

# Genomic and Transcriptomic Resources for Marker Development in *Synchytrium endobioticum*, an Elusive but Severe Potato Pathogen

Friederike Busse, Annette Bartkiewicz, Diro Terefe-Ayana, Frank Niepold, Yvonne Schleusner, Kerstin Flath, Nicole Sommerfeldt-Impe, Jens Lübeck, Josef Strahwald, Eckhard Tacke, Hans-Reinhard Hofferbert, Marcus Linde, Jarosław Przetakiewicz, and Thomas Debener

First, second, third, twelfth, and fourteenth authors: Institute for Plant Genetics, Leibniz Universität Hannover, Herrenhäuser Str. 2, 30419 Hannover, Germany; fourth author: Julius Kühn Institute (JKI), Institute of Epidemiology and Pathogen Diagnostics, Braunschweig, Germany; fifth, sixth, and seventh authors: JKI, Federal Research Centre for Cultivated Plants, Institute for Plant Protection of Field Crops and Grassland, Kleinmachnow, Germany; eighth and ninth authors: SaKa Pflanzenzucht GmbH & Co. KG, Windeby, Germany; tenth and eleventh authors: Böhm-Nordkartoffel Agrarproduktion GmbH & Co. OHG, Ebstorf, Germany; and thirteenth author: Plant Breeding and Acclimatization Institute-National Research Institute, Radzikow 05-870, Blonie, Poland.

Accepted for publication 3 November 2016.

## ABSTRACT

*Synchytrium endobioticum* is an obligate biotrophic fungus that causes wart diseases in potato. Like other species of the class Chytridiomycetes, it does not form mycelia and its zoospores are small, approximately 3 µm in diameter, which complicates the detection of early stages of infection. Furthermore, potato wart disease is difficult to control because belowground organs are infected and resting spores of the fungus are extremely durable. Thus, *S. endobioticum* is classified as a quarantine organism. More than 40 *S. endobioticum* pathotypes have been reported, of which pathotypes 1(D1), 2(G1), 6(O1), 8(F1), and 18(T1) are the most important in Germany. No molecular methods for the differentiation of pathotypes are available to date. In this work, we sequenced both genomic DNA and cDNA of the German pathotype 18(T1) from infected potato tissue and generated 5,422 expressed sequence tags (EST) and 423 genomic contigs. Comparative sequencing of

33 genes, single-stranded confirmation polymorphism (SSCP) analysis with polymerase chain reaction fragments of 27 additional genes, as well as the analysis of 41 simple sequence repeat (SSR) loci revealed extremely low levels of variation among five German pathotypes. From these markers, one sequence-characterized amplified region marker and five SSR markers revealed polymorphisms among the German pathotypes and an extended set of 11 additional European isolates. Pathotypes 8(F1) and 18(T1) displayed discrete polymorphisms which allow their differentiation from other pathotypes. Overall, using the information of the six markers, the 16 isolates could be differentiated into three distinct genotype groups. In addition to the presented markers, the new collection of EST from genus *Synchytrium* might serve in the future for molecular taxonomic studies as well as for analyses of the host–pathogen interactions in this difficult pathosystem.

*Synchytrium endobioticum* (Schilb.) Percival is the causative agent of potato wart disease. It is a soilborne obligate biotrophic organism that belongs to the fungal phylum Chytridiomycota. *S. endobioticum* is an economically important quarantine organism, and infection of the host can lead to unmarketable tubers and complete yield losses (Hampson 1993; Melnik 1998). Once fields are contaminated with *S. endobioticum*, potato production is forbidden until the soil is proven to be free of sporangia (Obidiegwu et al. 2014). However, the sporangia can survive in the soil for 30 years or longer (Arora and Khurana 2004; Laidlaw 1985). After 43 years, in favorable conditions, disease recurrence may occur even from single spores of *S. endobioticum* (Przetakiewicz 2015b). Because the thick-walled winter sporangia are extremely durable, effective chemical control measures are not available (Obidiegwu et al. 2014). Strict quarantine measures and the use of resistant cultivars have been used to control the fungus but local occurrence of the fungus has been reported in nearly all European countries (Çakır 2005; De Boer 2001; Flath et al. 2014; Przetakiewicz 2015a).


*S. endobioticum* exhibits a life cycle with a haploid asexual reproduction phase which can change under stress conditions into sexual reproduction with a diploid phase. In spring uninflagellate

zoospores of approximately 3 µm in diameter are released from infected tissue and may penetrate meristematic tissue of young tubers or potato stolons, leaving their flagellum outside. After infection, the potato host cell greatly enlarges and surrounding cells divide irregularly, forming the tumor-like tissue of the wart. The zoospore inside develops into a haploid, thick-walled sorus containing several prosori or zoosporangia. These so-called summer sporangia can release several hundred haploid zoospores which can infect new tissues. Under unfavorable conditions, the zoospores can act as isogametes and fuse to a diploid biflagellate zygote. The zygotes can penetrate host tissues in the same way as the zoospores. The infected cells are buried deep in the tissue, forming thick-walled resting or winter sporangia which are released to the soil. The diploid nucleus divides repeatedly, with a first reduction division forming uninucleate haploid zoospores which are released after rupturing of the cell wall (Curtis 1921; Lange and Olson 1981).

More than 40 *S. endobioticum* pathotypes have been reported (Baayen et al. 2006; Çakır et al. 2009; Przetakiewicz 2015a), and 5 of these—namely, pathotypes 1(D1), 2(G1), 6(O1), 8(F1), and 18(T1)—are the most important and widespread pathotypes in Europe (Anonymous 2004; Flath et al. 2014; Obidiegwu et al. 2014). At present, a conventional, time-consuming and laborious inoculation assay using differential cultivars is the only method available for distinguishing different *S. endobioticum* pathotypes (Anonymous 2004; Baayen et al. 2006; Obidiegwu et al. 2014). Classification using differential cultivars may also incur bias in the assignment of pathotypes because different countries use different potato differential cultivars, and the inoculation assay itself may result in inconsistencies

Corresponding author: D. Terefe-Ayana; E-mail address: terefe@genetik.uni-hannover.de

\*The e-Xtra logo stands for “electronic extra” and indicates that five supplementary figures and three supplementary tables are published online.

 Copyright © 2017 The Author(s). This is an open access article distributed under the CC BY-NC-ND 4.0 International license.

between different laboratories using the same cultivars (Flath et al. 2014). Therefore, there is a need to complement the use of differential cultivars with molecular diagnostic techniques.

As an alternative to the time-consuming detection methods, highly sensitive polymerase chain reaction (PCR)-based methods have been proposed. The described protocols include simple PCR and real-time PCR-based detection (Bonants et al. 2015; Niepold and Stachewicz 2004; van den Boogert et al. 2005; van Gent-Pelzer et al. 2010), microarray-based detection (Abdullahi et al. 2005), and colorimetric detection using a peptide nucleic acid probe and a cyanine dye (Duy et al. 2015). In addition, TaqMan PCR assays based on ribosomal DNA sequences have also been developed (Smith et al. 2014; van Gent-Pelzer et al. 2010). More recently, a real-time TaqMan PCR assay was described to differentiate pathotype 1(D1) from the other four pathotypes: 2(G1), 6(O1), 8(F1), and 18(T1). However, these methods are either inconsistent or need expensive instruments and consumables. Therefore, simple, affordable, and reliable molecular diagnostic techniques for *S. endobioticum* are still missing.

Genomic and transcriptomic sequence information from plant pathogens is important for the understanding of several aspects of plant–pathogen interaction. It may be used to discriminate pathotypes and is crucial to reveal genetic diversity within pathogen populations which, in turn, facilitates pathogen control strategies (McDonald and Linde 2002a,b). It may also lead to the identification of pathotype-specific effector genes (Jones and Dangl 2006) on which disease management measures can be based. However, at present, there are only eight *S. endobioticum* gene sequences available in public databases, including GenBank, European Nucleotide Archive (ENA), and DNA Data Bank of Japan (DDBJ), and these sequences originate from the 18S and 28S ribosomal RNA gene sequences. Although such nuclear ribosomal sequences have been reported as fungal DNA barcode markers (Schoch et al. 2012), the sequencing of intergenic nontranscribed spacers of ribosomal genes from four *S. endobioticum* pathotypes showed no sequence polymorphisms (Smith et al. 2014), indicating that ribosomal gene sequencing is poorly suited for dissecting *S. endobioticum* pathotypes.

Therefore, the objectives of the study presented here are to provide high-quality genomic and transcriptome sequence data of *S. endobioticum*, develop molecular markers for diagnosis, and assess polymorphisms and differentiate pathotype groups.

## MATERIALS AND METHODS

**Fungal pathotypes.** Eleven isolates representing different known pathotypes and five isolates of unknown *S. endobioticum* pathotypes from five countries were used in this study (Table 1). These five isolates (numbers SE4, SE5, SE6, SE7, and SE13) were from a newly characterized recent finding and are in the process of being assigned to a pathotype (J. Przetakiewicz, personal communication). Fourteen of the isolates were obtained as fresh wart tissue from the Julius Kühn Institute (Germany) and Plant Breeding and Acclimatization Institute (IHAR) (Poland). Two of the pathotypes (numbers SE2 and SE11) were obtained as wart compost from Germany and Northern Ireland (Table 1). The pathotype nomenclature is according to the standard code (Baayen et al. 2006; Langerfeld et al. 1994). The German *S. endobioticum* pathotype 18(T1) was selected for sequencing based on its aggressiveness and its wide distribution (Stachewicz 2002). Susceptible ‘Tomensa’ potato was used for propagation and maintenance of the fungus.

**DNA and RNA extraction.** Sixteen days postinoculation, the proliferating tumor-like tissue of 14 isolates (SE1, SE3 to SE10, and SE12 to SE16) (Table 1) was harvested in 100-mg aliquots and frozen in liquid nitrogen. Dried winter spores (30 mg) of two isolates (SE2 and SE11) (Table 1) were also aliquoted from the wart compost. The frozen tissue or dry winter spores were then homogenized with a bead mill (Retsch GmbH, Haan, Germany). DNA was extracted from the fine powder of the 16 isolates using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) according to the

manufacturer’s instructions. Similarly, 100 mg of ground tissue was used from the German pathotype 18(T1; number SE 16) (Table 1) for RNA extraction using the RNeasy Plant Mini Kit (Qiagen) according to the manufacturer’s instructions. Removal of the remaining DNA from the extracted RNA was performed using the DNase-free kit from Ambion (Ambion, Cambridgeshire, UK), as recommended by the manufacturer. The DNA and RNA concentrations were determined using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA), and the integrity was checked on a 1% agarose gel.

**Whole-genome or transcriptome shotgun sequencing.** High molecular weight DNA and RNA samples (20 µg) from the German pathotype 18(T1) were delivered on dry ice to Eurofins Genomics (Eurofins Genomics GmbH, Ebersberg, Germany), where genomic DNA and cDNA library preparations were performed and sequenced on an Illumina HiSeq 2500 in the 2-by-150-bp paired-end mode. Two channels of the Illumina HiSeq 2500 (one for the genomic library and the other for the cDNA library) were used for the sequencing. The whole-genome and transcriptome shotgun sequencing were performed using a single source of tissue from Tomensa potato infected with the German *S. endobioticum* pathotype 18(T1, SE 16) (Table 1).

**Sequence analyses and selection of *S. endobioticum* sequences.** The Illumina TruSeq adapters were trimmed from the raw reads using CLC Genomics Workbench 7.5 (Qiagen). The sequences with a Phred quality score of less than 30 were removed. The trimmed reads were mapped to the potato, tomato, and *Batrachochytrium dendrobatidis* genome sequences separately using the default parameters of CLC Genomics Workbench 7.5. In addition, the cDNA reads were mapped to potato and tomato transcripts. The potato genome sequence (*Solanum tuberosum* Group Phureja DM) and the tomato genome sequence (*S. lycopersicum* ‘Heinz 1706’) were downloaded from public websites (The Potato Genome Sequencing Consortium 2011; The Tomato Genome Consortium 2012). The draft genome of *B. dendrobatidis* was obtained from the Broad Institute public website (<https://www.broadinstitute.org/>).

The sequence reads that were not mapped to the potato and tomato genomes and transcriptomes were collected as non-Solanaceae and, therefore, putative *Synchytrium endobioticum* reads. The reads mapped to *B. dendrobatidis* were also collected and included in the non-Solanaceae reads.

TABLE 1. List of investigated *Synchytrium endobioticum* pathotypes and isolates

Number	Pathotypes/isolates	Country of collection	Source material
SE1	1(D1)	Germany	Fresh wart tissue
SE2	1(D1)	Northern Ireland	Wart compost
SE3	39(P1)	Poland	Fresh wart tissue
SE4	D12 <sup>a</sup>	Denmark	Fresh wart tissue
SE5	D14 <sup>a,b</sup>	Denmark	Fresh wart tissue
SE6	P2/15 <sup>a</sup>	Poland	Fresh wart tissue
SE7	D22 <sup>a</sup>	Denmark	Fresh wart tissue
SE8	2(Ch1)	Poland	Fresh wart tissue
SE9	3(M1)	Poland	Fresh wart tissue
SE10	2(G1)	Germany	Fresh wart tissue
SE11	2(G1)	Germany	Wart compost
SE12	6(O1)	Germany	Fresh wart tissue
SE13	D25 <sup>a,c</sup>	Denmark	Fresh wart tissue
SE14	8(F1)	Germany	Fresh wart tissue
SE15	18(T1)	Greece	Fresh wart tissue
SE16	18(T1) <sup>d</sup>	Germany	Fresh wart tissue

<sup>a</sup> Newly described isolates in the process of being assigned to a pathotype.

<sup>b</sup> Isolate from Denmark with the same genotype as the Polish pathotype 2(Ch1).

<sup>c</sup> Isolate from Denmark with the same genotype as the German pathotype 8(F1).

<sup>d</sup> The German pathotype 18(T1) used for the genome and transcriptome shotgun sequencing.

**Sequence assembly and annotation of the contigs.** The non-Solanaceae sequence reads and those mapped to *B. dendrobatidis* were assembled using CLC Genomics Workbench 7.5. The default assembly parameters were used, with the exception that the similarity fraction was adjusted to 90%. The assembly was performed for the selected cDNA and genomic DNA separately. The assembly was also performed for the whole trimmed raw reads of the cDNA and genomic DNA separately.

Automatic annotation of the assembled cDNA contigs was performed by a BLAST search of the National Center for Biotechnology Information (NCBI) database (Altschul et al. 1997) with the Blast2GO software (Götz et al. 2008) using a BLAST threshold value of 0.1. Gene ontology (GO) term analysis and mapping to the Enzyme Commission (EC) classification databases were also performed using the Blast2GO software (Götz et al. 2008). Further selection and refinement of *S. endobioticum* sequences were performed based on the BLAST hits.

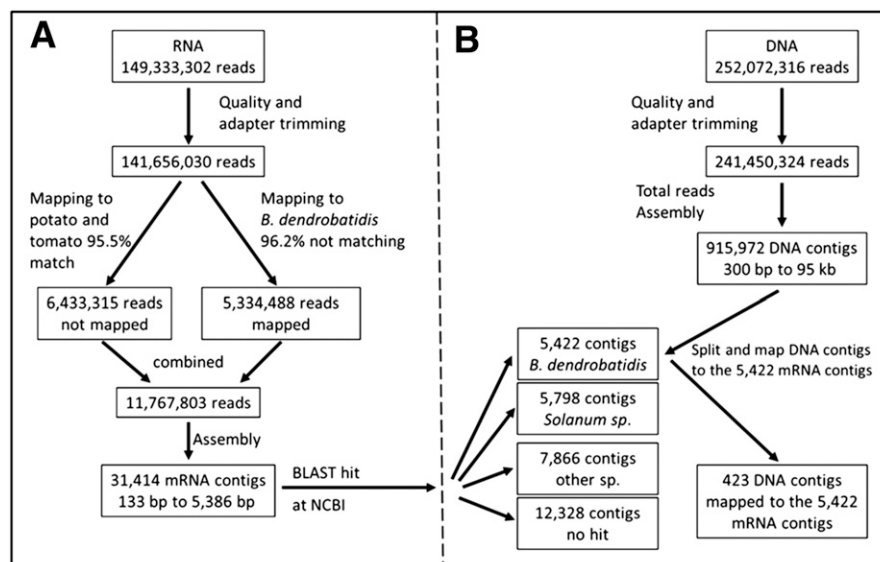
**Verification of selected *S. endobioticum* contigs by PCR.** Primer sets were designed for 96 randomly selected putative *S. endobioticum* transcriptome contig sequences and verified by PCR using DNA isolated from four German pathotypes—1(D1), 2(G1), 8(F1), 18(T1)—and uninfected Tomensa leaves. PCR mixtures were prepared using 20 ng of genomic DNA in a total volume of 25 µl containing 0.5 µM each of the forward and reverse primers, 1 U of Taq polymerase, and 0.12 mM dNTP in 1× buffer provided by the manufacturer (Bioline, Luckenwalde, Germany). The PCR conditions were as follows: initial denaturation for 4 min at 94°C; followed by 35 cycles of 45 s at 94°C, 1 min at 64°C, and 1.5 min at 72°C; and a final extension of 10 min at 72°C. The PCR products were mixed with 10% Orange G loading dye (0.25% Orange G, 30% glycerine, and 1 mM EDTA) and separated on a 1.5% agarose gel in 1× Tris-acetate-EDTA buffer.

**Sequence-characterized amplified region marker development.** Of the 96 primer sets that were verified by PCR above, 1 amplified a polymorphic fragment specific to pathotype 1(D1) and an additional fragment that was common to all five pathotypes (Supplementary Fig. S4A). The PCR product specific for the single pathotype 1(D1) was excised from the agarose gel, purified using a QIAquick PCR Purification Kit (Qiagen), and completely sequenced by Sanger sequencing. Using the sequence information of this pathotype 1(D1)-specific fragment, a sequence-characterized amplified region (SCAR) marker, called 14425P1, was developed. The sequences of the SCAR

marker 14425P1 primers were as follows: forward, 5' GGTTC GCAGCGAGCTAGATA 3' and reverse, 5' TGAGCTCGTCACT ATCATGG 3'. PCR mixtures were prepared using 20 ng of genomic DNA in a total volume of 25 µl containing 0.25 µM each of the forward and reverse primers, 1 U of Taq polymerase, and 0.12 mM dNTP in 1× buffer provided by the manufacturer (Bioline). The PCR conditions were as follows: initial denaturation for 1 min at 94°C; followed by 30 cycles of 30 s at 94°C, 30 s at 62°C, and 45 s at 72°C; and a final extension of 2 min at 72°C. The PCR products were mixed with 10% Orange G loading dye and separated on a 1.5% agarose gels in 1× Tris-acetate-EDTA buffer. In the PCR mixtures of the SCAR marker 14425P1, a primer pair from the sporulation protein RMD1 (forward, 5' GTTCTTGATGCACCCGA 3' and reverse, 5' TC TCATCGTGACTGCTCGTC 3') was included as an internal control.

**Additional resequencing and single-stranded confirmation polymorphism analyses.** Sixty of the genes that were verified by PCR and showed a single banding pattern were amplified from the DNA of the five German *S. endobioticum* pathotypes 1(D1), 2(G1), 6(O1), 8(F1), and 18(T1). The PCR was performed with an annealing temperature of 64°C using proofreading PrimeSTAR HS DNA polymerase (TAKARA BIO Inc., Saint-Germain-en-Laye, France) following the procedures described by the manufacturer. PCR products from 33 of the 60 genes from all five German pathotypes were purified using a QIAquick PCR Purification Kit (Qiagen) and completely sequenced by Sanger sequencing. Comparative sequence analyses among the five pathotypes were made for all 33 genes using CLC Genomics Workbench 7.5. PCR products of the remaining 27 genes from the five German pathotypes were analyzed by single-stranded confirmation polymorphism (SSCP) (Orita et al. 1989). The 27 genes were selected based on their smaller fragment size appropriate for an SSCP analysis. For the SSCP analyses, the PCR products were separated using nondenaturing polyacrylamide gel electrophoresis on 0.5 × mutation detection enhancement gel (Biozym, Hessisch Oldendorf, Germany).

**Selection of *S. endobioticum* genomic contigs.** *S. endobioticum* expressed sequence tags (EST) were used as a reference to capture the corresponding genomic sequences. First, the larger genomic contigs were split into 1,000-bp fragments with 100-bp overlaps using splitter (<http://emboss.bioinformatics.nl/cgi-bin/emboss/splitter>). The genomic contigs were then mapped to the *S. endobioticum* EST. The mapped *S. endobioticum* genomic contigs were then extracted and



**Fig. 1.** Schematic illustration of **A**, transcriptome sequence analyses steps for the preliminary enrichment of *Synchytrium endobioticum* cDNA sequences and **B**, steps to obtain the 423 DNA contigs corresponding to the 5,422 *S. endobioticum* mRNA contigs.

the corresponding larger contigs were obtained from the original assembly by using CLC Genomics Workbench 7.5.

**Identification of *S. endobioticum* simple sequence repeat motifs.** The selected *S. endobioticum* pathotype 18(T1) genomic contig sequences were screened for hexa-, penta-, tetra-, tri-, and dinucleotide simple sequence repeat (SSR) motifs using the “SSR locator” software (da Maia et al. 2008). Primer sets corresponding to 41 randomly selected SSR motifs were designed, and each of the forward primers was end-labeled with infrared fluorescent dye 700. PCR mixtures were prepared using 20 ng of genomic DNA in a total volume of 25 µl containing 0.25 µM each of the forward and reverse primers, 1 U of Taq polymerase, and 0.12 mM dNTP in 1× buffer provided by the manufacturer (Bioline). The PCR conditions were as follows: initial denaturation for 1 min at 94°C; followed by 30 cycles of 30 s at 94°C, 30 s at 64°C, and 45 s at 72°C; and a final extension of 2 min at 72°C. The PCR products were separated on 6% polyacrylamide gel (Sequagel XR; National Diagnostics Inc., Charlotte, NC) using a DNA Analyzer (LI-COR Biosciences, Lincoln, NE).

## RESULTS

**Whole-transcriptome shotgun sequencing and selection of *S. endobioticum* sequences.** In total, 149,333,302 reads were obtained from the whole-transcriptome shotgun sequencing of proliferated tissue from Tomensa potato infected with the German

*S. endobioticum* pathotype 18(T1). After quality and adapter trimming, 141,656,030 reads remained (Fig. 1A).

The mapping of these reads to the potato and tomato genomic and transcriptomic sequences revealed 95.5% matching sequences. The 4.5% (6,433,315 reads) sequences that were not mapped to these sequences were selected as putative *S. endobioticum* reads. In parallel, the 141,656,030 trimmed reads were compared using BLAST to the *B. dendrobatidis* genome and transcriptome, and 3.8% (5,334,488 reads) produced a significant match. These 5,334,488 reads were also selected as putative *S. endobioticum* reads. Although the majority of the reads that were mapped to *B. dendrobatidis* could have already been represented in the 6,433,315 nonpotato and nontomato reads, some additional reads were obtained (Fig. 1A).

The 6,433,315 nonpotato and nontomato reads and the 5,334,488 reads that matched *B. dendrobatidis* were combined and assembled into 31,414 contigs, which ranged in length from 133 to 5,386 bp (Fig. 1A) and had an N50 (an assembled contig size above which the size of 50% of the whole assembled contigs are represented) of 328 bp.

**Annotation of the contigs through a sequence similarity search of the NCBI database.** A BLAST similarity search identified 5,422 contigs with hits to *B. dendrobatidis*, 5,798 contigs with hits to *Solanum* spp., 12,328 contigs without any hits and 7,866 contigs with hits to other species.

The 5,422 mRNA contigs with a BLAST hit to *B. dendrobatidis* sequences are hereafter referred to as *S. endobioticum* EST. The sizes of these EST ranged from 200 to 2,979 bp, and 3,849 EST were

TABLE 2. Descriptions of the 41 simple sequence repeat motifs and the corresponding primer sets evaluated using DNA from the five German *Synchytrium endobioticum* pathotypes 1(D1), 2(G1), 6(O1), 8(F1), and 18(T1)

Primer set	Forward	Reverse	Motif	Product size (bp)
2978	5' ACGTATGCACGGATCAAACA 3'	5' TACAACCCACCTTTTCCTG 3'	(TA)11	323/330
11176	5' TGCACCTTGTACCAGCCAAG 3'	5' TATACCGACCTGTCCAAGCC 3'	(AT)7	384/390
4865	5' CTACAAGCCAGTACCAGTCA 3'	5' CCACCGAATATCCCATCATC 3'	(GGAT)8	392/402
16161	5' GGCTTCGTCTTCAGCGATAC 3'	5' CCATCTCCGGTCTGTTTCAT 3'	(GGTGTC)4	369/377
818598	5' GGCCAGACGTCGATATTTCT 3'	5' GTACGACCTCCCAGCATTGT 3'	(ATAGC)6	294/304
4761	5' TGTGCACACTTTGAGAAAGG 3'	5' GGTTCACGACGTTGGTAGT 3'	(CT)4-(AT)4	383
7711a	5' GACCTTGTGTCCCGGAGTAA 3'	5' TAGCTTCGCGGATAGGAAGA 3'	(GT)11	196
7711b	5' ACCAAGAAGCGACCATATT 3'	5' AATAAGCGAGGAATCAGCGA 3'	(TA)11	286
459	5' AAAGCTAATGGCATGGTTGG 3'	5' AGACGCCATCACTATGCCTC 3'	(TC)5-(GA)5	364
t322060	5' GATCAACATGTGCTGGATCG 3'	5' ATGATTATGACACGGTGCGA 3'	(TC)6-(GA)4	217
3714	5' AGAACGTCTCTGCTGCCAAT 3'	5' TTTGGTGTGCCCCTGTGTGA 3'	(AT)11	350
492	5' AACTGACCTTCTCCACATC 3'	5' TGATGGATGATGCCTGCT 3'	(TA)8	384
4846	5' AGTATGGGACGTCAGAACCG 3'	5' TAACCGTGGGATATCATGGC 3'	(AC)4-(AC)5	223
9030	5' GGACGGATCTGCCATAGTA 3'	5' GTCATGACTTAGGGGGCAA 3'	(AT)4-(TA)4	315
3881	5' TCGTAGATGTCGTGGTGCAG 3'	5' CAGTAGGACAGGGGACCAGA 3'	(ACA)7	287
t107447a	5' AACAGAAAGCGTGGTGTGG 3'	5' ATCAGTGAGAGCCCAAG 3'	(TA)8	393
t107447b	5' GTATTCGGAGGTGTCGCTGT 3'	5' TTGTGTCCACCGAGTATCCA 3'	(CGT)9	267
t107447c	5' AAGGCGTGCACGAGATAGTT 3'	5' TAATCTGACTTGCTGTGCG 3'	(GCGT)7	297
639	5' TGGTCTTCCAATGGCCCTAAC 3'	5' AAATGCCAACCACGTAGGAG 3'	(CAT)7	327
3129	5' AATCCACACGAGTGGACATC 3'	5' AATTCCAACATGTAGGCG 3'	(AAG)8	358
2826	5' GAATACGCCCCGAACCTAACCA 3'	5' TTTCTGCACCTTGTGCGACTG 3'	(TACGGA)4	369
2035	5' GTCGCCCTTGATTCATGTTT 3'	5' AACTGTCGCTGAAGAGGGAA 3'	(CCCA)7	259
2035A	5' TTCCCTCTCAGCGACAGTT 3'	5' ATTTACACTGGGGGTAGGGG 3'	(CTA)11	328
2035B	5' CATGTGTTTGACCGAGGATG 3'	5' TGCAATGAAAGCTATGCTGG 3'	(GCAAAA)5	317
3244	5' ATCTATCATGCATCCCGTCC 3'	5' CACTCCACTTGTCCGACTGA 3'	(CGGT)6	361
5274	5' TTGCAATGCGTTACTTGCTC 3'	5' GTTGCGCGATGATCACTAGA 3'	(GGCGTG)4	262
2733	5' ATTGGTGTGAAGATGAGGGC 3'	5' ATCCATGACGTGCAACCTC 3'	(GGTGTG)4	106
2894	5' CAACACACGAGTGGACATC 3'	5' ATGTGTAATGGTGTGCGGA 3'	(CAAC)5	313
2785	5' GAGGTCGAGTCTTTGCGTTC 3'	5' AGGTTCTCTGTCCATCGTG 3'	(CAGCCG)5	276
2723	5' TCGTTGCTGCACCTACCTG 3'	5' GTTTCGATTGGGCATCAACT 3'	(TGCT)6	307
693	5' ATTAATTTGGCCAAGCAACG 3'	5' CCTCTATAGGAATGCGCAGC 3'	(TGTCGC)6	371
705	5' TGCAAGGAATGCACACTATTG 3'	5' CTCGTGCTGCTGATTGT 3'	(ACAGG)5	348
345443	5' TGGACTTCCCTACGTTGGTC 3'	5' CCATGATGGATGTGGATGAA 3'	(TGGTGC)7	399
442158	5' AATCGGTACAGCTTGGGATG 3'	5' GACGAGGAGTTTGGATGAA 3'	(GTGCC)5	263
213091	5' TCGAGCTCCTCTAGTGC 3'	5' TTAATGGGGTCTCGTCAGG 3'	(CTTCT)4	248
4884	5' CGCCATCATGATCCTTCTTT 3'	5' CACTGTTCCGGTCTTTTTC 3'	(CATCC)5	384
t397081	5' CGGGAGTCGTATGGCAGTAT 3'	5' GGTGTGTGATGCAGGTATGC 3'	(TA)12	315
29983	5' TCTTGCATGGCCAGTAGACA 3'	5' TATAGTCATCCTCCCTGGCC 3'	(TA)13	203
2671	5' TGAGAAGAGCCGAAGATTCC 3'	5' CATCCATCCATCCATCCATC 3'	(TAGA)6	100
24845	5' AAACCTTACCGTGGCCCTT 3'	5' CCATGCATTTGCATCTCTCT 3'	(AT)14	261
626	5' TGTTGTTGTTCTCTGCTGCTC 3'	5' GGAGGCTTTGATATTTCCCC 3'	(CTA)7-(CTA)13	291

longer than 300 bp. This Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the accession GDBM00000000. The version described in this article is the first version, GDBM01000000.

Blast2Go analysis retrieved GO terms for 3,583 of the 5,422 EST. Accordingly, 434 EST were grouped into GO terms for metabolic and oxidation-reduction processes within the biological process category (Supplementary Table S1). Within the molecular function category, 1,010 EST were grouped into GO terms for ATP and protein binding and, within the cellular component category, 452 EST were grouped into GO terms for nucleus and cytoplasm. Among the EST for which GO term descriptions were assigned, 1,007 EST were effectively grouped into the six EC classifications. The majority of the EC-annotated EST belonged to EC 3 (hydrolases), followed by EC 2 (transferases), EC 1 (oxidoreductases), and EC 6 (ligases) (Supplementary Fig. S1).

**PCR verification of some of the 5,422 putative *S. endobioticum* EST.** Primer sets corresponding to 95 of the 96 EST produced the desired PCR fragments specifically from the DNA isolated from proliferating Tomensa tissue infected with *S. endobioticum* pathotypes but not from the uninfected Tomensa DNA (Supplementary Fig. S2). Primer sets from 94 of the 95 EST produced fragments of identical size in all five pathotypes: 1(D1), 2(G1), 6(O1), 8(F1), and 18(T1). However, 1 of the 95 primer sets (14425) amplified a 772-bp fragment specific to only pathotype 1(D1) and a 446-bp fragment common to the five German pathotypes 1(D1), 2(G1), 6(O1), 8(F1), and 18(T1) (Supplementary Fig. S4A). Sixty of the verified genes showed a single banding pattern in each of the pathotypes.

**Screening for polymorphism between *S. endobioticum* pathotypes.** Sequencing of the pathotype 1(D1)-specific 772-bp fragment that was obtained using the primer set 14425 revealed a 326-bp insertion into the 446-bp sequence. The 446-bp sequence was common to all five pathotypes: 1(D1), 2(G1), 6(O1), 8(F1), and 18(T1) (Supplementary Fig. S3). Interestingly, the 326-bp sequence was characterized by a 7- and 65-bp direct repeat. The EST detected by this primer set has a sequence similarity to a transmembrane segment (TMS) membrane protein/tumor differentially expressed gene.

PCR using the SCAR marker 14425P1 confirmed a 367-bp fragment in pathotype 1(D1) which was absent in all other German pathotypes: 2(G1), 6(O1), 8(F1), and 18(T1) (Supplementary Fig. S4B). Further tests using DNA from 11 additional pathotypes and isolates obtained from Denmark, Greece, Northern Ireland, and Poland indicated the presence of the 367-bp fragment in pathotypes 39 (P1) and 2(Ch1) and four isolates. The internal control primer from the sporulation protein RMD1 amplified a 175-bp fragment in all 16 isolates, indicating the intactness of all DNA and the optimum performance of the PCR procedures.

Sequencing 33 of 60 EST showing a single banding pattern resulted in sequences with a length ranging from 285 to 4,615 bp. However, no sequence variation was detected between the five pathotypes for all 33 EST (Supplementary Table S2). The sequence from these 33 EST covered a region with a size of 37,127 bp. SSCP analysis of the remaining 27 of 60 EST showing a single banding pattern and with less than 690 bp also showed no polymorphisms among the five German *S. endobioticum* pathotypes.

**Whole-genome shotgun sequencing and assembly.** Whole-genome shotgun sequencing of tissue from Tomensa potato infected with the German *S. endobioticum* pathotype 18(T1) resulted in 252,072,316 reads. After quality and adapter trimming, 241,450,324 reads remained. The assembly of the 241,450,324 reads resulted in 915,972 contigs longer than 300 bp, with a maximum contig size of 95 kb.

**Selection of *S. endobioticum* genomic contigs corresponding to the 5,422 EST sequences.** In total, 256,210 contigs were split into 1,000-bp fragments with a 100-bp overlapping sequence. Mapping to the 5,422 EST identified 423 *S. endobioticum* pathotype 18(T1) genomic contigs (Fig. 1B). The maximum size of the genomic contigs identified was 43,140 bp and, together, the 423 contigs covered a size of 2,081,121 bp with a GC content of 47%. This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LDOR00000000. The version described in this article is version LDOR01000000.

**Identification of *S. endobioticum* SSR motifs.** Screening of the 423 putative *S. endobioticum* pathotype 18(T1) genomic contig sequences resulted in 389 motifs with 3 to 26 repeats (Supplementary Table S3). Forty-one primer sets corresponding to 41 randomly selected SSR motifs were successfully evaluated using the five German *S. endobioticum* pathotypes (Table 2). Of the 41 SSR primer sets, 5 sets (2978, 11176, 4865, 16161, and 818598) revealed polymorphisms (Supplementary Fig. S5). These five primer sets also resulted in two different banding patterns on the 11 additional pathotypes and isolates obtained from Denmark, Greece, Northern Ireland, and Poland. The first banding pattern was specific to *S. endobioticum* pathotypes 1(D1), 2(G1), 6 (O1), 2(Ch1), 39(P1), and 3(M1) and the four isolates D12, D14, D22, and P2/15, whereas the second banding pattern was specific to *S. endobioticum* pathotypes 8(F1) and 18(T1) and isolate D25.

## DISCUSSION

**Identification of *S. endobioticum* EST and genomic contigs.** In the study presented here, a small-scale next-generation transcriptome and genome sequencing approach was used for the first time on *S. endobioticum*, a quarantine organism for which only

TABLE 3. Genotypes and grouping of the 16 isolates of *Synchytrium endobioticum* based on the six polymorphic markers

Number	Pathotypes, isolates	14425P1 <sup>a</sup>	2978 <sup>b</sup>	11176 <sup>b</sup>	4865 <sup>b</sup>	16161 <sup>b</sup>	818598 <sup>b</sup>	Group
SE1	1(D1)	+	I	I	I	I	I	1
SE2	1(D1)	+	I	I	I	I	I	1
SE3	39(P1)	+	I	I	I	I	I	1
SE4	D12	+	I	I	I	I	I	1
SE5	D14	+	I	I	I	I	I	1
SE6	P2/15	+	I	I	I	I	I	1
SE7	D22	+	I	I	I	I	I	1
SE8	2(Ch1)	+	I	I	I	I	I	1
SE9	3(M1)	-	I	I	I	I	I	2
SE10	2(G1)	-	I	I	I	I	I	2
SE11	2(G1)	-	I	I	I	I	I	2
SE12	6(O1)	-	I	I	I	I	I	2
SE13	D25	-	II	II	II	II	II	3
SE14	8(F1)	-	II	II	II	II	II	3
SE15	18(T1)	-	II	II	II	II	II	3
SE16	18(T1)	-	II	II	II	II	II	3

<sup>a</sup> Symbols + or - indicate the presence or absence, respectively, of the 367-bp fragment for the sequence-characterized amplified region marker 14425P1.

<sup>b</sup> Presence of the banding pattern "I" or "II" is indicated for the five polymorphic simple sequence repeat loci.

eight 18S and 28S ribosomal gene sequences were available thus far in the whole public sequence databases. Of course, the isolation of sequence information from *S. endobioticum* is complicated by the fact that producing a pure culture based on single-spore isolates is not possible because the fungus is an obligate biotroph and impossible to grow on artificial media. *S. endobioticum* produces sporangia carrying motile spores but no hyphae or other fruiting structures (Anonymous 2004; Webster and Weber 2007). Therefore, we decided to isolate the total RNA and DNA from freshly infected tissue and to discriminate *S. endobioticum* sequences from potato sequences through bioinformatics methods. To verify the identified sequences, we designed PCR primer sets for 96 EST sequences and tested them by PCR on four German *S. endobioticum* pathotypes. All except one primer set amplified the expected product from all four German pathotypes.

The lack of an *S. endobioticum* reference genome and the fact that the reference species *B. dendrobatidis* used in this study belongs to a different order than *S. endobioticum* led to fewer BLAST hits compared with, for example, the RNAseq data from *Botrytis cinerea*-infected lettuce tissue (de Cremer et al. 2013). However, the identification of 5,422 EST and the additional 423 genomic contigs of *S. endobioticum* substantially extends the limited information available to date. Many of the remaining 12,328 mRNA contigs without any BLAST hits may also belong to *S. endobioticum*, even though they cannot be identified as such due to the lack of a reference genome that is sufficiently taxonomically close to reveal sequence similarity.

**Polymorphic marker-based grouping of *S. endobioticum* pathotypes.** More than 40 *S. endobioticum* pathotypes are currently known (Baayen et al. 2006; Çakır et al. 2009; Przetakiewicz 2015a), and they are distinguished based on their infection patterns on differential sets of host genotypes. More than half of these pathotypes are either extinct or of minor relevance (Baayen et al. 2006). Pathotype 1(D1) is the oldest pathotype to which most European potato cultivars are resistant (Anonymous 2004; Flath et al. 2014). Pathotypes 8(F1) and 18(T1) seem more virulent than the other common pathotypes. For instance, 7 of the 10 differential cultivars were susceptible to pathotype 18(T1), compared with only 3 and 5 being susceptible to pathotypes 1(D1) and 6(O1), respectively (Anonymous 2004). An independent test showed that pathotypes 8(F1) and 18(T1) are more aggressive than the other pathotypes (K. Flath, personal communication). The SCAR primer set 14425P1 and the five SSR markers from different, nonoverlapping genomic contigs classified the 16 studied *S. endobioticum* collections into three groups (Table 3). The more virulent pathotypes 8(F1) and 18(T1) and isolate D25 are identified in the same group distinct from all the other used pathotypes and isolates. Based on the SCAR and SSR markers as well as by using differential cultivars (J. Przetakiewicz, personal communication), isolate D25 from Denmark is found to be the same genotype as the German pathotype 8(F1) and the same pathotype. Hence, isolate D25 can be represented by pathotype 8(F1). The markers of the pathotype 8(F1) and 18(T1) group may be used in a preevaluation of previously uncharacterized *S. endobioticum* isolates.

**Low levels of diversity between *S. endobioticum* pathotypes.** A comparative sequence analysis of 33 contigs covering a region with a size of 37,127 bp and an SSCP analysis of 27 EST covering a size of 10,059 bp revealed no sequence polymorphisms between the five German *S. endobioticum* pathotypes 1(D1), 2(G1), 6(O1), 8(F1), and 18(T1), indicating that *S. endobioticum* pathotypes may be exceptionally similar to each other. It should be noted that the method used to select amplicons for sequencing was extremely conservative to ensure that all sequences originated from *S. endobioticum*, and this may result in bias toward the sequencing of conserved regions. However, our observation was corroborated by a recent report, in which extremely low levels of diversity were detected among eight isolates by using complexity reduction of

polymorphic sequences analyses (Bonants et al. 2015). This result is astonishing given the clear differentiation of pathotypes based on infection assays. The low sequence diversity is probably due to the development of new *S. endobioticum* pathotypes by very limited changes in avirulence factors rather than extensive genetic recombination between divergent genotypes.

## CONCLUSIONS

For the first time, a large number of *S. endobioticum* genome and transcriptome sequences are presented in this study. Compared with the very limited information available to date, the large set of EST sequences and more than two million bases of genomic sequences reported here will assist future studies in increasing the understanding of the mechanism of tumor formation and in elucidating the nature of the interaction with the potato host. A simple SCAR marker, 14425P1, is presented that distinguishes the German *S. endobioticum* pathotype 1(D1) from the other pathotypes commonly present in Germany. Five SSR primer sets are identified that clearly distinguish the more virulent pathotypes 8(F1) and 18(T1) from the other European pathotypes and isolates. We hope that the data presented here will help to narrow the current gaps in scientific knowledge of *S. endobioticum*.

## ACKNOWLEDGMENTS

The study was supported by the Federal Ministry of Food and Agriculture via the Fachagentur Nachwachsende Rohstoffe e.V. (grant numbers 22036111, 22036211, 22036311, and 22036411). We thank G. Little for providing a pathotype from Northern Ireland.

## LITERATURE CITED

- Abdullahi, I., Koerber, M., Stachewicz, H., and Winter, S. 2005. The 18S rDNA sequence of *Synchytrium endobioticum* and its utility in microarrays for the simultaneous detection of fungal and viral pathogens of potato. *Appl. Microbiol. Biotechnol.* 68:368-375.
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* 25: 3389-3402.
- Anonymous. 2004. EPPO Standards. Diagnostic protocols for regulated pests: *Synchytrium endobioticum*. *Bull. OEPP/EPPO Bull.* 34:213-218.
- Arora, R. K., and Khurana, S. M. P. 2004. Major fungal and bacterial diseases of potato and their management. Pages 189-231 in: *Fruit and Vegetable Diseases. Disease Management of Fruits and Vegetables, Vol. 1.* K. G. Mukerji, ed. Springer Netherlands, Dordrecht.
- Baayen, R., Cochius, G., Hendriks, H., Meffert, J., Bakker, J., Bekker, M., van den Boogert, P., Stachewicz, H., and van Leeuwen, G. 2006. History of potato wart disease in Europe—a proposal for harmonisation in defining pathotypes. *Eur. J. Plant Pathol.* 116:21-31.
- Bonants, P. J. M., van Gent-Pelzer, M. P. E., van Leeuwen, G. C. M., and van der Lee, T. A. J. 2015. A real-time TaqMan PCR assay to discriminate between pathotype 1(D1) and non-pathotype 1(D1) isolates of *Synchytrium endobioticum*. *Eur. J. Plant Pathol.* 143:495-506.
- Çakır, E. 2005. First report of potato wart disease in Turkey. *Plant Pathol.* 54: 584.
- Çakır, E., van Leeuwen, G. C. M., Flath, K., Meffert, J. P., Janssen, W. A. P., and Maden, S. 2009. Identification of pathotypes of *Synchytrium endobioticum* found in infested fields in Turkey. *Bull. OEPP/EPPO Bull.* 39:175-178.
- Curtis, M. 1921. The life-history and cytology of *Synchytrium endobioticum* (Schilb) Perc the cause of wart disease in potato. *Philos. Trans. R. Soc. Lond. B* 210:409-478.
- da Maia, L. C., Palmieri, D. A., de Souza, V. Q., Kopp, M. M., de Carvalho, F. I. F., and de Oliveira, A. C. 2008. SSR Locator: Tool for simple sequence repeat discovery integrated with primer design and PCR simulation. *Int. J. Plant Genomics* 2008:Article ID 412696.
- De Boer, S. H. 2001. Occurrence of potato wart caused by *Synchytrium endobioticum* on Prince Edward Island, Canada. *Plant Dis.* 85:1292.
- de Cremer, K., Mathys, J., Vos, C., Froenicke, L., Michelmore, R. W., Cammue, B. P. A., and de Coninck, B. 2013. RNAseq-based transcriptome analysis of *Lactuca sativa* infected by the fungal necrotroph *Botrytis cinerea*. *Plant Cell Environ.* 36:1992-2007.

- Duy, J., Smith, R. L., Collins, S. D., and Connell, L. B. 2015. Rapid Colorimetric Detection of the Fungal Phytopathogen *Synchytrium endobioticum* Using Cyanine dye-Indicated PNA Hybridization. *Am. J. Potato Res.* 92:398-409.
- Flath, K., Przetakiewicz, J., van Rijswijk, P. C. J., Ristau, V., and van Leeuwen, G. C. M. 2014. Interlaboratory tests for resistance to *Synchytrium endobioticum* in potato by the Glynne-Lemmerzahl method. *Bull. OEPP/EPPO Bull.* 44:510-517.
- Götz, S., Garcia-Gomez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., Robles, M., Talon, M., Dopazo, J., and Conesa, A. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 36:3420-3435.
- Hampson, M. C. 1993. History, biology and control of potato wart disease in Canada. *Can. J. Plant Pathol.* 15:223-244.
- Jones, J. D. G., and Dangl, J. L. 2006. The plant immune system. *Nature* 444:323-329.
- Laidlaw, W. M. R. 1985. A method for the detection of the resting sporangia of the potato wart disease (*Synchytrium endobioticum*) in the soil of old outbreak sites. *Potato Res.* 28:223-232.
- Lange, L., and Olson, L. W. 1981. Germination and parasitism of the resting sporangia of *Synchytrium endobioticum*. *Protoplasma* 166:69-82.
- Langerfeld, E., Stachewicz, H., and Rintelen, J. 1994. Pathotypes of *Synchytrium endobioticum* in Germany. *Bull. OEPP/EPPO Bull.* 24:799-804.
- McDonald, B. A., and Linde, C. 2002a. Pathogen population genetics, evolutionary potential, and durable resistance. *Annu. Rev. Phytopathol.* 40:349-379.
- McDonald, B. A., and Linde, C. 2002b. The population genetics of plant pathogens and breeding strategies for durable resistance. *Euphytica* 124:163-180.
- Melnik, P.A. 1998. Wart disease of potato, *Synchytrium endobioticum* (Schilbersky) Percival. EPPO Technical Documents 1032. Paris.
- Niepold, F., and Stachewicz, H. 2004. PCR-detection of *Synchytrium endobioticum* (Schilb.) Perc. *J. Plant Dis. Prot.* 111:313-321.
- Obidiegwu, J. E., Flath, K., and Gebhardt, C. 2014. Managing potato wart: A review of present research status and future perspective. *Theor. Appl. Genet.* 127:763-780.
- Orita, M., Suzuki, Y., Sekiya, T., and Hayashi, K. 1989. Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* 5:874-879.
- Przetakiewicz, J. 2015a. First report of new pathotype 39(P1) of *Synchytrium endobioticum* causing potato wart disease in Poland. *Plant Dis.* 99:285.
- Przetakiewicz, J. 2015b. The viability of winter sporangia of *Synchytrium endobioticum* (Schilb.) Perc. from Poland. *Am. J. Potato Res.* 92:704-708.
- Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Lévesque, C. A., and Chen, W. 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc. Natl. Acad. Sci. USA* 109:6241-6246.
- Smith, D. S., Rocheleau, H., Chapados, J. T., Abbott, C., Ribero, S., Redhead, S. A., Lévesque, C. A., and De Boer, S. H. 2014. Phylogeny of the genus *Synchytrium* and the development of TaqMan PCR assay for sensitive detection of *Synchytrium endobioticum* in soil. *Phytopathology* 104:422-432.
- Stachewicz, H. 2002. Bedeutung und Bekämpfung des Kartoffelkrebses in Deutschland. *Mitt. Biol. Bundesanst. Land-Forstwirtschaft.* 388:16-24.
- The Potato Genome Sequencing Consortium. 2011. Genome sequence and analysis of the tuber crop potato. *Nature* 475:189-195.
- The Tomato Genome Consortium. 2012. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485:635-641.
- van den Boogert, P., van Gent-Pelzer, M. P. E., Bonants, P. J. M., De Boer, S. H., Wander, J. G. N., Lévesque, C. A., van Leeuwen, G. C. M., and Baayen, R. P. 2005. Development of PCR-based detection methods for the quarantine phytopathogen *Synchytrium endobioticum*, causal agent of potato wart disease. *Eur. J. Plant Pathol.* 113:47-57.
- van Gent-Pelzer, M., Krijger, M., and Bonants, P. J. M. 2010. Improved real-time PCR assay for detection of the quarantine potato pathogen, *Synchytrium endobioticum*, in zonal centrifuge extracts from soil and in plants. *Eur. J. Plant Pathol.* 126:129-133.
- Webster, J., and Weber, R. 2007. Pages 127-164 in: *Introduction to Fungi*, 3rd ed. University Press, Cambridge.