Journal of Microbiological Methods 178 (2020) 106054

Contents lists available at ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

A simple method for high molecular-weight genomic DNA extraction suitable for long-read sequencing from spores of an obligate biotroph oomycete

Charlotte Penouilh-Suzette[a], Sandra Fourré[b], Guillaume Besnard[c], Laurence Godiard[a], Yann Pecrix[a,d,*]

[a] LIPM (Laboratoire des Interactions Plantes Microorganismes), INRAE, CNRS, Université de Toulouse, 24 Chemin de Borde-Rouge, BP 52627, F-31326 Castanet-Tolosan, France
[b] GeT-PlaGe, INRAE Auzeville, US 1426, 24 Chemin de Borde-Rouge, BP 52627, F-31326 Castanet-Tolosan, France
[c] CNRS, Université Paul Sabatier, IRD, UMR 5174 EDB (Laboratoire Évolution et Diversité Biologique), 118 route de Narbonne, F-31062 Toulouse, France
[d] CIRAD, UMR 53 Peuplements Végétaux et Bioagresseurs en Milieu Tropical (PVBMT), Pole de Protection des Plantes, 7 chemin de l'IRAT, F-97410 Saint Pierre, Réunion, France

ARTICLE INFO

Keywords:
Oomycete
Downy mildew
Spores
High molecular weight genomic DNA
PacBio
Long-read sequencing

ABSTRACT

Long-read sequencing technologies are having a major impact on our approaches to studying non-model organisms and microbial communities. By significantly reducing the cost and facilitating the genome assembly pipelines, any laboratory can now develop its own genomics program regardless of the complexity of the genome studied. The most crucial current challenge is to develop efficient protocols for extracting genomic DNA (gDNA) with high quality and integrity adapted to the organism of interest. This can be particularly complex for obligate pathogens that must maintain intimate interactions inside infected host tissues. Here we propose a simple and cost-effective method for high molecular weight gDNA extraction from spores of *Plasmopara halstedii*, an obligate biotroph oomycete pathogen responsible for downy mildew in sunflower. We optimized the yield, the quality and the integrity of the extracted gDNA by fine-tuning three critical parameters, the grinding, the lysis temperature and the lysis duration. We obtained gDNA with a fragment size distribution reaching a peak ranging from 79 to 145 kb. More than half of the extracted gDNA consisted of DNA fragments larger than 42 kb, with 23% of fragments larger than 100 kb. We then demonstrated the relevance of this protocol for long-read sequencing using PacBio RSII technology. With this protocol, we were able to obtain a mean read length of 9.3 kb, a max read length of 71 kb and an N50 of 13.3 kb. The development of such DNA extraction protocols is an essential prerequisite for fully exploiting technologies requiring high molecular weight gDNA (*e.g.* long-read sequencing or optical mapping). These technological advances will help generate data to answer questions such as the role of newly duplicated gene clusters, repeated regions, genomic structural variations or to define number of chromosomes that still remains undefined in many species of pathogenic fungi and oomycetes.

1. Introduction

The biological sciences have undergone a revolution thanks firstly to Sanger sequencing and then to high-throughput second-generation sequencing also called Next Generation Sequencing (NGS) which has enabled to decipher genomic DNA (gDNA) from any organism to be deciphered and consequently to broaden the use of genome-wide studies. Until now, the most widespread NGS technologies were based on short read sequencing, typically of 50–300 nucleotides. These technologies required the development of complex bioinformatic pipelines to obtain the assembly of millions of short sequences into genomes. However, the limit of NGS had been reached since no assembler program can solve the problem of repeated elements. Therefore, short read sequencing can only give a fragmented genome view since whole genomic regions can be masked and structural variations, such as indels or complex chromosomal rearrangements, are highly underestimated

* Corresponding author at: CIRAD, UMR 53 Peuplements Végétaux et Bioagresseurs en Milieu Tropical (PVBMT), Pole de Protection des Plantes, 7 chemin de l'IRAT, F-97410 Saint Pierre, Réunion, France
E-mail addresses: charlotte.penouilh@gmail.com (C. Penouilh-Suzette), sandra.fourre@inrae.fr (S. Fourré), guillaume.besnard@univ-tlse3.fr (G. Besnard), laurence.godiard@inrae.fr (L. Godiard), yann.pecrix@cirad.fr (Y. Pecrix).

https://doi.org/10.1016/j.mimet.2020.106054
Received 3 May 2020; Received in revised form 9 August 2020; Accepted 7 September 2020
Available online 12 September 2020

(Sedlazeck et al., 2018).

These last years have witnessed a technological breakthrough with the development of a third-generation sequencers developed by Pacific Biosciences (Roberts et al., 2013) or Oxford Nanopore Technologies (Jain et al., 2016) using single-molecule real time sequencing (SMRT) or nanopore-based sequencing technologies, respectively. These new technologies have significantly increased the length of sequencing by providing average reads of 10 kb. These advances in long-read single molecule sequencing considerably facilitated the unambiguous assembly of complete genomes allowing exploration of new regions that were inaccessible thus far, revealing their true complexity. However, against all odds, gDNA extraction has become today the most limiting step for genome sequencing since to fully exploit the potential of long-read sequencing, it is critical to obtain high molecular weight gDNA of sufficient quality and quantity.

The extraction of genomic DNA respecting the quality controls required for long-read sequencing can be particularly challenging for certain organisms such as some pathogens that are completely reliant on host tissues. This is especially the case for fungi (*e.g.* Feehan et al., 2017; Lee et al., 2017) and oomycetes, a class of fungal like organisms that include a large number of pathogens responsible for e*conomically important plant and animal diseases* (Kamoun et al., 2015; Derevnina et al., 2016). *Numerous oomycetes species are obligate parasites and must maintain a close interaction with their host* (Fawke et al., 2015). *Most of the time, this lifestyle prevents in vitro culture. Consequently, the harvest of samples* such as mycelium that are devoid of host tissues and easy to handle for gDNA extraction *is complicated. Therefore to obtain gDNA without host contaminant, the choice of tissues is limited to dissemination structures like sporangia.* This biological constraint is particularly true for *Plasmopara halstedii*, a devastating obligate biotroph oomycete responsible for sunflower downy mildew (Gascuel et al., 2015). *In recent decades, the sudden genetic diversification of P. halstedii has led to the emergence of new isolates that are particularly virulent, constituting a serious threat to the sunflower cultivation* (Ahmed et al., 2012). *Recently, the identification of P. halstedii RXLR effectors and the use of some of them accurately selected in breeding programs demonstrated that they could be a powerful tool to identify new genetic plant resistances* (Gascuel et al., 2016; Pecrix et al., 2018a; Pecrix et al., 2019). *Thus, obtaining a complete version of the P. halstedii genome is a major challenge for sunflower cultivation, since it will provide (i) an exhaustive view of the effector repertoire of the pathogen, and (ii) a better understanding of* mechanisms of genomic change at the point of origin of new virulences.

In this work we present a simple method to extract high molecular weight gDNA from *P. halstedii* sporangia and spores. We have investigated three parameters, the grinding duration, lysis temperature and duration, affecting DNA yield, quality and integrity. We finally provided a rapid optimized protocol that was more efficient than an often-used commercial kit and showed that this gDNA extraction method was suitable for long-read sequencing.

2. Materials and methods

2.1. Culture and sporangia harvesting

Spores of *P. halstedii* were collected from the susceptible sunflower variety *Helianthus annuus* L., cv. Peredovik. The sunflower seeds were germinated for two days at 25 °C in the dark in Petri dishes between two discs of Whatman paper moistened with distilled water. The seedlings were soaked in a *P. halstedii* spore solution (10^5 spores/ml) for 4 h at 18 °C. The infected seedlings were then sown in soilless compost and grown in uncovered germination trays (52 × 42 × 9 cm, Stewart Plastics Ltd., Croydon, UK) at 18 °C, 16/8 h day/night photoperiod and 80% relative humidity, allowing the colonization of the plantlets by *P. halstedii*. Germination trays were then covered with transparent lids for 48 h at the cotyledon stage, before the first pair of leaves appear, to induce sporulation.

Sporulating cotyledons were detached from the plants, placed in Petri dishes and immersed in distilled water. The spores were carefully harvested by rubbing the cotyledons with a brush. The suspension was pipetted into 2 ml round-bottomed centrifuge tubes and the spores were pelleted by centrifugation at 1000*g* for 1 min. The supernatant was discarded, and the pellet was air-dried 10 min at room temperature. The samples were then frozen using liquid nitrogen and stored at −80 °C.

2.2. Genomic DNA extraction

Two stainless-steel beads (3 mm diameter) were added to each microcentrifuge tube containing the pellet of spores. The samples were then placed in liquid nitrogen and subsequently ground using a Mixer Mill MM 400 grinder (Retsch, Haan, Germany) with tubes adapters racks cooled to −80 °C beforehand. Different grinding durations of 10, 30 or 90 s with a vibrational frequency set to 30 Hz (1800 rpm) were tested.

In the following protocol, we excluded the use of vortex mixer and pipetting steps were performed slowly with caution to limit DNA shearing. The gDNA was extracted from up to 50 mg of ground spores obtained from four to eight infected cotyledons depending on the intensity of sporulation. The lysis was performed by adding 800 µl of preheated lysis buffer (100 mM Tris-HCl pH 8, 10 mM EDTA pH 8, 1 M NaCl, 1% SDS) to samples supplemented with 2 µl of proteinase K (Roche, 20 mg/ml). The mix was incubated at 50 °C, 60 °C or 70 °C for 30 min, 180 min or 16 h (different temperatures and durations tested in this work). During the first 30 min of the lysis, the sample was mixed manually by inverting the tube twice every 5 min. To precipitate proteins and polysaccharides, 270 µl of potassium acetate 5 M were added to the lysate, the resulting mix was incubated on ice for 10 min and centrifuged at 5000 *g* for 10 min at 4 °C. The supernatant was transferred into a new 2 ml tube, and 2 µl of RNAse A (Qiagen, 100 mg/ml) was added to degrade the ribonucleic acids. The samples were mixed by inversion and incubated 30 min at 37 °C. The gDNA was precipitated by adding one volume of isopropanol (about 1 ml), inverting the tube carefully 10 times and letting incubate 30 min at room temperature. The gDNA was pelleted by centrifugation at 10,000 *g* for 2 min. The supernatant was discarded and the pellet of gDNA was washed with 1 ml of ethanol 70% freshly prepared. The gDNA was pelleted by centrifugation at 10,000 *g* for 1 min and the supernatant was discarded. The pellet of gDNA was air dried for 1 min and resuspended in 50 µl of TE buffer (10 mM Tris pH 8, 1 mM EDTA pH 8).

2.3. Genomic DNA quantification and quality control

The gDNA was quantified with a Qubit 3.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) after staining with Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA). The purity of gDNA was estimated with a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) by calculating A260/A280 and A260/A230 ratios. The gDNA fragment sizes were first investigated with a Fragment Analyzer (Advanced Analytical Technologies Inc., USA) using the HS Large Fragment 50 kb Kit, DNF-464 (Advanced Analytical Technologies Inc., USA). In the case of gDNA samples composed of very large fragments the sizes were measured with a Femto Pulse System (Advanced Analytical Technologies Inc., USA) using the Genomic DNA 165 kb Kit, FP-1002-0275 (Advanced Analytical Technologies Inc., USA). The data were processed using the PROSize 3.0.1.5 software (Advanced Analytical Technologies Inc., USA).

2.4. PacBio sequencing

The sequencing was performed by the GeT-PlaGe core facility (INRAE, Toulouse, France; https://get.genotoul.fr). A library was

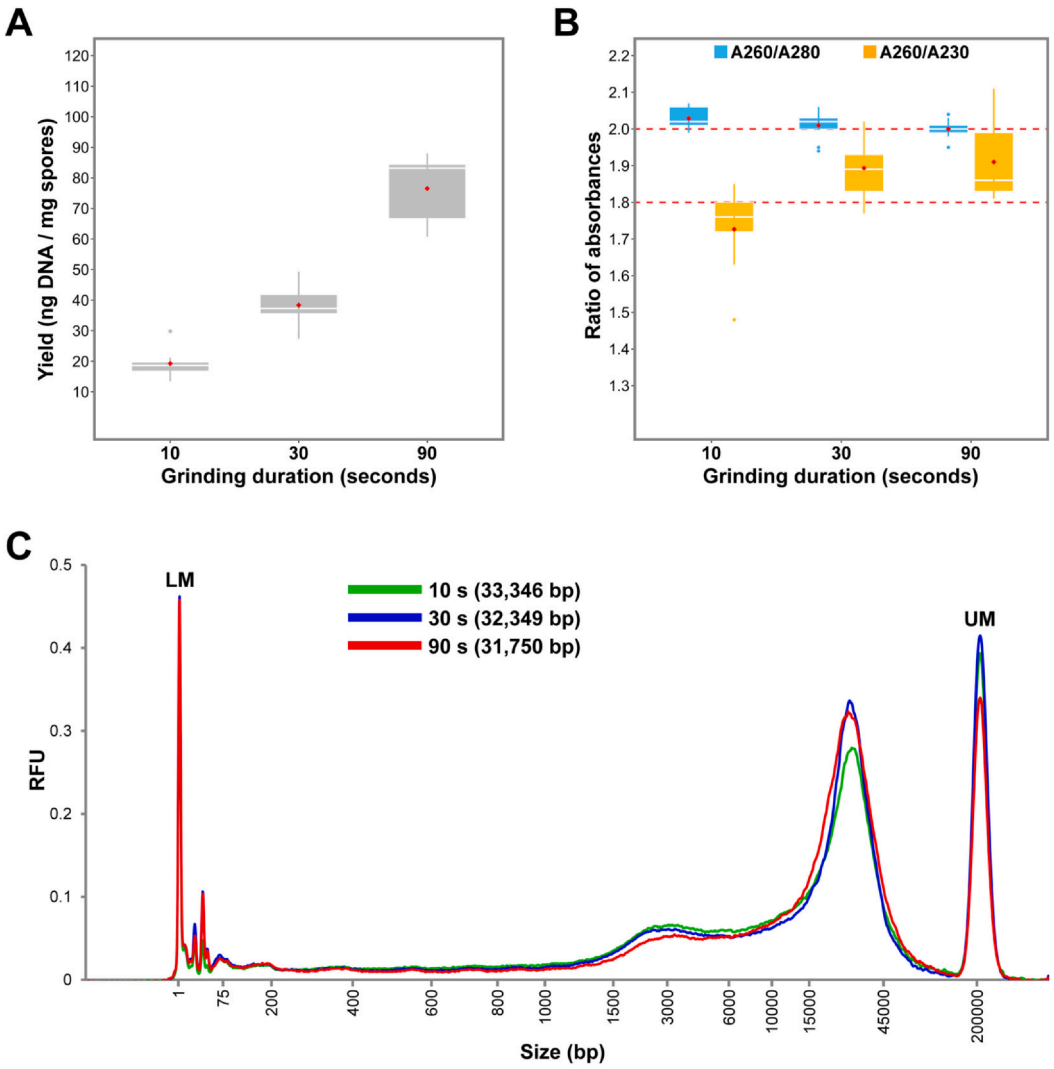

Fig. 1. Grinding duration effect on the yield, quality and integrity of the extracted genomic DNA.
(A) DNA Yield. (B) DNA quality. Blue and yellow boxplots represent A260/A280 and A260/A230 ratios of absorbance, respectively. The red dotted lines indicate the quality thresholds required by Pacific Bioscience for long-read sequencing. The 260/280 ratio must range between 1.8 and 2.0 and the 260/230 ratio should be close to 2.0. (A) and (B) For each condition, the white lines in the boxplots represent the median and the red dots indicate the mean. Each measurement was obtained from nine replicates. (C) DNA Integrity. Size distribution of extracted genomic DNA measured with a Fragment Analyzer. For each measurement, equal amounts of gDNA sample from nine replicates were mixed for analysis. LM = Lower Marker, UM = Upper Marker. For each condition, the size of the peak is indicated in brackets. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

produced with the SMRTbell Template Prep Kit 1.0, from 1 µg of gDNA and sequenced by a PacBio RS II SMRT DNA Sequencing System (Pacific Biosciences, Menlo Park, CA, USA) with nine SMRT Cells using the P6/C4 chemistry.

3. Results

To optimize the gDNA extraction protocol we attempted to define the best conditions of grinding and lysis. The objective was to obtain sufficient quantities of gDNA and to maximize the quality and the sizes of gDNA fragments in compliance with the PacBio manufacturer's recommendations, *i.e.* (i) an A260/A280 ratio greater than 1.8, (ii) an A260/A230 ratio close to 2, and (iii) gDNA mean fragment size in the range of 10 kb to 100 kb and ideally > 40 kb.

3.1. Effect of grinding duration on gDNA extraction

The grinding of the samples was carried out using a mechanical grinder at 30 rpm (30 Hz) for 10, 30 and 90 s. Whatever the grinding duration, the lysis was performed at 50 °C for 30 min (Fig. 1). Unsurprisingly, grinding was essential for an efficient gDNA extraction since 10 s of grinding led to an extract with insufficient amounts of gDNA with an average of 19.3 ng/mg (nanogram of gDNA obtained per milligram of spores), while the yield increased from 38.4 ng/mg after 30 s to 76.5 ng/mg after 90 s of grinding (Fig. 1A). The A260/A280 ratio measurements for all extractions were over 1.8 and were thus not indicative of presence of proteins or other contaminants absorbing at or near 280 nm (Fig. 1B). By contrast, the A260/A230 ratio measurements indicated that grinding had a significant impact on gDNA purity since a sufficient quality was only reached from 30 s of grinding. This qualitative gain was certainly the consequence of the yield improvement which decreased the proportion of contaminants, such as carbohydrates, absorbing at or near A230. Most importantly, the grinding did not dramatically damage the gDNA. The size distribution of gDNA fragments, assessed by Fragment Analyzer, displayed a peak between 31.8 and 33.3 kb (Fig. 1C). Moreover, the proportions of gDNA fragments larger than 10 kb, estimated by the area under the curve calculation, were 56%, 59% and 65% for 10, 30 and 90 s of grinding

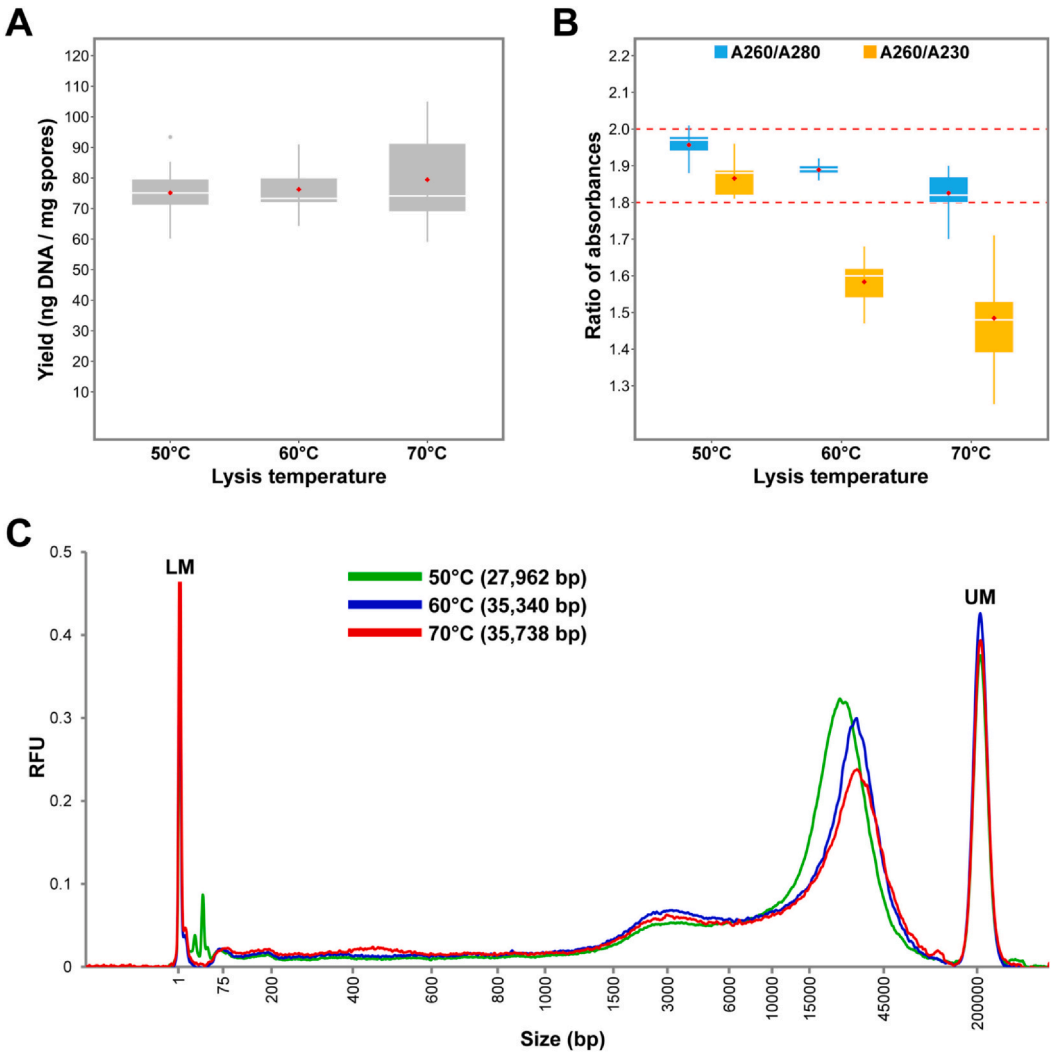

Fig. 2. Lysis temperature effect on the yield, quality and integrity of the extracted genomic DNA.
(A) DNA Yield. (B) DNA quality. Blue and yellow boxplots represent A260/A280 and A260/A230 ratios of absorbance, respectively. The red dotted lines indicate the quality thresholds required by Pacific Bioscience for long-read sequencing. (A) and (B) For each condition, the white lines in the boxplots represent the median and the red dots indicate the mean. Each measurement was obtained from nine replicates. (C) DNA Integrity. Size distribution of extracted genomic DNA measured with a Fragment Analyzer. For each measurement, equal amounts of gDNA sample from nine replicates were mixed for analysis. LM = Lower Marker, UM = Upper Marker. For each condition, the size of the peak is indicated in brackets. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

duration, respectively. Based on these results, all the following grindings were carried out at 30 rpm for 90 s.

3.2. Effect of lysis temperature on gDNA extraction

To estimate the impact of the lysis temperature, three conditions were tested, 50 °C, 60 °C and 70 °C. To amplify a possible effect of the temperature, the lysis duration was increased and set to 90 min (Fig. 2). Increasing the lysis temperature did not improve the gDNA extraction with yields reaching an average of 75.1, 76.3 and 79.4 ng/mg at 50 °C, 60 °C and 70 °C, respectively (Fig. 2A). On the other hand, increasing temperature affected gDNA purity. The A260/A280 ratio was significantly reduced when comparing 50 °C and 70 °C but all met quality criteria (Fig. 2B). However, this decrease was more marked for the A260/A230 ratio with average values of 1.87, 1.58 and 1.48 at 50 °C, 60 °C and 70 °C, respectively. The increase in temperature affected the gDNA fragments size whose distribution peaks ranged from 28 kb at 50 °C to 35.7 kb at 70 °C. However, the increase in temperature was accompanied by an accentuated gDNA degradation since the proportion of larger fragments of 10 kb were 65%, 57% and 56% for 50 °C, 60 °C and 70 °C, respectively (Fig. 2C). Given the negative impact of temperature on the quality and integrity of the extracted gDNA, the following lyses were performed at 50 °C.

3.3. Effect of lysis duration on gDNA extraction

The lysis duration effect was estimated by comparing 30 min, 180 min and overnight lysis with a temperature set at 50 °C (Fig. 3). The lengthening of lysis had no impact on the yields which were on average 74.5, 78.4 and 79.6 ng/mg respectively for 30, 180 min and overnight (Fig. 3A). On the other hand, the impact on gDNA purity was severe at 180 min and overnight, especially for the A260/A280 ratio, which suggested that prolonged lysis could solubilize contaminants like proteins (Fig. 3B). In addition, the prolongation of the lysis could dramatically affect the integrity of the gDNA. While the distribution size of gDNA fragments, extracted with a 30 min lysis reached a peak at 30.7 kb, those obtained with 180 min and overnight lysis decreased respectively from 22.5 to 17.6 kb (Fig. 3C). Accordingly, the lysis time for the subsequent long-read sequencing, will be of 30 min.

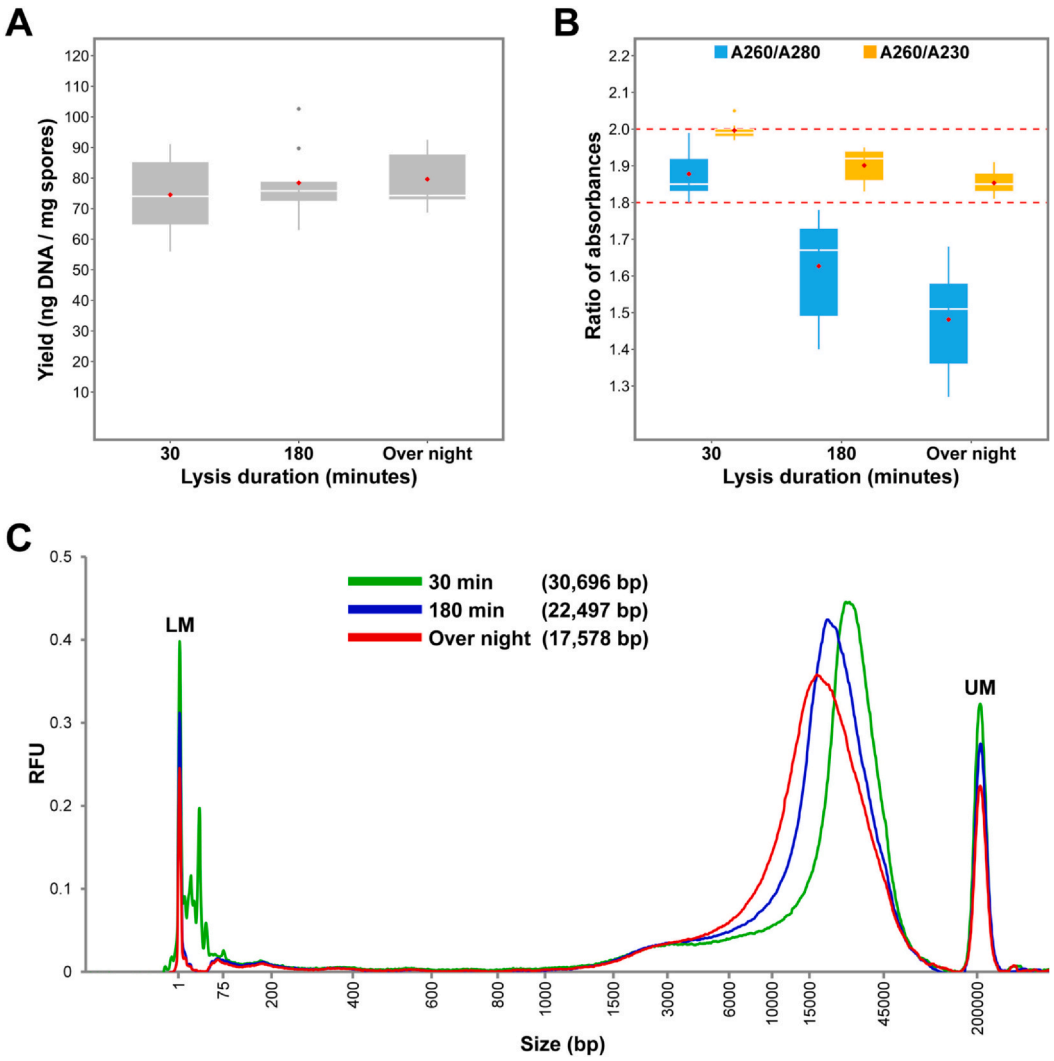

Fig. 3. Lysis duration effect on the yield, quality and integrity of the extracted genomic DNA.
(A) DNA Yield. (B) DNA quality. Blue and yellow boxplots represent A260/A280 and A260/A230 ratios of absorbance, respectively. The red dotted lines indicate the quality thresholds required by Pacific Bioscience for long-read sequencing. (A) and (B) For each condition, the white lines in the boxplots represent the median and the red dots indicate the mean. Each measurement was obtained from nine replicates. (C) DNA Integrity. Size distribution of extracted genomic DNA measured with a Fragment Analyzer. For each measurement, equal amounts of gDNA sample from nine replicates were mixed for analysis. LM = Lower Marker, UM = Upper Marker. For each condition, the size of the peak is indicated in brackets. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.4. Optimized protocol and long-read sequencing

Taking all the previous results into account summarized in Table 1, we defined an optimized protocol by fixing the grinding duration at 90 s and the lysis temperature at 50 °C. To minimize the negative impact of lysis on the size of the gDNA fragments, we reduced the lysis time to 10 min. A condensed version for bench of this optimized protocol is available in Fig. S1.

With these settings we obtained average yield of 85.3 ± 3.1 ng/mg, and average A260/A280 and A260/A230 ratios of 2.02 ± 0.01 and 1.95 ± 0.03, respectively. These results were more satisfying than those obtained with the DNeasy Plant Mini Kit (supplied by QIAGEN), the most commonly method used for gDNA extraction from oomycetes (Zelaya-Molina et al., 2011), with which we obtained average yield of

Table 1
Summary of protocol optimizations.

	Grinding duration optimization			Lysis temperature optimization			Lysis duration optimization		
Grinding duration (seconds)	10	30	90	90	90	90	90	90	90
Lysis temperature	50 °C	50 °C	50 °C	50 °C	60 °C	70 °C	50 °C	50 °C	50 °C
Lysis duration (minutes)	30	30	30	90	90	90	30	180	Over night
Yield (ng DNA/mg spores)	19.3	38.4	76.5	75.1	76.3	79.4	74.5	78.4	79.6
A260/A280	2.0	2.0	2.0	2.0	1.9	1.8	2.0	1.9	1.9
A260/A230	1.7	1.9	1.9	1.9	1.6	1.5	1.9	1.6	1.5
DNA fragments > 10 kb	55.5%	59.2%	64.5%	64.5%	57.5%	55.7%	64.6%	62.9%	59.0%

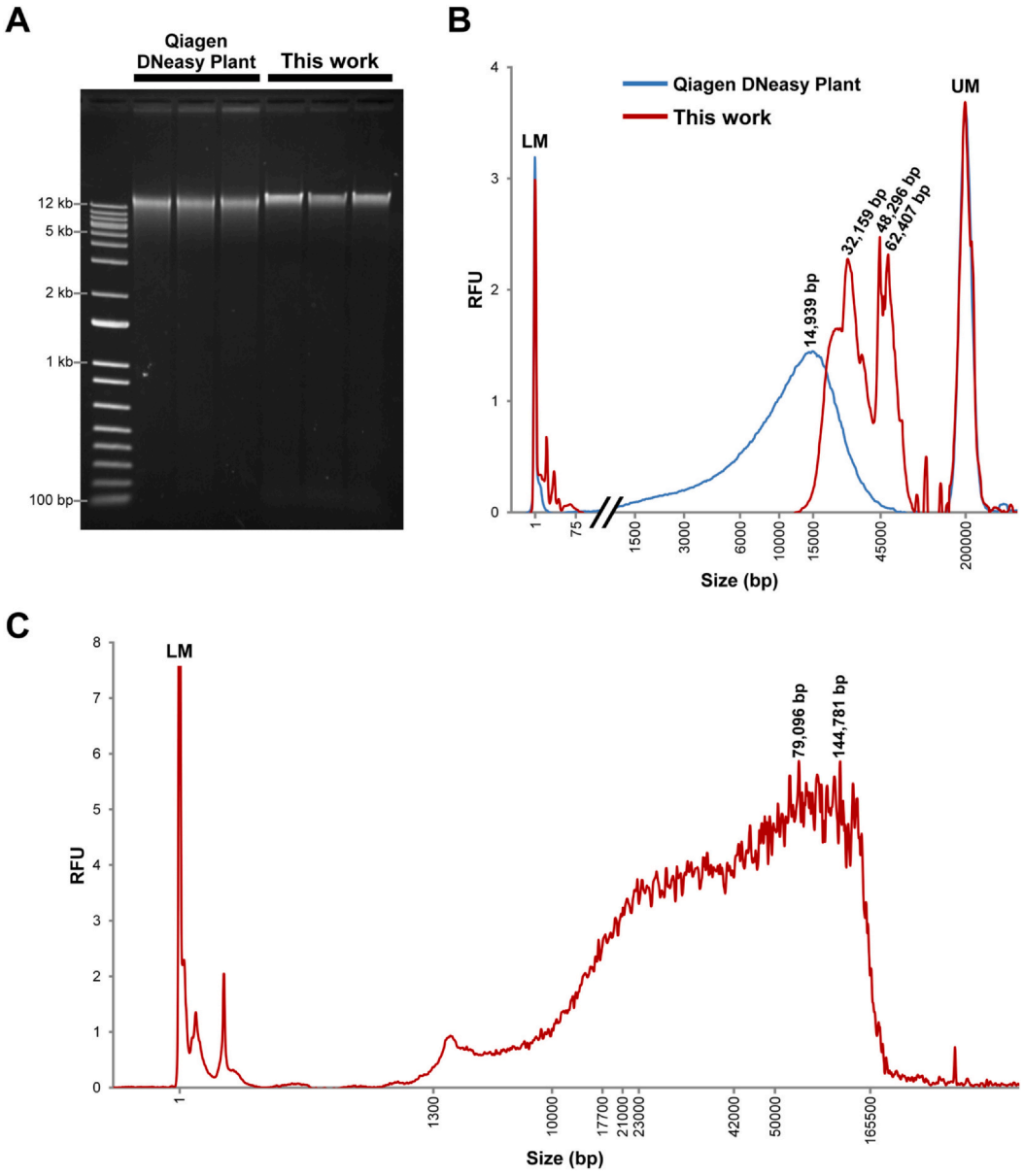

Fig. 4. Optimized genomic DNA extraction protocol.
(A) Agarose gel electrophoresis of gDNA samples obtained with the DNeasy Plant Mini Kit (Qiagen) or with our protocol. For both methods, three replicates are presented. (B) Size distribution of extracted genomic DNA measured with a Fragment Analyzer. The multi-peak profile is characteristic of a sample with DNA fragments sizes out of range. (C) Size distribution of extracted genomic DNA obtained with our protocol and measured with a Femto Pulse System, a more suitable device for very high molecular weight DNAs. For each measurement, equal amounts of gDNA sample from three replicates were mixed for analysis. LM = Lower Marker, UM = Upper Marker.

31.8 ± 4.0 ng/mg, and average A260/A280 and A260/A230 ratios of 1.82 ± 0.01 and 1.27 ± 0.02, respectively. Agarose gel migration showed that the gDNA extracted with our protocol appeared to be less degraded (Fig. 4A). To accurately compare these two methods, we investigated the gDNA integrity by measuring the size distribution of extracted gDNA with a Fragment Analyzer (Fig. 4B). For samples extract with DNeasy Plant Mini Kit, we obtained an average gDNA fragment distribution reaching a peak at 14.9 kb and 50% of the extracted gDNA was composed of fragments larger than 11.4 kb. In contrast, with our protocol we obtained a multi-peak profile which is characteristic of a sample with gDNA fragment sizes out of range for this apparatus (Fig. 4B). We then measured the gDNA size using the Femto Pulse System, an instrument suitable for very high molecular weight DNA fragments through 165 kb. The gDNA displayed a fragment size distribution reaching a peak ranging from 79 to 145 kb and fragments gDNA size over 10 kb, 42 kb and 100 kb represented respectively 93%, 50% and 23% of total extracted gDNA (Fig. 4C). This gDNA quality is optimal for long-read sequencing by PacBio RS II. Although, according to manufacturer's instructions and sequencing facilities, libraries of insert sizes ranging from 500 bp to greater than 20 kb are required to obtain 10 kb raw reads on average, they highly recommend generating these libraries from larger fragments, ideally 40 kb on average.

Based on this protocol, we extracted 1 µg of gDNA from the *P. halstedii* pathotype 710 for sequencing by PacBio RS II technology. In total 1.1 million subreads were generated with an N50 of 13.3 kb and a mean length of 9.3 kb. The targeted genome coverage of 140× was obtained with 10.5 Gb of raw sequence (Table 2).

Table 2
Raw sequencing metrics obtained with PacBio RS II sequencer.
NUM: total number of reads, MAX: size of the longest read in bp; N50 BP: half of the data in reads with length greater than this value; N50 NUM: number of reads to get half of the data; N90 BP: 90% of the data in reads with length greater than this value; N90 NUM: number of reads to get 90% of the data; MEAN: mean read length, MEDIAN: median read length; BP: total number of base pairs sequenced; Coverage: ratio BP/size of *Plasmopara halstedii* genome estimated at 75 Mb (Pecrix et al., 2019).

Isolate	NUM	MAX	N50 BP	N50 NUM	N90 BP	N90 NUM	MEAN	MEDIAN	BP	Coverage
Plhal710	1,131,228	71,339	13,315	286,525	5390	739,606	9284	8236	10,503,316,161	140×

4. Discussion

In this study we have developed an efficient method for high molecular-weight gDNA extraction from spores of an obligate biotroph oomycete. To optimize the protocol we have varied three parameters, the grinding duration, the temperature and the duration of the lysis. After each extraction, the yield as well as the quality and integrity of the gDNA were estimated.

For the grinding of spores we preferred the use of a mechanical grinder instead of manual grinding with mortar and pestle. This automated device has ensured the reproducible modulation of grinding duration. For the quantifications of extracted gDNA, the Qubit Fluorometer was preferred to NanoDrop for its sensitivity and accuracy. The Qubit Fluorometer uses a fluorescent dye that is specific to double-stranded DNA whereas NanoDrop quantification is only based on the UV absorbance at 260 nm. Many components other than DNA absorb at 260 nm, often resulting in overestimation of the amount of DNA measured by NanoDrop. NanoDrop is however a suitable approach to detect the presence of contaminants that can absorb at a wavelength of 280 nm (protein, phenol) or 230 nm (carbohydrate, polysaccharide, phenol, guanidine, glycogen) and to estimate the purity of DNA by calculating A260/A280 and A260/A230 ratios.

The DNA sizing and estimation of its degradation is classically performed by conventional agarose gel electrophoresis using a static electric field. However, this method is routinely used to separate and size DNA fragments $\leq$20 kb and is not reliable for high molecular weight DNA (Ferris et al., 2004). We thus preferred the Fragment Analyzer to evaluate DNA integrity, an automated capillary electrophoresis instrument that enables a DNA sizing range of 75 bp to 48,500 bp or Femto Pulse System for larger DNA fragments up to 165 kb.

We then were able to show that the grinding step is crucial to improve yields and quality without affecting the gDNA integrity. To obtain pure gDNA samples composed of large fragments, we pointed out the importance of limiting the lysis temperature and lysis duration that might promote the solubilization of contaminants and the degradation of gDNA.

We finally confirmed that this protocol is suitable for third-generation long-range DNA sequencing by producing long-reads using the PacBio RS II technology (associated with the P6/C4 chemistry). The sequencing metrics were indicative of a high sequencing quality and were consistent with those obtained with the same sequencing technology and chemistry for other complex eukaryotic genomes (Shi et al., 2016; Badouin et al., 2017; Miller et al., 2018; Pecrix et al., 2018b; Fletcher et al., 2019).

5. Conclusions

The development of such protocols of large gDNA molecules extraction is an essential prerequisite for gapless genome assembly that has now become a major issue (Thomma et al., 2016). Although in recent years, hundreds of genomes sequenced using short-read technologies were released, these are generally highly fragmented, like the two *Plasmopara halstedii* genomes currently available (Sharma et al., 2015; Pecrix et al., 2019). This quality of assembly has so far been sufficient to analyze and compare gene contents but these technologies have now shown their limits when investigating genome structure. The first long-read sequencing of genomes including those of fungi or oomycetes have glimpsed higher levels of complexity in genome functions (Miller et al., 2018; Yang et al., 2018; Dussert et al., 2019; Shen et al., 2019). Thus, the roles of newly duplicated gene clusters, repeated regions or genomic structural variations, in the virulence of pathogenic fungi and oomycetes can now be fully unraveled.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mimet.2020.106054.

Authors' contributions

YP, LG and GB conceived the project and guided the research work. CPS and YP produced and harvested the spores of *Plasmopara halstedii*. CPS and YP extracted the genomic DNAs. YP and SF performed the quality control assays and analyzed the data. LG, GB critically revised and contributed to the manuscript. YP wrote the manuscript. All authors read and approved the final manuscript.

Declaration of Competing Interest

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

Acknowledgements

This work was supported by the French ANR grant EFFECTOORES (ANR-13-ADAP-0003) and by the Inter-Units grant EVOPLASMO funded by the French Laboratory of Excellence project TULIP (Labex TULIP, ANR-10-LABX-41; ANR-11-IDEX-0002-02) and the Agrobiosciences, Interactions and Biodiversity Research Federation (FR AIB, FR3450). We thank the Sunflower Biological Resources Centre of the French National Institute for Agricultural Research (INRAE) for providing sunflower seeds. We thank the Get-PlaGe platform for DNA measurements and sequencing, and the LIPM Bioinformatics platform for performing metrics calculations. GB is member of the EDB laboratory supported by the French Laboratory of Excellence project CEBA (Labex CEBA, ANR-10-LABX-25-01) and Labex TULIP, managed by the French ANR.

References

Ahmed, S., de Labrouhe, D.T., Delmotte, F., 2012. Emerging virulence arising from hybridisation facilitated by multiple introductions of the sunflower downy mildew pathogen *Plasmopara halstedii*. Fungal Genet. Biol. 49, 847–855. https://doi.org/10.1016/j.fgb.2012.06.012.

Badouin, H., Gouzy, J., Grassa, C.J., Murat, F., Staton, S.E., Cottret, L., Lelandais-Brière, C., Owens, G.L., Carrère, S., Mayjonade, B., Legrand, L., Gill, N., Kane, N.C., Bowers, J.E., Hubner, S., Bellec, A., Bérard, A., Bergès, H., Blanchet, N., Boniface, M.-C., Brunel, D., Catrice, O., Chaidir, N., Claudel, C., Donnadieu, C., Faraut, T., Fievet, G., Helmstetter, N., King, M., Knapp, S.J., Lai, Z., Le Paslier, M.-C., Lippi, Y., Lorenzon, L., Mandel, J.R., Marage, G., Marchand, G., Marquand, E., Bret-Mestries, E., Morien, E., Nambeesan, S., Nguyen, T., Pegot-Espagnet, P., Pouilly, N., Raftis, F., Sallet, E., Schiex, T., Thomas, J., Vandecasteele, C., Varès, D., Vear, F., Vautrin, S., Crespi, M., Mangin, B., Burke, J.M., Salse, J., Muños, S., Vincourt, P., Rieseberg, L.H., Langlade, N.B., 2017. The sunflower genome provides insights into oil metabolism, flowering

and Asterid evolution. Nature 546, 148–152. https://doi.org/10.1038/nature22380.
Derevnina, L., Petre, B., Kellner, R., Dagdas, Y.F., Sarowar, M.N., Giannakopoulou, A., De La Concepcion, J.C., Chaparro-Garcia, A., Pennington, H.G., Van West, P., Kamoun, S., 2016. Emerging oomycete threats to plants and animals. Philos. Trans. R. Soc. B Biol. Sci. 371, 20150459. https://doi.org/10.1098/rstb.2015.0459.
Dussert, Y., Mazet, I.D., Couture, C., Gouzy, J., Piron, M.-C., Kuchly, C., Bouchez, O., Rispe, C., Mestre, P., Delmotte, F., 2019. A high-quality grapevine downy mildew genome assembly reveals rapidly evolving and lineage-specific putative host adaptation genes. Genome Biol. Evol. 11, 954–969. https://doi.org/10.1093/gbe/evz048.
Fawke, S., Doumane, M., Schornack, S., 2015. Oomycete interactions with plants: infection strategies and resistance principles. Microbiol. Mol. Biol. Rev. 79, 263–280. https://doi.org/10.1128/MMBR.00010-15.
Feehan, J.M., Scheibel, K.E., Bourras, S., Underwood, W., Keller, B., Somerville, S.C., 2017. Purification of high molecular weight genomic DNA from powdery mildew for long-read sequencing. J. Vis. Exp., e55463. https://doi.org/10.3791/55463.
Ferris, M.M., Yan, X., Habbersett, R.C., Shou, Y., Lemanski, C.L., Jett, J.H., Yoshida, T.M., Marrone, B.L., 2004. Performance assessment of DNA fragment sizing by high-sensitivity flow cytometry and pulsed-field gel electrophoresis. J. Clin. Microbiol. 42, 1965–1976. https://doi.org/10.1128/JCM.42.5.1965-1976.2004.
Fletcher, K., Gil, J., Bertier, L.D., Kenefick, A., Wood, K.J., Zhang, L., Reyes-Chin-Wo, S., Cavanaugh, K., Tsuchida, C., Wong, J., Michelmore, R., 2019. Genomic signatures of heterokaryosis in the oomycete pathogen *Bremia lactucae*. Nat. Commun. 10, 2645. https://doi.org/10.1038/s41467-019-10550-0.
Gascuel, Q., Martinez, Y., Boniface, M.-C., Vear, F., Pichon, M., Godiard, L., 2015. The sunflower downy mildew pathogen *Plasmopara halstedii*. Mol. Plant Pathol. 16, 109–122. https://doi.org/10.1111/mpp.12164.
Gascuel, Q., Buendia, L., Pecrix, Y., Blanchet, N., Muños, S., Vear, F., Godiard, L., 2016. RXLR and CRN effectors from the sunflower downy mildew pathogen *Plasmopara halstedii* induce hypersensitive-like responses in resistant sunflower lines. Front. Plant Sci. 7, 1887. https://doi.org/10.3389/fpls.2016.01887.
Jain, M., Olsen, H.E., Paten, B., Akeson, M., 2016. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biol. 17, 239. https://doi.org/10.1186/s13059-016-1103-0.
Kamoun, S., Furzer, O., Jones, J.D.G., Judelson, H.S., Ali, G.S., Dalio, R.J.D., Roy, S.G., Schena, L., Zambounis, A., Panabières, F., Cahill, D., Ruocco, M., Figueiredo, A., Chen, X.-R., Hulvey, J., Stam, R., Lamour, K., Gijzen, M., Tyler, B.M., Grünwald, N.J., Mukhtar, M.S., Tomé, D.F.A., Tör, M., Van Den Ackerveken, G., McDowell, J., Daayf, F., Fry, W.E., Lindqvist-Kreuze, H., Meijer, H.J.G., Petre, B., Ristaino, J., Yoshida, K., Birch, P.R.J., Govers, F., 2015. The top 10 oomycete pathogens in molecular plant pathology. Mol. Plant Pathol. 16, 413–434. https://doi.org/10.1111/mpp.12190.
Lee, M.K., Park, H.S., Han, K.H., Hong, S.B., Yu, J.H., 2017. High molecular weight genomic DNA mini-prep for filamentous fungi. Fungal Genet. Biol. 104, 1–5. https://doi.org/10.1016/j.fgb.2017.04.003.
Miller, M.E., Zhang, Y., Omidvar, V., Sperschneider, J., Schwessinger, B., Raley, C., Palmer, J.M., Garnica, D., Upadhyaya, N., Rathjen, J., Taylor, J.M., Park, R.F., Dodds, P.N., Hirsch, C.D., Kianian, S.F., Figueroa, M., 2018. *De Novo* assembly and phasing of Dikaryotic genomes from two isolates of *Puccinia coronata* f. sp. *avenae*, the causal agent of oat crown rust. mBio 9https://doi.org/10.1128/mBio.01650-17. e01650-17.
Pecrix, Y., Penouilh-Suzette, C., Muños, S., Vear, F., Godiard, L., 2018a. Ten broad Spectrum resistances to downy mildew physically mapped on the sunflower genome. Front. Plant Sci. 9, 1780. https://doi.org/10.3389/fpls.2018.01780.
Pecrix, Y., Staton, S.E., Sallet, E., Lelandais-Brière, C., Moreau, S., Carrère, S., Blein, T., Jardinaud, M.-F., Latrasse, D., Zouine, M., Zahm, M., Kreplak, J., Mayjonade, B., Satgé, C., Perez, M., Cauet, S., Marande, W., Chantry-Darmon, C., Lopez-Roques, C., Bouchez, O., Bérard, A., Debellé, F., Muños, S., Bendahmane, A., Bergès, H., Niebel, A., Buitink, J., Frugier, F., Benhamed, M., Crespi, M., Gouzy, J., Gamas, P., 2018b. Whole-genome landscape of *Medicago truncatula* symbiotic genes. Nat. Plants 4, 1017–1025. https://doi.org/10.1038/s41477-018-0286-7.
Pecrix, Y., Buendia, L., Penouilh-Suzette, C., Maréchaux, M., Legrand, L., Bouchez, O., Rengel, D., Gouzy, J., Cottret, L., Vear, F., Godiard, L., 2019. Sunflower resistance to multiple downy mildew pathotypes revealed by recognition of conserved effectors of the oomycete *Plasmopara halstedii*. Plant J. 97, 730–748. https://doi.org/10.1111/tpj.14157.
Roberts, R.J., Carneiro, M.O., Schatz, M.C., 2013. The advantages of SMRT sequencing. Genome Biol. 14, 405. https://doi.org/10.1186/gb-2013-14-6-405.
Sedlazeck, F.J., Lee, H., Darby, C.A., Schatz, M.C., 2018. Piercing the dark matter: bioinformatics of long-range sequencing and mapping. Nat. Rev. Genet. 19, 329–346. https://doi.org/10.1038/s41576-018-0003-4.
Sharma, R., Xia, X., Cano, L.M., Evangelisti, E., Kemen, E., Judelson, H., Oome, S., Sambles, C., van den Hoogen, D.J., Kitner, M., Klein, J., Meijer, H.J.G., Spring, O., Win, J., Zipper, R., Bode, H.B., Govers, F., Kamoun, S., Schornack, S., Studholme, D.J., Van den Ackerveken, G., Thines, M., 2015. Genome analyses of the sunflower pathogen *Plasmopara halstedii* provide insights into effector evolution in downy mildews and *Phytophthora*. BMC Genomics 16, 741. https://doi.org/10.1186/s12864-015-1904-7.
Shen, D., Tang, Z., Wang, C., Wang, J., Dong, Y., Chen, Y., Wei, Y., Cheng, B., Zhang, M., Grenville-Briggs, L.J., Tyler, B.M., Dou, D., Xia, A., 2019. Infection mechanisms and putative effector repertoire of the mosquito pathogenic oomycete *Pythium guiyangense* uncovered by genomic analysis. PLoS Genet. 15, e1008116. https://doi.org/10.1371/journal.pgen.1008116.
Shi, L., Guo, Y., Dong, C., Huddleston, J., Yang, H., Han, X., Fu, A., Li, Q., Li, N., Gong, S., Lintner, K.E., Ding, Q., Wang, Z., Hu, J., Wang, D., Wang, F., Wang, L., Lyon, G.J., Guan, Y., Shen, Y., Evgrafov, O.V., Knowles, J.A., Thibaud-Nissen, F., Schneider, V., Yu, C.-Y., Zhou, L., Eichler, E.E., So, K.-F., Wang, K., 2016. Long-read sequencing and *de novo* assembly of a Chinese genome. Nat. Commun. 7, 12065. https://doi.org/10.1038/ncomms12065.
Thomma, B.P.H.J., Seidl, M.F., Shi-Kunne, X., Cook, D.E., Bolton, M.D., van Kan, J.A.L., Faino, L., 2016. Mind the gap; seven reasons to close fragmented genome assemblies. Fungal Genet. Biol. 90, 24–30. https://doi.org/10.1016/j.fgb.2015.08.010.
Yang, M., Duan, S., Mei, X., Huang, H., Chen, W., Liu, Y., Guo, C., Yang, T., Wei, W., Liu, X., He, X., Dong, Y., Zhu, S., 2018. The *Phytophthora cactorum* genome provides insights into the adaptation to host defense compounds and fungicides. Sci. Rep. 8, 6534. https://doi.org/10.1038/s41598-018-24939-2.
Zelaya-Molina, L.X., Ortega, M.A., Dorrance, A.E., 2011. Easy and efficient protocol for oomycete DNA extraction suitable for population genetic analysis. Biotechnol. Lett. 33, 715–720. https://doi.org/10.1007/s10529-010-0478-3.