critical number of reads (Votintseva et al., 2017). These express diagnostics rates of MinION cannot be beaten by instruments of other HTS platforms that require several hours for library preparation and at least one day for sequencing (except 6 hours for Ion Torrent; Reuter et al., 2015).

However, there is a clear trade-off between overall analysis time and data reliability in nanopore sequencing. Shortened DNA extraction protocols tend to yield lower quality and quantity of DNA, whereas culled incubation and centrifugation steps in library preparation may result in dilute and poorly indexed libraries overrepresented by short fragments. Although we successfully identified potato pathogens from a library 13-fold more dilute than recommended, it may be useful to add a certain amount of so-called carrier DNA to secure preparation of high-quality libraries (cf. Mojarro et al., 2018). Sequencing time is almost linearly related to sequencing depth and thus quality of consensus reads and genome or taxonomic coverage. Sample preparation, bioinformatics and interpretation processes take longer for multiplexed samples, which may be necessary to reduce the overall analytical costs of 500-800 EUR per sample to ca 100 EUR per sample using Oxford Nanopore's commercial multiplexing kits or to 2-3 EUR per sample by using custom multiplexing methods of indexed primers. For example, analytical costs for the ONT1 and ONT2 runs were roughly 10 EUR per sample considering triple use of flowcells. To reduce the chances of carryover of previous molecules, contamination-aware indexes (different indexes across runs) could be used.

Technical and analytical issues Although several authors report species-level identification of bacterial taxa in complex communities using MinION (e.g. Shin et al., 2016; Kerkhof et al., 2017), these interpretations are not convincing, because mapping of sequences with 10-25% errors to reference reads of high similarity is inaccurate (see above). We demonstrate that even when using the relatively rapidly evolving fungal ITS region and reference sequences differing from each other by at least 2%, positive control samples and plant material yield multiple OTUs, sometimes recovering strongest hits to different genera. Conversely, species absent from databases are mapped to one or more closest species, which may provide wrong taxonomic implications. This is of particular importance for molecular diagnostics, necessitating inclusion of marker genes of all potential target species in the reference database to prevent incorrect interpretation. This issue is particularly pressing for long rRNA amplicons and single-copy genes. The metagenomics approach requires a comprehensive database of genomes of all potential target organisms, which strongly depends on whole genome sequencing initiatives. Exome compilations are suboptimal, because much of the eukaryotic DNA is non-coding. Besides target organisms, metagenomics databases also require inclusion of potential contaminants such as specific interacting taxa (e.g. potato) and human, and various bacteria that contribute much to the metagenomic DNA.

A major concern with novel sequencing techniques is the paucity of reports on analytical shortfalls, and nature of artefacts, which may partly be derived from the lack of controls or inappropriate sampling design (Pontefract et al., 2018). The MinION has been used for five years, but so far there is very little information about analytical errors and biases, and very few authors mention about checking chimeras, index switching artefacts or unsuccessful runs (but see White et al., 2017 Cusco et al., 2018; Xu et al., 2018). The virtual lack of constructive (self)criticism echoes a misleading message about the multi-purpose, non-problematic use of the method, serving the interests of the manufacturer and journal editors in an unjustified manner. Users of MinION, many of which have no prior experience with other HTS techniques and related problems, heavily struggle to squeeze reasonable data out of the device. There are thousands of academic users, but only a few hundred papers out. For example, our team purchased MinION instruments with extra troubleshooting service; yet, the company responded only to technical questions but not to troubleshooting about index switching, contamination and sequence carry-over. It should be the responsibility of researchers and editors to include such problematic issues and solutions in

publications to prevent the research community and specialists of governmental and private institutions from wasting countless time (re)falling into the same analytical holes.

Perspectives of third-generation sequencing technologies Both MinION and Sequel are rapidly evolving in terms of read length and base calling accuracy. At the moment, Sequel seems to be the best choice for metabarcoding regular-size (600-1000 bp) and long (up to 3 kb) amplicons (Cline et al., 2015; Kyaschenko et al., 2017; Tedersoo et al., 2018; Nilsson et al., 2019; Tedersoo et al., 2019) and for barcoding ultra-long markers (up to 7 kb; Wurzbacher et al., 2019). The forthcoming M8 chemistry will reduce the costs of PacBio sequencing roughly two-fold. Declining costs, greater throughput, read length and quality continue to make Sequel more attractive for metabarcoding and seek supporters from metagenome and metatranscriptome researchers (Rard, 2018). However, it may be hard to convince users of Illumina MiSeq to switch to another platform and adopt alternative bioinformatics workflows.

Use of MinION for metabarcoding looks relatively less promising considering the current state-of-the-art technology with unacceptably high error rates. The error rates should be reduced to <0.1%, i.e. to the level of Illumina and circular consensus of Sequel for use in routine metabarcoding. Several methods of tandem repeat (concatemer) sequencing enable to reduce error rates to 1-3% (Li et al., 2016; Calus et al., 2018; Volden et al., 2018). Double-barcoding of each size-selected RNA or DNA molecule followed by generation of consensus sequences yields quality improvements comparable to tandem repeat sequencing (Karst et al., 2018), but it would require ultra-high sequencing depth to reach 1% error rate and to be able to multiplex over several biological samples. Combining these methods may facilitate reducing error rates towards the critical threshold.

For regular barcoding, the third-generation sequencing tools offer great promise in situations where their throughput and read length are much superior compared to double-pass Sanger sequencing, i.e. for barcodes >1000 bases and multiplexing hundreds of samples to secure cost-efficiency (Hebert et al., 2018; Srivathsan et al., 2018). Sanger sequencing handles poorly the alleles or copies of markers with read length polymorphism, which is common in non-coding regions of eukaryotes. The third-generation HTS technologies are able to recover various alleles as well as pseudogenes (Cornelis et al., 2019) from mixed or contaminated samples by sequencing

single DNA molecules. Sequel and MinION are capable of handling DNA amplicons of >7000 bases, requiring generation of consensus reads for reliable results (Wurzbacher et al., 2019). Although we could reach 99.5% accuracy with over 100 reads, Pomerantz et al. (2018) estimated that 100 MinION reads suffice for principally error-free generation of barcodes for animals using sequences from a complementary strand. For PacBio, a single read may be enough for reads around 2000 bp, but three or more may be needed for longer fragments and to average over PCR errors.

Unlike Sequel and other HTS technologies, MinION is well-suited to rapid diagnostics of pathogens and invasive species especially in groups that are well-known and well-referenced in public sequence databases. Besides detection of pathogenic species, MinION has a potential to recover antibiotic resistance genes and pathogenesis-related genes as well as single mutations in the metagenomics mode (Bradley et al., 2015; Cornelis et al., 2019). By using multiplex amplicons or metagenomics/genomics approach, it will be possible to detect harmful organisms and their specific pathogenicity-related genomic features in less than one day (Schmidt et al., 2017). Besides enabling to trace taxon/gene exchange between different habitats (Bahram et al., 2018), this approach has important implications for improving diagnosis and implementing countermeasures, e.g. releasing biocontrol agents, spraying biocides or arranging quarantine.

Nanopore-based detection methods are flexible for incorporating additional options such as recording epigenetic modifications (Jain et al., 2016; Simpson et al., 2017) and primary structure of RNA (Garalde et al., 2018), proteins (Perez-Restrepo et al., 2018) and polysaccharides (Karawdeniya et al., 2018). Alternative nanopore-based DNA sequencing methods are also evolving (Goto et al., 2018). The potential of different nanopores to record various biomolecules indicates great promise of nanopore-based diagnostics in the future.

Conclusion We demonstrate that the MinION device is well-suited for rapid and accurate diagnosis of pathogens, which may take as little as 150 minutes from sample preparation (including DNA extraction, library preparation, sequencing, bioinformatics and data interpretation). However, care should be taken to secure profound reference sequence data to avoid misdiagnosis. Amplicon-based diagnostics takes longer time, but is more accurate if genomes of potential pathogens are unavailable. For whole-community metabarcoding, Sequel performs much

better than MinION in terms of data quality and analytical biases. Although development of tandem repeat sequencing and read consensus sequencing have been developed for MinION, their error rate of 1-3% is still insufficient for exploratory metabarcoding analyses of biodiversity, but this may change in the coming years.

MATERIALS AND METHODS

Sample preparation The potato subset includes 27 samples of leaves and 10 samples of tubers with symptoms of disease (Table 5). We also included a DNA sample of two Australian Tuberaceae species as a positive control. The conifer system includes 36 distinct needle samples from eight species of *Pinus* and two species of *Picea* that represent material with symptoms of needle blight or no symptoms. The samples of natural, planted or recently imported trees were collected in Estonia in 2011-2018 (Table 6). We included a cultured isolate of *D. pini* (146889), *D. septosporum* (150931) and closely related *Lecanosticta acicola* (150943) as positive controls. Unequal mixture of DNA from these cultures served as a simple mock community. In both systems, we included a negative control sample.

In needle samples, DNA from 0.2 g plant material and fungal cultures was extracted using the Thermo Scientific GeneJET Genomic DNA Purification Kit (Thermo Fisher Scientific, EU). In potato samples, DNA from 0.1 g diseased fresh leaf tissue was extracted with a lysis buffer (0.8 M Tris-HCl, 0.2 M $(NH_4)_2SO_4$, 0.2% w/v Tween-20; Solis BioDyne, Tartu, Estonia). For some additional analyses using potato samples, we also used Phire Plant Direct PCR Kit (Thermo Scentific, Waltham, MA, USA).

Molecular identification Needle samples were screened for specific pathogenic fungi *Dothiostroma pini, D. septosporum* and *Lecanosticta acicola,* using species-specific primer pairs following the developers' protocols (Ioos et al., 2010). Potato samples were amplified using the ITS1F + ITS4 primer pair (White et al., 1990; Gardes & Bruns, 1993) and sequenced using the ITS5 primer (White et al., 1990).

For the metabarcoding approach, we used a forward primer ITS1catta (5'-ACCWGCGGARGGATCATTA-3'; Tedersoo et al. submt.) and a reverse primer ITS4ngsUni (Tedersoo &

Lindahl, 2016) to be able to selectively amplify fungal DNA and simultaneously avoid the 18S rRNA gene intron bias. Located in the terminus of the 18S rRNA gene, the ITS1catta primer covers nearly all Ascomycota and Basidiomycota as well as selected groups of 'zygomycetes' and 'early diverging lineages', but discriminates against plants and most other eukaryote groups (incl. fungal taxa Mortierellomycota and Tulasnellaceae) in the last position. To specifically detect Oomycota, we used the ITS1Oo primer (Riit et al., 2016, 2018) in combination with the ITS4ngsUni primer for the potato data set. Forward primers were tagged with one of the 12-base Golay indexes with at least four differences to any other index (Tedersoo et al., 2018). Because of issues in data recovery, we also amplified a subset of eight potato samples (KL001-KL008) using ITS1catta and ITS4ngsUni primers in which the forward primer was equipped with tandem repeat barcode of double length (securing at least 8 base difference) to increase resolution among samples. Because the $1D^2$ nanopore sequencing method requires DNA fragments of >2kb, we amplified these potato samples (>3 kb amplicons) using the indexed ITS1catta primer combined with the LR14 primer (Vilgalys & Hester, 1990). For comparing the metabarcoding approach to metagenomics method, we used ITS1catta in combination with the LR11 primer (Vilgalys & Hester, 1990) that yielded stronger amplicons compared with LR14. We used negative and positive controls as above.

PCR was performed in 25 µl volume comprising 0.5 µl each of the tagged primer (20 µM), 5 µl HOT FIREPol Blend Master Mix (Solis Biodyne), 1 µl DNA extract and 18 µl ddH_20. Thermocycling conditions included an initial 15 min denaturation at 95 °C, 30 cycles of 30 sec denaturation, 30 sec annealing at 55 °C and 60 sec elongation at 72 °C, and a final 10 min elongation before hold at 4 °C. The number of cycles was increased to 35 or 38 for some samples to yield a visible amplicon on 1% agarose gel. For the ITS1Oo + ITS4ngsUni primer combination, 40 PCR cycles at 50 °C annealing was used to secure greater product recovery. The two replicate amplicons were pooled, checked on a gel, and mixed with amplicons of other samples in roughly equal proportions based on visual estimates of band size.

Third-generation sequencing The mixed amplicons of potato and those of needles were separately split into library preparation for Sequel and MinION. The two PacBio libraries were sequenced on a Sequel instrument using the same SMRT cell (SMRT cell 1M, v2; Sequel Polymerase v2.1, sequencing chemistry v2.1., loading by

diffusion, movie time 600 min, pre-extension time 30 min). PacBio CCS reads (minPasses=3, MinAccuracy=0.9) were generated using SMRT Link v 6.0.0.47841 (Pacific Biosciences). Subsequent bioinformatics were performed using PipeCraft 1.0 (Anslan et al., 2017) that included steps of demultiplexing (2 mismatches to primer and 1 mismatch to tag), extraction of the ITS region (ITSx: default options; Bengtsson-Palme et al., 2013), chimera removal (UCHIME: de novo and reference-based methods combined; Edgar et al., 2011; and additional filtering option „remove primer artefacts“ that removes chimeric sequences were the full length primer is found somewhere in the middle of the read), clustering (UPARSE: 98% sequence similarity threshold; Edgar, 2013), taxonomic identification (BLASTn: e-value=0.001, word size=7, reward=1, penalty=-1, gap open=1, gap extend=2; Camacho et al., 2009) against UNITE 7.2 (Kõljalg et al., 2013) reference database. We used the criteria of blast e-value <e-40 and sequence similarity >75% for the kingdom level identification, and e-value <e-50 for phylum and class-level identification.

For the MinION instrument, amplicon library preparation was performed using the 1D amplicon/cDNA by Ligation (SQK-LSK109) kit (Oxford Nanopore) using R9.4 flowcells following manufacturer's instructions. For long fragments, we also used the $1D^2$ sequencing of genomic DNA (SQK-LSK308) kit on R9.5 flowcell, following the producer's protocols. Flowcells were used 1-3 times, being cleaned once or twice using the Wash Kit (EXP-WSH002; Oxford Nanopore). Sequencing was performed in the laboratory at room temperature, connecting the MinION device to a plugged-in, internet-connected laptop computer with four processors and 32 GB RAM. For base calling in MinKnow 3.1.19 software (Oxford Nanopore), we used the default phred score = 11, which placed the reads into 'passed' and 'failed' bins. The 'passed' fasta-formatted reads (additionally 'failed' reads in some analyses) were further subjected to bioinformatics using PipeCraft and WIMP (Juul et al., 2015) in parallel. The options in PipeCraft included demultiplexing of metabarcoding reads allowing no mismatches to the barcode, followed by blastn search using default settings. The sequencing adaptors were removed by a custom script.

Given the poor overall sequence quality, traditional OTU-based approaches are not applicable to the MinION data; therefore, we mapped reads based on their best matches to database sequences in the UNITE

reference database, following previous nanopore sequencing studies (Benitez-Paez et al., 2016; Kerkhof et al., 2017). Limitations of this approach are outlined in the Discussion.

To maximize the speed of pathogen diagnosis, we used a metagenomics-based approach with the MinION instrument. For this, we concentrated the genomic DNA of select potato samples using FavorPrep Gel/PCR Purification kit (Favorgen Biotech Corp., Vienna, Austria). For library preparation, the Rapid Sequencing kit (SQK-RAD004) and Rapid Barcoding Sequencing kit (SQK-RBK004) were used following manufacturer's instructions. Base calling was performed as described above. Both 'passed' and 'failed' sequences were used in further bioinformatics analyses as implemented in the Pipecraft. Taxonomic reference libraries included UNITE 7.2 and SILVA 132 (Quast et al., 2013), and INSD for extracted rRNA gene reads and other reads separately. The UNITE database was merged with a database of oomycetes created based on ITS sequences in INSD. Using this reference, chimera checking was performed using Uchime on demultiplexed reads. Specific information about the amount of initial material, sequencing time and sequencing runs is given in Table 1.

Sanger sequences of potato samples have been deposited in the UNITE database (https://unite.ut.ee/; accessions UDBxxxxx-UDBxxxxx). Raw sequence data of MinION and Sequel are available from the PlutoF digital repository. Sample-by-OTU tables used in these analyses are given in supplementary material (Tables S1 and S2).

ACKNOWLEDGEMENTS

We thank V. Kisand for constructive comments on the manuscript and A. Tooming-Klunderud for running PacBio sequencing. Financial contribution was provided by the Estonian Science Foundation (grants PUT1399, PUT1317, PSG136, IUT21-04, IUT 36-2, MOBERC13, ECOLCHANGE). Author contributions: KL, RD and LT designed the study; RK, KL, KA and RD provided material; RP and KL performed HTS analyses; MB, SA and LT analyzed data; LT wrote the paper with input from all co-authors.

APPENDIX

Additional File 1:

Table S1 Sample-by-OTU table of metabarcoding studies of conifer needles.

Table S2 Sample-by-OTU table of metabarcoding studies of potato tissues.

REFERENCES

Anslan S, Bahram M, Hiiesalu I, Tedersoo L. 2017. PipeCraft: Flexible open-source toolkit for bioinformatics analysis of custom high-throughput amplicon sequencing data. Mol Ecol Res 17:e234–e240.

Ashikawa S, Tarumoto N, Imai K, Sakai J, Kodana M, Kawamura T, Ikebuchi K, Murakami T, Mitsutake K, Maesaki S, Maeda T. Rapid identification of pathogens from positive blood culture bottles with the MinION nanopore sequencer. J. Med Microbiol 67:1589-1595.

Bengtsson-Palme J, Ryberg M, Hartmann M, Branco S, Wang Z, Godhe A, de Wit P, Sanchez-Garcia M, Ebersberger I, de Sousa F, Amend AS, Jumpponen A, Unterseher M, Kristiansson E, Abarenkov K, Bertrand YJK, Sanli K, Eriksson KM, Vik U, Veldre V, Nilsson RH. 2013. Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. Meth Ecol Evol 4:914-919.

Benitez-Paez A, Portune KJ, Sanz Y. 2016. Species-level resolution of 16S rRNA gene amplicons sequenced through the MinION™ portable nanopore sequencer. GigaScience 5:4.

Benitez-Paez A, Sanz Y. 2017. Multi-locus and long amplicon sequencing approach to study microbial diversity at species level using the MinION portable nanopore sequencer. GigaScience 6:1-12.

Bik HM, Porazinska D, Creer S, Caporaso JG, Knight R, Thomas WK. 2012. Sequencing our way towards understanding global eukaryotic biodiversity. Trends Ecol Evol 27:233-243.

Bradley P, Gordon NC, Walker TM, Dunn L, Heys S, Huang B, Earle S, Pankhurst LJ, Anson L, De Cesare M, Piazza P. 2015. Rapid antibiotic-resistance predictions from genome sequence data for Staphylococcus aureus and Mycobacterium tuberculosis. Nature Comm 6:10063.

Bronzato Badial A, Sherman D, Stone A, Gopakumar A, Wilson V, Schneider W, King J. 2018. Nanopore sequencing as a surveillance tool for plant pathogens in plant and insect tissues. Plant Dis 20:PDIS-04.

Calus ST, Ijaz UZ, Pinto AJ. 2018. NanoAmpli-Seq: A workflow for amplicon sequencing for mixed microbial communities on the nanopore sequencing platform. GigaScience 7:140.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinform 10:421.

Castro-Wallace SL, Chiu CY, John KK, Stahl SE, Rubins KH, McIntyre AB, Dworkin JP, Lupisella ML, Smith DJ, Botkin DJ, Stephenson TA. 2017. Nanopore DNA sequencing and genome assembly on the International Space Station. Sci Rep 7:18022.

Charalampous T, Richardson H, Kay GL, Baldan R, Jeanes C, Rae D, Grundy S, Turner DJ, Wain J, Leggett RM, Livermore DM. 2018. Rapid Diagnosis of Lower Respiratory Infection using Nanopore-based Clinical Metagenomics. BioRxiv 2018:387548.

Cline LC, Zak DR. 2015. Initial colonization, community assembly and ecosystem function: fungal colonist traits and litter biochemistry mediate decay rate. Mol Ecol 24:5045–5058.

Comtet T, Sandionigi A, Viard F, Casiraghi M. 2015. DNA (meta)barcoding of biological invasions: a powerful tool to elucidate invasion processes and help managing aliens. Biol Invas 17:905–922.

578 Cornelis S, Gansemans Y, Vander Plaetsen AS, Weymaere J, Willems S, Deforce D, Van Nieuwerburgh F. 2019.
579 Forensic tri-allelic SNP genotyping using nanopore sequencing. For Sci Int Genet 38:204-210.
580 Costello M, Fleharty M, Abreu J, Farjoun Y, Ferriera S, Holmes L, Granger B, Green L, Howd T, Mason T,
581 Vicente G. 2018. Characterization and remediation of sample index swaps by non-redundant dual indexing on
582 massively parallel sequencing platforms. BMC Genomics 19:332.
583 Cusco A, Catozzi C, Vines J, Sanchez A, Francino O. 2018. Microbiota profiling with long amplicons using
584 Nanopore sequencing: full-length 16S rRNA gene and whole rrn operon. F1000Research. 2018:7.
585 Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. 2011. UCHIME improves sensitivity and speed of
586 chimera detection. Bioinformatics 27:2194-2200.
587 Edgar RC. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature Meth 10:996–
588 998.
589 Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A. 2009.
590 Real-time DNA sequencing from single polymerase molecules. Science 323:133-138.
591 Garalde DR, Snell EA, Jachimowicz D, Sipos B, Lloyd JH, Bruce M, Pantic N, Admassu T, James P, Warland A,
592 Jordan M. 2018. Highly parallel direct RNA sequencing on an array of nanopores. Nature Meth 15:201-208.
593 Giordano F, Aigrain L, Quail MA, Coupland P, Bonfield JK, Davies RM, Tischler G, Jackson DK, Keane TM, Li
594 J, Yue JX. De novo yeast genome assemblies from MinION, PacBio and MiSeq platforms. Sci Rep. 2017 Jun
595 21;7(1):3935.
596 Goto Y, Yanagi I, Matsui K, Yokoi T, Takeda KI. 2018. Identification of four single-stranded DNA
597 homopolymers with a solid-state nanopore in alkaline CsCl solution. Nanoscale 10:20844-50.
598 Hebert PD, Braukmann TW, Prosser SW, Ratnasingham S, Ivanova NV, Janzen DH, Hallwachs W, Naik S,
599 Sones JE, Zakharov EV. 2018. A Sequel to Sanger: amplicon sequencing that scales. BMC Genomics 19:219.
600 Hildebrand F, Tadeo R, Wright AY, Bork P, Raes J. 2014. LotuS: an efficient and user-friendly OTU processing
601 pipeline. Microbiome 2:30.
602 Hu Y, Green G, Milgate A, Stone E, Rathjen J, Schwessinger B. 2019. Pathogen detection and microbiome
603 analysis of infected wheat using a portable DNA sequencer. bioRxiv 2019:429878.
604 Hyde KD, Udayanga D, Manamgoda DS, Tedersoo L, Larsson E, Abarenkov K, Bertrand Y, Oxelman B,
605 Hartmann M, Kauserud H, Ryberg M, Kristiansson E, Nilsson RH. 2013. Incorporating molecular data in fungal
606 systematics: a guide for aspiring researchers. Curr Res Environ Appl Mycol 3:1-32.
607 Jain M, Koren S, Miga KH, Quick J, Rand AC, Sasani TA, Tyson JR, Beggs AD, Dilthey AT, Fiddes IT, Malla S.
608 2018. Nanopore sequencing and assembly of a human genome with ultra-long reads. Nature Biotechnol 36:338.
609 Johnson SS, Zaikova E, Goerlitz DS, Bai Y, Tighe SW. 2017. Real-time DNA sequencing in the Antarctic dry
610 valleys using the Oxford Nanopore sequencer. J Biomol Techn 28:2.
611 Juul S, Izquierdo F, Hurst A, Dai X, Wright A, Kulesha E, Pettett R, Turner DJ. 2015. What's in my pot? Real-
612 time species identification on the MinION. bioRxiv 2015:030742.
613 Karawdeniya BI, Bandara YN, Nichols JW, Chevalier RB, Dwyer JR. 2018. Surveying silicon nitride nanopores
614 for glycomics and heparin quality assurance. Nature comm 9:3278.
615 Karlsson I, Edel Hermann V, Gautheron N, Durling MB, Kolseth AK, Steinberg C, Persson P, Friberg H. 2016.
616 Genus-specific primers for study of *Fusarium* communities in field samples. Appl Environ Microbiol 82:491-
617 501.
618 Karst SM, Dueholm MS, McIlroy SJ, Kirkegaard RH, Nielsen PH, Albertsen M. 2016. Retrieval of a million
619 high-quality, full-length microbial 16S and 18S rRNA gene sequences without primer bias. Nature Biotechnol
620 36:190–195.
621 Kashyap PL, Rai P, Kumar S, Chakdar H, Srivastava AK. 2017b. DNA barcoding for diagnosis and monitoring of
622 fungal plant pathogens. In: Singh BP, Gupta VK (eds). Molecular Markers in Mycology. Pp. 87-122.
623 Kerkhof LJ, Dillon KP, Häggblom MM, McGuinness LR. 2017. Profiling bacterial communities by MinION
624 sequencing of ribosomal operons. Microbiome 5:116.
625 Kilianski A, Haas JL, Corriveau EJ, Liem AT, Willis KL, Kadavy DR, Rosenzweig CN, Minot SS. 2016.
626 Bacterial and viral identification and differentiation by amplicon sequencing on the MinION nanopore
627 sequencer. Gigascience 4:12.

Kõljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AFS, Bahram M et al. 2013. Towards a unified paradigm for sequence-based identification of Fungi. Mol Ecol 22:5271–5277.

Kyaschenko Y, Clemmensen K, Hagenbo A, Karltun E, Lindahl B. 2017. Shift in fungal communities and associated enzyme activities along an age gradient of managed *Pinus sylvestris* stands. ISME J 11:863–874.

Lu H, Giordano F, Ning Z. 2016. Oxford Nanopore MinION sequencing and genome assembly. Gen Prot & Bioinform 14:265-279.

Mikheyev AS, Tin MMY. 2014. A first look at the Oxford Nanopore MinION sequencer. Mol Ecol Res 14:1097–1102.

Mojarro A, Hachey J, Ruvkun G, Zuber MT, Carr CE. 2018. CarrierSeq: a sequence analysis workflow for low-input nanopore sequencing. BMC Bioinform 19:108.

Mosher JJ, Bowman B, Bernberg EL, Shevchenko O, Kan J, Korlach J. 2014. Improved performance of the PacBio SMRT technology for 16S rDNA sequencing. J of Microbiol Meth 104:59–60.

Nilsson RH, Anslan S, Bahram M, Wurzbacher C, Baldrian P, Tedersoo L. 2019. Mycobiome diversity: high-throughput sequencing and identification of fungi. Nature Rev Microbiol, in press.

Pomerantz A, Peñafiel N, Arteaga A, Bustamante L, Pichardo F, Coloma LA, Barrio-Amorós CL, Salazar-Valenzuela D, Prost S. 2018. Real-time DNA barcoding in a rainforest using nanopore sequencing: opportunities for rapid biodiversity assessments and local capacity building. GigaScience 7:033.

Pontefract A, Hachey J, Zuber MT, Ruvkun G, Carr CE. 2018. Sequencing nothing: Exploring failure modes of nanopore sensing and implications for life detection. Life Sci Space Res 18:80-86.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucl Ac Res 41:D590–D596.

Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, Bore JA, Koundouno R, Dudas G, Mikhail A, Ouédraogo N. 2016. Real-time, portable genome sequencing for Ebola surveillance. Nature 530:228-232.

Quick J, Quinlan AR, Loman NJ. 2014. A reference bacterial genome dataset generated in the MinION portable single-molecule nanopore sequencer. GigaScience 3:22.

Rard T. 2018. Pacific Biosciences of California, Inc. Announces Third Quarter 2018 Financial Results. Sci Technol News 2018:11.1

Restrepo-Pérez L, Joo C, Dekker C. 2018. Paving the way to single-molecule protein sequencing. Nature Nanotechnol 13:786.

Reuter JA, Spacek DV, Snyder MP. 2015. High-throughput sequencing technologies. Mol Cell 58:586-597.

Rhoads A, Au KF. 2015. PacBio sequencing and its applications. Gen Prot Bioinform 13:278-289.

Riit T, Tedersoo L, Drenkhan R, Runno-Paurson E, Kokko H, Anslan S. 2016. Oomycete-specific ITS primers for identification and metabarcoding. MycoKeys 14:17-30.

Riit T, Tedersoo L, Drenkhan R, Runno-Paurson E, Kokko H, Anslan S. 2018. Corrigendum for:"Oomycete-specific ITS primers for identification and metabarcoding" published in MycoKeys. MycoKeys 41:119.

Schloss PD, Westcott SL, Jenior ML, Highlander SK. 2016. Sequencing 16S rRNA gene fragments using the PacBio SMRT DNA sequencing system. Peer J 4:e1869.

Schmidt K, Mwaigwisya S, Crossman LC, Doumith M, Munroe D, Pires C, Khan AM, Woodford N, Saunders NJ, Wain J, O'grady J. 2017. Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. J Antimicrob Chemother 72:104–114.

Schnell IB, Bohmann K, Gilbert MTP. 2015. Tag jumps illuminated – reducing sequence-to-sample misidentifications in metabarcoding studies. Mol Ecol Res 15:1289–1303.

Shin J, Lee S, Go MJ, Lee SY, Kim SC, Lee CH, Cho BK. 2016. Analysis of the mouse gut microbiome using full-length 16S rRNA amplicon sequencing. Sci Rep 6:29681.

Simpson JT, Workman RE, Zuzarte PC, David M, Dursi LJ, Timp W. 2017. Detecting DNA cytosine methylation using nanopore sequencing. Nature Meth 14:407.

Srivathsan A, Baloğlu B, Wang W, Tan WX, Bertrand D, Ng AH, Boey EJ, Koh JJ, Nagarajan N, Meier R. 2018. A Min ION™-based pipeline for fast and cost-effective DNA barcoding. Mol Ecol Res 18:1035-1049.

Tedersoo L, Drenkhan R, Anslan S, Morales-Rodrigues C, Cleary M. 2019. High-throughput identification and diagnostics of pathogens and pests: overview and practical recommendations. Mol Ecol Res 19: 47–76.

679 Tedersoo L, Lindahl B. 2016. Fungal identification biases in microbiome projects. Environ Microbiol Rep. 8:774-
680 779.
681 Tedersoo L, Tooming-Klunderud A, Anslan S. 2018. PacBio metabarcoding of fungi and other eukaryotes: biases
682 and perspectives. New Phytol 217:1370-1385.
683 Wagner J, Coupland P, Browne HP, Lawley TD, Francis SC, Parkhill J. 2016. Evaluation of PacBio sequencing
684 for full-length bacterial 16S rRNA gene classification. BMC Microbiol 16:274.
685 Walder F, Schlaeppi K, Wittwer R, Held AY, Vogelgsang S, van der Heijden MG. 2017. Community profiling of
686 Fusarium in combination with other plant-associated fungi in different crop species using SMRT sequencing.
687 Front Plant Sci 2017:8.
688 Weirather JL, de Cesare M, Wang Y. 2017. Comprehensive comparison of Pacific Biosciences and Oxford
689 Nanopore Technologies and their applications to transcriptome analysis. F1000Research 6:100.
690 Veldre V, Abarenkov K, Bahram M, Martos F, Selosse M-A, Tamm H, Kõljalg U, Tedersoo L. 2013. Evolution
691 of nutritional modes of Ceratobasidiaceae (Cantharellales, Basidiomycota) as revealed from publicly available
692 ITS sequences. Fung Ecol 6:256-268.
693 White R, Pellefigues C, Ronchese F, Lamiable O, Eccles D. 2017. Investigation of chimeric reads using the
694 MinION. F1000Research 6:631.
695 Vilgalys R, Hester M. 1990. Rapid genetic identification and mapping of enzymatically amplified ribosomal DNA
696 from several Cryptococcus species. J Bacteriol 172:4238-4246.
697 Volden R, Palmer T, Byrne A, Cole C, Schmitz RJ, Green RE, Vollmers C. 2018. Improving nanopore read
698 accuracy with the R2C2 method enables the sequencing of highly multiplexed full-length single-cell cDNA.
699 Proc Natl Acad Sci USA 115:9726-9731.
700 Votintseva AA, Bradley P, Pankhurst L, del Ojo Elias C, Loose M, Nilgiriwala K, Chatterjee A, Smith EG,
701 Sanderson N, Walker TM, Morgan MR. 2017. Same-day diagnostic and surveillance data for tuberculosis via
702 whole genome sequencing of direct respiratory samples. J Clin Microbiol 8:02483.
703 Wurzbacher C, Larsson E, Bengtsson-Palme J, Van den Wyngaert S, Svantesson S, Kristiansson E, Kagami M,
704 Nilsson RH. 2018. Introducing ribosomal tandem repeat barcoding for fungi. Mol Ecol Res 19:118-127.
705
706

FIGURE LEGENDS [and figures]

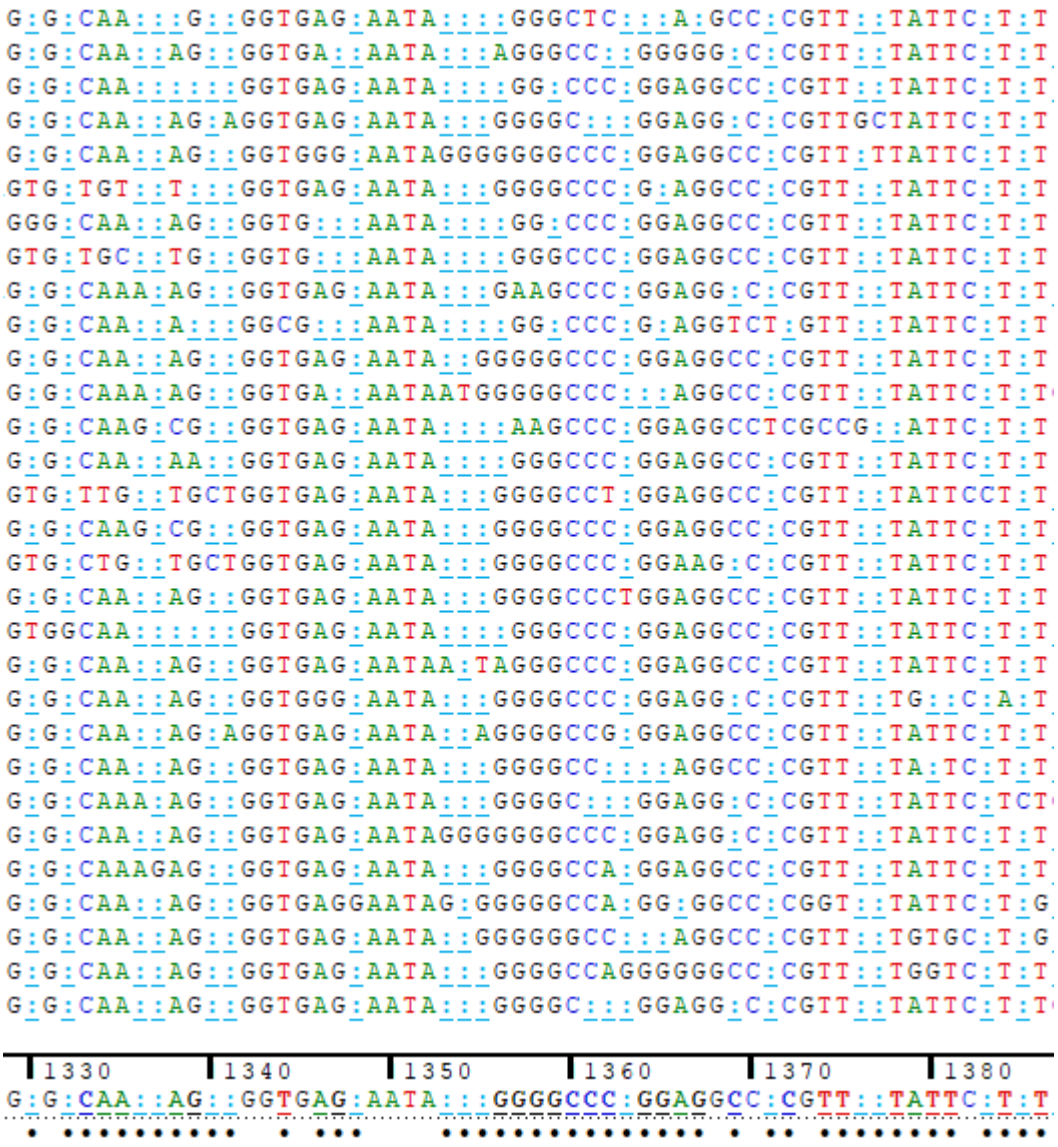

FIG 1. Screenshot example of multiple sequence alignment of MinION reads mapped to the contaminant *Coniothyrium* sp. using Sequencher 5.1 software (GeneCodes Corp., Ann Arbor, MI, USA). Note the error-infested double homopolymeric region (center) and a relatively accurately recorded region upstream.

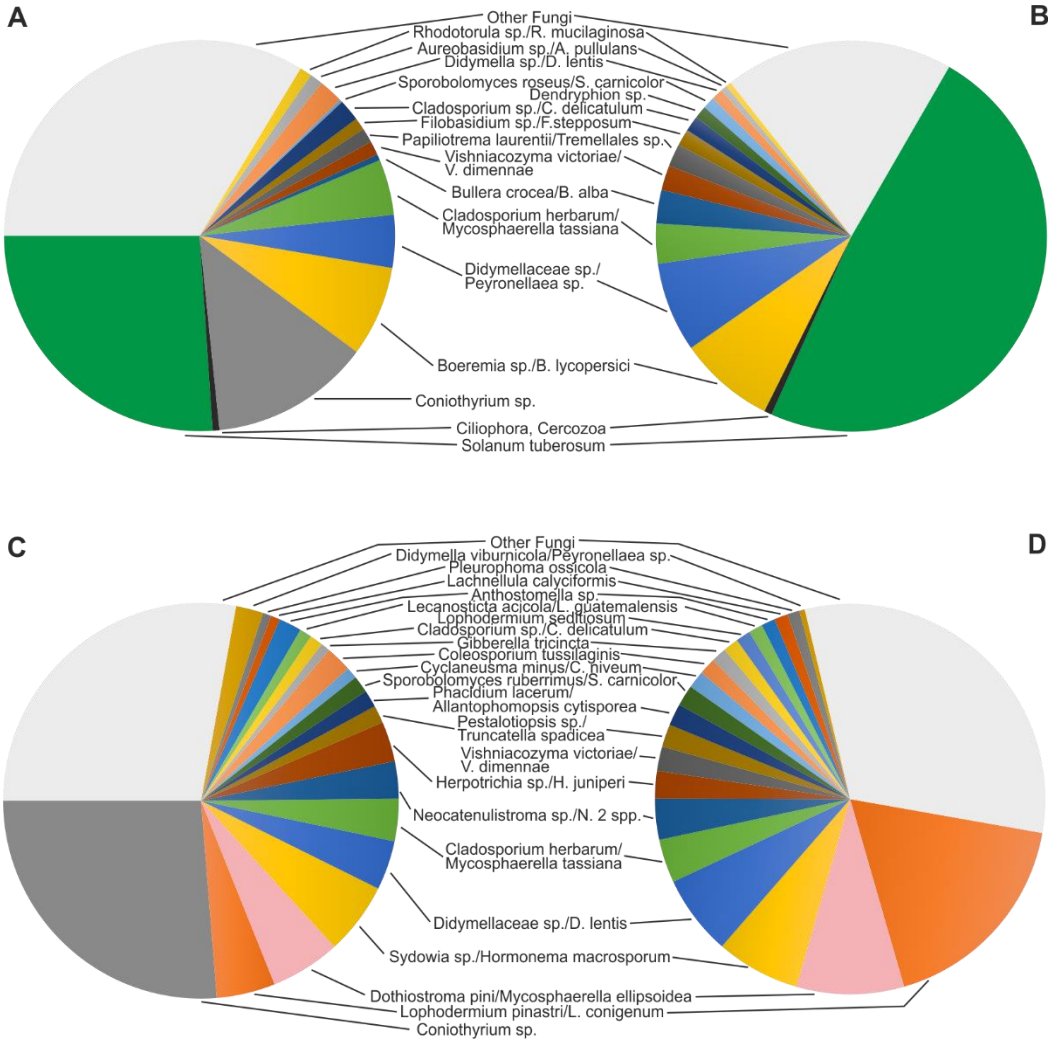

FIG 2 Pie diagrams demonstrating dominance of higher taxa and OTUs in a) ONT1 (needles; MinION), b) ONT1S (needles; Sequel), c) ONT2 (potato; MinION) and ONT2S runs (potato; Sequel) based on relative abundance of identified reads in the ITS1catta + ITS4ngsUni amplicons. Taxon names in Sequel and MinION data sets have been put into correspondence (/).

Table 1 Detailed information about MinION and Sequel runs

Run ID (DNA quantity, ng)	Samples (n)	Primers	Chemistry	Cell	N° reads obtained/ qualified (time, min)
ONT1 (1165)	Conifer needles (36)	ITS1catta + ITS4ngsUni	MinION 1D: SQK-LSK109	Flowcell #1 (new)	1,053,693/ 186,586 (1440)
ONT1S (1000)	Conifer needles	ITS1catta +	Sequel	SMRT cell #1	167,864/ 121,965

	(36)	ITS4ngsUni			(600)
ONT2 (2002)	Potato leaves and tubers (35)	ITS1catta + ITS4ngsUni; ITS1Oo + ITS4ngsUni	MinION 1D: SQK-LSK109	Flowcell #1 (2nd use)	1,194,242/ 36,779 (343)
ONT2S (1000)	Potato leaves and tubers (35)	ITS1catta + ITS4ngsUni; ITS1Oo + ITS4ngsUni	Sequel	SMRT cell #2	177,635/ 75,573 (600)
ONT2a (1076)	Potato leaves (8)	ITS1catta) + ITS4ngsUni	MinION 1D: SQK-LSK109	Flowcell #1 (3rd use)	130,130/ 14,155 (260)
ONT2b (926)	Potato leaves (8)	ITS1catta + ITS4ngsUni	MinION 1D: SQK-LSK109	Flowcell #2 (new)	Failed
ONT2f (473)	Potato leaves (8)	ITS1catta + LR1	MinION 1D[2]: SQK-LSK308	Flowcell #3 (new)	5433/ 265 (77)
ONT2g (69)	Potato tuber (1)	Metagenome	MinION 1D: SQK-RAD004	Flowcell 2 (2nd use)	466,488/ 66,133 (251)
ONT2h (448)	Potato tuber (1)	ITS1catta + LR11	MinION 1D: SQK-LSK109	Flowcell 2 (3rd use)	767,611/ 342,923 (165)
ONT2i (31)	Potato tuber (1)	Metagenome	MinION 1D: SQK-RAD004	Flowcell 4 (new)	1142/ 436 (50)

Table 2 Identification of Fungi in needle samples. Numbers and percentages in Sequel and MinION columns indicate the number of fungal reads and per cent of sequences assigned to particular OTUs. Asterisks indicate taxon names that correspond to each other based on >98% sequence similarity.

Sample ID	Species-specific primers:	Sequel (reads per sample: dominant taxa)	MinION (reads per sample: dominant taxa)
115	Negative	6275: LoPi 35%, CoTu 27%, RhySp 3%	16543: CoSp 44%, CoTu 21%, HeJu 2%
117	Negative	5048: LoCo 49 %, DiSp 25%, ViVi 9%	14709: LoCo 42%, DiLe 13%, CoSp 12%
118	LeAc	5511: NeGe* 36%, DiSp 10%, SySp 9%	8109: NeAb* 17%, NeMi* 15%, CoSp 9%
119	Negative	4460: LoPi 34%, RhySp 16%, CySp 11%	8911: CoSp 49%, ChaeSp 9%, PhEu 3%
123	Negative	6121: DiSp* 49%, PlOs 19%, SySp 19%	7519: DiLe* 25%, PlOS 14%, HoMa 14%
125	Negative	2753: SySp* 56%, ViVi 8%, SpRu 7%	2500: HoMa* 49%, CoSp 13%, MyTa 5%
127	DoSe	1801: LaCa 68%, SySp 25%, InSp 6%	1450: LaCa 67%, HoMa 21%, CoSp 2%
139	Negative	1248: ClHe* 22%, AuSp 12%, LaCa 8%	2005: MyTa* 20%, CoSp 18%, AuPu 11%
141	Negative	5545: DiSp* 21%, SySp** 21%, SpRu 16%	11966: HoMa** 17%, DiLe* 12%, CoSp 10%
142	Negative	2421: LoPi 40%, NeGe* 28%, DoPi 9%	1590: CoSp 43%, NeMi* 12%, MyEl 7%
148	Negative	4762: ClHe* 40%, HerSp** 35%, SpRu 15%	8556: HeJu** 37%, MyTa* 35%, SpCa 9%
154	Negative	5892: ViVi 18%, LoCo 13%, HeSp* 12%	13465: CoSp 31%, HeJu *12%, LoCo 10%
2404	NA	3777: DoPi* 100%	5281: PaCa* 76%, CoSp 9%, PsOp 2%
3904	NA	598: LeAc* 52%, AlAl 23%, HaOr 23%	1413: LeGu* 44%, CoSp 24%, AlIr 19%
3906	NA	282: DoPi* 68%, DoSe** 32%	907: MyEl** 70%, CoSp 7%, PaCa* 4%
4154	DoPi, DoSe, LeAc	2741: LoPi 94%, AnCo* 3%, NeGe 1%	5281: CoSp 86%, AnFo* 3%, HoMa 1%
4162	Negative	3674: LoPi 62%, DoPi 9%, CeFe 6%	1951: CoSp 57%, CeFe 10%, MyEl 9%
4180	DoSe	602: DoPi* 32%, LoPi 22%, ClHe 18%	1027: MyEl* 27%, CoSp 26%, MyTa 21%
4181	Negative	2853: LoPi 63%, LoSp 6%, LoCo 5%	4582: CoSp 62%, LoSp 6%, LoCo 5%
4192	DoSe	5311: LoPi 32%, PhLa 18%, CyMi 13%	6552: CoSp 33%, AlCy 13%, CyNi 10%
4194	LeAc	4388: PeSp* 44%, NeGe 15%, RhiSp 11%	4001: TrSp* 40%, NeAb 8%, ScSp 7%
4195	LeAc	1935: SySp* 52%, LoPi 11%, PeSp 10%	4938: HoMa* 44%, CoSp 15%, TrSp 10%
4197	DoSe	4387: LoPi* 46%, DoPi** 25%, RhySp 8%	6031: CoSp 49%, MyEl** 19%, LoSp* 14%
4220	Negative	4349: LoPi 58%, AnSp 32%, PhLa 5%	11258: CoSp 56%, AnSp 28%, AlCy 3%
4221	DoSe	671: LoPi 47%, ClHe* 43%, AuSp 3%	2039: CoSp 47%, MyTa* 37%, MyEl 4%
4222	DoSe	5232: LoPi 48%, DoPi*7%, EuSp 6%	5333: CoSp 49%, MyEl* 7%, PhSp 5%
4223	DoSe, LeAc	1029: SySp* 25%, DiSp 10%, DoPi 8%	1949: HoMa* 20%, CoSp 8%, MyEl 8%
5136	Negative	158: AsSy* 73%, DiVi 27%	825: AsSy* 36%, CoSp 13%, CoSp 11%
5137	Negative	1149: HeAn* 56%, DoPi19%, DoSe 13%	674: HeAb* 45%, MyEl 31%, CoSp 3%
5146	Negative	5414: GiTr 17%, DiSp 17%, CeSp 14%	6450: CeSp 14%, AlCy 11%, GiTr 10%
5148	Negative	463: DiSp* 45%, ArSp 11%, GiTr 11%	259: DiLe* 29%, PeySp 10%, GiTr 7%

5151	Negative	3877: CyMi* 22%, ClHe 11%, ClSp 9%	2503: CyNi* 18%, CoSp 15%, MyTa 11%
5186	Negative	474: ClHe* 70%, ViVi 8%, DiSp 5%	268: MyTa*, 66%, CoSp 15%, AuPu 4%
5194	Negative	925: DiVi* 26%, HelSp 20%, MaOb 11%	505: PeSp* 17%, MyTa 10%, DoSp 10%
5195	Negative	320: RaHy* 57%, ZySp1 19%, ExSp 18%	130: RaSp* 41%, CaSp 12%, ZyVe 11%
5297	Negative	4508: ZySp2* 18%, RhiSp 16%, RhoSp 5%	3963: ZyVe* 17%, CoSp 10%, ScSp 10%
5307	Negative	1986: ZySp1* 20%, ExSp 12%, ClHe 10%	765: ZyVe* 13%, CoSp 11%, MyTa 10%
14374	ND[2]	1637: LeAc* 38%, DoSp 13%, TeSp 10%	1374: LeGu* 31%, PeIn 10%, CoSp 7%
14378	ND	3561: SySp* 46%, NeSp 23%, ChSp 6%	2861: HoMa* 42%, PlSt 16%, CoSp 8%

[1]Abbreviations for species: AlAl, *Alternaria alternata*; AlCy, *Allantophomopsis cytisporea*; AlIr, *Alternaria iridiaustralis*; AnCo, *Anthostomella conorum*; AnFo, *Anthostomella formosa*; AnSp, *Anthostomella* sp.; ArSp, *Articulospora* sp.; AsSy, *Aspergillus sydowii*; AuPu, *Aureobasidium pullulans*; AuSp, *Aureobasidium* sp.; CaSp, Capnodiales sp.; CeFe, *Cenangium ferruginosum*; CeSp, Ceratobasidiaceae sp.; ChSp, *Chalara* sp.; ChaeSp, *Chaetothyriales* sp.; ClHe, *Cladosporium herbarum*; ClSp, *Cladosporium* sp.; CoSp, *Coniothyrium* sp.; CoTu, *Coleosporium tussilaginis*; CyMi, *Cyclaneusma minus*; CyNi, *Cyclaneusma niveum*; CySp, *Cyphellophora* sp.; DiLe, *Didymella lentis*; DiSp, Didymellaceae sp.; DiVi, *Didymella viburnicola*; DoPi, *Dothistroma pini*; DoSe, *Dothistroma septosporum*; DoSp, Dothideomycetes sp.; EuSp, Eurotiomycetes sp.; ExSp, *Extremus* sp.; GiTr, *Gibberella tricincta*; HaOr, *Hannaella oryzae*; HeAb, *Heterobasidion abietinum*; HeAn, *Heterobasidion annosum*; HeJu, *Herpotrichia juniperi*; HelSp, Helotiales sp.; HerSp, *Herpotrichia* sp.; HoMa, *Hormonema macrosporum*; InSp, Insecta sp.; LaCa, *Lachnellula calyciformis*; LeAc, *Lecanosticta acicola*; LeGu, *Lecanosticta guatemalensis*; LoCo, *Lophodermium conigenum*; LoPi, *Lophodermium pinastri*; LoSp, *Lophodermium* sp.; MaOb, *Malassezia obtusa*; MyEl, *Mycosphaerella ellipsoidea*; MyTa, *Mycosphaerella tassiana*; NeAb, *Neocatenulostroma abietis*; NeGe, *Neocatenulostroma germanicum*; NeMi, *Neocatenulostroma microsporum*; NeSp, *Nectria* sp.; PaCa, *Passalora californica*; PeIn, *Perusta inaequalis*; PeSp, *Pestalotiopsis* sp.; PeySp, *Peyronellaea* sp.; PhEu, *Phaeococcomyces eucalypti*; PhLa, *Phacidium lacerum*; PhSp, *Phaeomoniella* sp.; PlOs, *Pleurophoma ossicola*; PlSt, *Pleonectria strobi*; RaHy, *Ramularia hydrangeae-macrophyllae*; RaSp, *Ramularia* sp.; RhiSp, *Rhizosphaera* sp.; RhoSp, *Rhodotorula* sp.; RhySp, *Rhytismataceae* sp.; ScSp, *Scleroconidioma sphagnicola*; SpCa, *Sporobolomyces carnicolor*; SpRu, *Sporobolomyces ruberrimus*; SySp, *Sydowia* sp.; ZySp1, *Zymoseptoria* sp.; ZySp2, *Zymoseptoria* sp.; ZyVe, *Zymoseptoria verkleyi*; TeSp, Teratosphaeriaceae sp.; TrSp, *Truncatella spadicea*; ViVi, *Vishniacozyma victoriae;*

[2]ND, not determined

Table 3 Identification of Fungi in potato samples. Numbers and percentages in Sequel and MinION columns indicate the number of fungal reads and per cent of sequences assigned to particular OTUs. Asterisks indicate taxon names that correspond to each other based on >98% sequence similarity.

Sample	Sanger[1]	Sequel	MinION
KL001	failed	429: ClHe 10%, BuCr 8%, SpRo 7%	310: CoSp 24%, CerSp 16%, LeGu 11%
KL002	failed	282: KoCh 27%, PeEx 13%, SpRo 8%	271: MyEl 27%, KoCh 9%, CoSp 8%
KL003	failed	770: BuAl 21%, FiSp 14%, ClSp 12%	570: BuAl 17%, ClDe 12%, CoSp 11%
KL004	failed	209: EpNi 12%, ClHe 10%, SpRo 10%	444: NeAb 13%, CoSp 11%, NeMi 9%
KL005	FiWi	820: FiWi 41%, DioSp 5%, BuAu 5%	848: FiWi 26%, MyTa 13%, CoSp 10%
KL006	failed	433: SpRo 17%, LeSp 8%, ViVi 7%	275: CoSp 15%, HoMa 11%, MyTa 8%
KL007	failed	1226: FiSp* 26%, SpSp 11%, ClHe 7%	745: CoSp 15%, FiSt* 13%, SpSp 7%
KL008	failed	1948: PaLa* 49%, SpRo 9%, ClHe 6%	579: TreSp* 36%, CoSp 14%, MyTa 8%
KL009	failed	1753: DiSp* 15%, DiPo 12%, ClHe 11%	735: PeySp* 12%, CoSp 11%, MyTa 8%
KL010	BoEx	868: BoSp* 47%, FiSp 12%, SpRo 5%	970: BoLy* 26%, CoSp 18%, FiSt 5%
KL011	BoEx	1131: BoSp* 71%, ExEq 9%, ClHe 4%	868: BoLy* 52%, ExEq 8%, MyTa 4%
KL012	BoEx	1421: BoSp* 96%, ViTe 2%, MySp 1%	1288: BoLy* 50%, CoSp 14%, HeJu 3%
KL013	DioSp	3148: ClHe* 17%, BuCr 9%, ViVi 8%	2599: CoSp 24%, MyTa* 13%, HyaSp 4%
KL014	failed	2437: CyMa 14%, ViVi* 10%, LeSp 7%	2549: CoSp 30%, ViDi* 10%, CoTu 6%
KL015	BoEx	890: BoSp* 92%, BuCr 1%, FiSp 1%	1197: BoLy* 37%, LoCo 19%, CoSp 8%
KL016	failed	1110: ViVi 18%, PaLa* 15%, LeSp 12%	1005: CoSp 29%, TreSp* 10%, HoMa 7%
KL017	BoEx	1325: BoSp* 65%, FiSp 8%, ViVi 4%	519: BoLy* 44%, CoSp 15%, TrSp 7%
KL018	failed	426: ClHe* 35%, PlSp 13%, FiSp 13%	338: CoSp 39%, MyTa* 11%, MyEl 9%

KL019	failed	1399: AuSp* 13%, BuCr 11%, ViTe 11%	596: CoSp 32%, AuPu* 13%, BoLy 7%
KL020	BoEx	531: BoSp* 49%, SpSp 9%, FiSp 9%	317: CoSp 29%, BoLy* 17%, SpSp 5%
KL021	failed	657: ClHe* 35%, ViVi 15%, ClSp 10%	589: MyTa* 30%, CoSp 20%, AuPu 9%
KL022	ClSp	400: ClHe* 42%, BuCr 13%, AuSp 9%	970: RhMu 33%, MyTa* 23%, CoSp 11%
KL023	failed	64: AuSp 42%, ClHe 33%, AlAl 8%	457: CoSp 46%, AnSp 24%, AuSp 7%
KL024	ClSp	111: ClHe* 51%, SpRo 9%, DiBu 6%	107: MyTa* 29%, CoSp 20%, MyEl 10%
KL025	failed	821: ViVi 18%, DiSp* 18%, ClSp 13%	850: CoSp 30%, PeySp* 12%, ClDe 8%
KL026	ClSp	772: ClHe* 16%, SuGr 15%, BuCr 12%	479: CoSp 42%, MyTa* 11%, AtSp 5%
KL027	DioSp	1099: BuCr 65%, ViVi 9%, ClHe 7%	165: CoSp 75%, CeFe 4%, BlGr 3%
KL028	failed	2740: DiSp* 97%, HaVe 3%, CuMo 0%	1473: PeySp* 35%, DiLe 22%, PhBu 6%
KL029	failed	1245: DiSp* 54%, BoSp** 46%, PlCu 0%	465: BoLy** 28%, PeySp* 21%, DiLe 11%
KL030	failed	8: MoSp* 100%	122: PlCu* 17%, ZyVe 11%, ScSp 8%
KL031	failed	1931: DiSp* 93%, PeBi 4%, BoSp 1%	420: PeySp* 37%, DiLe 13%, PlSp 6%
KL032	failed	308: PlCu 43%, PsSp 22%, CuMo 17%	364: PlCu 23%, PlOr 11%, GeAs 10%
KL033	failed	1422: DeSp 50%, MoSp* 38%, NeSp 11%	309: PlCu* 41%, PlSp 36%, NeSp 8%
KL034	failed	223: PenSp 80%, MoSp 8%, PlCu 3%	364: CeSp 13%, AlCy 12%, GiTr 8%
KL035	failed	1059: PeBi 41%, PenSp 38%, PeBr 10%	185: PeBi 42%, PeAe 23%, PenGl 11%

[1]Abbreviations for species: AlAl, *Alternaria alternata*; AlCy, *Allantophomopsis cytisporea*; AnSp, *Anthostomella* sp.; AtSp, Atheliaceae sp.; AuPu, *Aureobasidium pullulans*; AuSp, *Aureobasidium* sp.; BlGr, *Blumeria graminis*; BoEx, *Boeremia exigua*; BoLy, *Boeremia lycopersici*; BoSp, *Boeremia* sp.; BuAl, *Bullera alba*; BuAu, *Buckleyzyma aurantiaca*; BuCr, *Bullera crocea*; CeFe, *Cenangium ferruginosum*; CerSp, Cercozoa sp.; CeSp, Ceratobasidiaceae sp.; ClDe, *Cladosporium delicatulum*; ClHe, *Cladosporium herbarum*; ClSp, *Cladosporium* sp.; CoSp, *Coniothyrium* sp.; CoTu, *Coleosporium tussilaginis*; CuMo, *Cutaneotrichosporon moniliiforme*; CyMa, *Cystofilobasidium macerans*; DeSp, *Dendryphion* sp.; DiBu, *Dioszegia butyracea*; DiLe, *Didymella lentis*; DioSp, *Dioszegia* sp.; DiPo, *Didymella pomorum*; DiSp, Didymellaceae sp.; CeSp, *Celosporium* sp.; EpNi, *Epicoccum nigrum*; ExEq, *Exophiala equina*; ExPi, *Exobasidium pieridis-ovalifoliae*; FiSp, *Filobasidium* sp.; FiSt, *Filobasidium stepposum*; FiWi, *Filobasidium wieringae*; GeAs, *Geomyces asperulatus*; GiTr, *Gibberella tricincta*; HaVe, *Harzia velata*; HeJu, *Herpotrichia juniperi*; HoMa, *Hormonema macrosporum*; HyaSp, Hyaloscyphaceae sp.; KoCh, *Kondoa changbaiensis*; LaCa, *Lachnellula calyciformis*; LeGu, *Lecanosticta guatemalensis*; LeSp, *Leucosporidium* sp.; LoCo, *Lophodermium conigenum*; MoSp, *Monographella* sp.; MyEl, *Mycosphaerella ellipsoidea*; MySp, Mycosphaerellaceae sp.; MyTa, *Mycosphaerella tassiana*; NeAb, *Neocatenulostroma abietis*; NeMi, *Neocatenulostroma microsporum*; NeSp, *Nectria* sp.; PaCa, *Passalora californica*; PaLa, *Papiliotrema laurentii*; PeAe, *Penicillium aethiopicum*; PeBi, *Penicillium bialowiezense*; PeBr, *Penicillium brevicompactum*; PeEx, *Penicillium expansum*; PenSp, *Penicillium* sp.; PeySp, *Peyronellaea* sp.; PhBu, Phoma bulgarica; PhLa, Phacidium lacerum; PlCu, Plectosphaerella cucumerina; PlOr, Plectosphaerella oratosquillae; PlSp, Pleosporales sp.; PsSp, *Pseudogymnoascus* sp.; RaHy, *Ramularia hydrangeae-macrophyllae*; RhiSp, *Rhizosphaera* sp.; RhMu, *Rhodotorula mucilaginosa*; ScSp, *Scleroconidioma sphagnicola*; SpRo, *Sporobolomyces roseus*; SpSp, *Sporobolomyces* sp.; SuGr, *Suillus granulatus*; ZyVe, *Zymoseptoria verkleyi*; TreSp, Tremellales sp.; TrSp, *Truncatella spadicea*; ViDi, *Vishniacozyma dimennae*; ViTe, *Vishniacozyma tephrensis*; ViVi, *Vishniacozyma victoriae*

Table 4 Identification of Stramenopila in potato samples based on the ITS1Oo + ITS4ngsUni primers. Numbers and percentages in Sequel and MinION columns indicate the number of all reads and per cent of sequences assigned to particular OTUs. Asterisks indicate taxon names that correspond to each other based on >98% sequence similarity. Samples with no PCR product and no sequences are excluded. Notable, plant and fungal sequences contributed on average 10% to MinION data (probably index switch artefacts from the fungal data set; not shown).

Sample	Sequel	MinION
KL003	262: *Phytophthora andina* 90%, *Peronospora radii* 10%	63: *Phytophthora infestans* 84%, *Peronospora radii* 11%
KL004	155: Xanthophyceae sp. 100%	92: Stramenopila sp. 86%
KL005	73: *Peronospora agrestis* 73%, Xanthophyceae 27%	29: *Peronospora agrestis* 83%, Stramenopila sp. 17%

KL006	143: *Peronospora* sp. 85%, *Eustigmatos* sp 15%	77: *Peronospora* sp. 79%, Eustigmataceae sp 13%
KL007	102: Chromulinaceae sp. 84%, *Hyaloperonospora parasitica* 16%	43: Chromulinaceae sp. 66%, *Hyaloperonospora parasitica* 9%
KL008	13: *Peronospora violae* 100%	6: *Peronospora violae* 84%
KL010	71: Chromulinaceae sp 100%	69: Chromulinaceae sp. 93%,
KL013	23: *Hyaloperonospora parasitica* 100%	7: *Hyaloperonospora parasitica* 100%
KL014	13: Xanthophyceae sp. 100%	2: Stramenopila sp. 50%
KL021	43: *Peronospora variabilis* 100%	19: -
KL022	27: *Peronospora variabilis* 100%	0: -
KL024	124: *Peronospora variabilis* 100%	0: -

Table 5 Details of needle samples.

Sample ID	Collection locality	Collection date	Host	Substrate	Disease symptoms
115	Tallinn Botanic Garden	17.11.2011	*Pinus sylvestris*	Needle	*Dothistroma*-like
117	Tallinn Botanic Garden	17.11.2011	*P. sylvestris*	Needle	*Dothistroma*-like
118	Pirita	17.11.2011	*P. mugo*	Needle	*Dothistroma*-like
119	Tallinn Botanic Garden	17.11.2011	*P. sylvestris*	Needle	*Dothistroma*-like
123	Tallinn Botanic Garden	17.11.2011	*P. uncinata*	Needle	*Dothistroma*-like
125	Tallinn Botanic Garden	17.11.2011	*P. rigida*	Needle	*Dothistroma*-like
127	Tallinn Botanic Garden	17.11.2011	*P. contorta*	Needle	*Dothistroma*-like
139	Tallinn Botanic Garden	15.08.2011	*P. x rotundata*	Needle	*Dothistroma*-like
141	Tallinn Botanic Garden	15.08.2011	*P. mugo*	Needle	*Dothistroma*-like
142	Tallinn Botanic Garden	15.08.2011	*P. x rotundata*	Needle	*Dothistroma*-like
148	Tallinn Botanic Garden	15.09.2011	*P. mugo* var. *pumilio*	Needle	*Dothistroma*-like
154	Tallinn Botanic Garden	15.09.2011	*P. rhaetica*	Needle	*Dothistroma*-like
2404	Mykolaiv, The Ukraine	10.09.2013	*P. nigra* subsp. *pallasiana*	Living culture: DoPi	NA[1]
3904	Kärevere	20.01.2015	*P. mugo*	Living culture: LeAc	NA
3906	Kärevere	20.01.2015	*P. mugo*	Living culture: DoSe	NA
4154	NA	09.10.2014	NA	Mock: DoPi, DoSe, LeAc	NA
4162	Levala	09.10.2014	*P. sylvestris*	Needle	*Dothistroma*-like
4180	Kolli	13.10.2014	*P. sylvestris*	Needle	*Dothistroma*-like
4181	Mustumetsa	13.10.2014	*P. sylvestris*	Needle	*Dothistroma*-like
4192	Soohara	07.10.2014	*P. sylvestris*	Needle	*Dothistroma*-like
4194	Värska	07.10.2014	*P. sylvestris*	Needle	*Dothistroma*-like
4195	Vastse-Kuuste	07.10.2014	*P. mugo*	Needle	*Dothistroma*-like
4197	Partsi	07.10.2014	*P. sylvestris*	Needle	*Dothistroma*-like
4220	Sääre	15.10.2014	*P. sylvestris*	Needle	*Dothistroma*-like
4221	Unimäe	15.10.2014	*P. sylvestris*	Needle	*Dothistroma*-like
4222	Tori	16.10.2014	*P. mugo*	Needle	*Dothistroma*-like
4223	Tori	16.10.2014	*P. mugo*	Needle	*Dothistroma*-like
5136	Imported: Netherlands	03.11.2015	*P. mugo* var. *pumilio*	Needle	Asymptomatic
5137	Imported: Germany	17.12.2015	*Picea omorika*	Needle	Asymptomatic
5146	Imported: Netherlands	03.11.2015	*P. mugo*	Needle	Asymptomatic
5148	Imported: Netherlands	03.11.2015	*P. mugo*	Needle	Asymptomatic
5151	Imported: Netherlands	03.11.2015	*P. sylvestris*	Needle	Asymptomatic
5186	Imported: Netherlands	26.10.2015	*P. peuce*	Needle	Asymptomatic
5194	Imported: Netherlands	26.10.2015	*P. koraiensis*	Needle	Asymptomatic
5195	Imported: Netherlands	26.10.2015	*P. mugo*	Needle	Asymptomatic
5297	Imported: Germany	17.12.2015	*Picea pungens*	Needle	Asymptomatic
5307	Imported: Germany	17.12.2015	*Picea omorika*	Needle	Asymptomatic
14374	Agali	16.02.2018	*P. sylvestris*	Needle	*Lecanosticta*-like
14378	Agali	16.02.2018	*P. mugo*	Needle	*Lecanosticta*-like

[1]NA, not applicable

Table 6 Details of potato samples.

Sample ID	Collection locality	Collection date	Potato cultivar	Substrate	Disease symptoms
KL001	Õssu	02.08.2017	Ants	Leaf	Dark circular lesions
KL002	Õssu	02.08.2017	Ants	Leaf	Dark circular lesions
KL003	Õssu	02.08.2017	Ants	Leaf	Dark circular lesions
KL004	Õssu	02.08.2017	Ants	Leaf	Dark circular lesions
KL005	Õssu	02.08.2017	Sarpo Mira	Leaf	Dark circular lesions
KL006	Õssu	02.08.2017	Sarpo Mira	Leaf	Dark circular lesions
KL007	Õssu	02.08.2017	Sarpo Mira	Leaf	Dark circular lesions
KL008	Õssu	02.08.2017	Sarpo Mira	Leaf	Dark circular lesions
KL009	Õssu	02.08.2017	Toluca	Leaf	Dark circular lesions
KL010	Õssu	02.08.2017	**Toluca**	Leaf	Dark circular lesions
KL011	Õssu	02.08.2017	**Toluca**	Leaf	Dark circular lesions
KL012	Õssu	02.08.2017	**Toluca**	Leaf	Dark circular lesions
KL013	Õssu	02.08.2017	Makhai	Leaf	Dark circular lesions
KL014	Õssu	02.08.2017	Makhai	Leaf	Dark circular lesions
KL015	Õssu	02.08.2017	Makhai	Leaf	Dark circular lesions
KL016	Õssu	02.08.2017	Makhai	Leaf	Dark circular lesions
KL017	Õssu	02.08.2017	Kelly	Leaf	Dark circular lesions
KL018	Õssu	02.08.2017	Kelly	Leaf	Dark circular lesions
KL019	Õssu	02.08.2017	Kelly	Leaf	Dark circular lesions
KL020	Õssu	02.08.2017	Kelly	Leaf	Dark circular lesions
KL021	Karala	02.08.2017	unknown	Leaf	Dark circular lesions
KL022	Karala	02.08.2017	unknown	Leaf	Dark circular lesions
KL023	Karala	02.08.2017	unknown	Leaf	Dark circular lesions
KL024	Karala	02.08.2017	unknown	Leaf	Dark circular lesions
KL025	Metsaküla	12.08.2017	unknown	Leaf	Dark circular lesions
KL026	Metsaküla	12.08.2017	unknown	Leaf	Dark circular lesions
KL027	Metsaküla	12.08.2017	unknown	Leaf	Dark circular lesions
KL028	Väljataguse	11.04.2018	Elfe	Tuber	Potato gangrene
KL029	Väljataguse	11.04.2018	Elfe	Tuber	Potato gangrene
KL030	Väljataguse	11.04.2018	Elfe	Tuber	Potato gangrene
KL031	Õssu	11.04.2018	Laura	Tuber	Potato gangrene
KL032	Suur-Rahula	11.04.2018	Gala	Tuber	Potato gangrene
KL033	Tagaküla	11.04.2018	Laura	Tuber	Potato gangrene
KL034	Tagaküla	11.04.2018	Laura	Tuber	Potato gangrene
KL035	Padise	11.04.2018	Marabel	Tuber	Potato gangrene
KL036[1]	Õssu	30.08.2018	Carolus	Tuber	*Rhizoctonia*-like
KL037[2]	Õssu	30.08.2018	Carolus	Tuber	*Rhizoctonia*-like

[1]used only for the ONT2g and ONT2h runs;

[2]used only for the ONT2i run