

From metagenome to gene

Identification of the first *Synchytrium endobioticum* effector through comparative genomics

B.T.L.H. van de Vossenberg



Propositions

1. The composition of *Synchytrium endobioticum* isolates, which are genetic populations, is not fixed after diversifying selection.
(this thesis)
2. Pathotype grouping concepts that do not identify the full virulence potential of pathogen isolates undermine phytosanitary control measures.
(this thesis)
3. The Great Food Transformation (Willet *et al.*, (2019) *The Lancet*, 10170; 447) is a comprehensive roadmap to host 10 billion people on earth in 2050, but implementation requires a sense of urgency that seems currently lacking.
4. Big data science might not only help with solving the birth of our planetary system, but it allows to study human behaviour under effect of musical genres (Park *et al.* (2019) *Nature Human Behaviour*, 3; 230) beyond merely verifying stereotypes.
5. The only tradition we need to hold on to, is the tradition of change.
6. A board game is a very powerful and simple way to inform the general public about plant health and the associated risks.

Propositions belonging to the PhD thesis, entitled:

“From metagenome to gene; identification of the first *Synchytrium endobioticum* effector through comparative genomics”

Bart T.L.H. van de Vossenber
Wageningen, 1 July 2019

From metagenome to gene

Identification of the first *Synchytrium endobioticum* effector
through comparative genomics

Bartholomeus Theodorus Leonardus Henricus
van de Vossenberg

Thesis committee

Promotor

Prof. Dr Richard G.F. Visser
Professor of Plant Breeding
Wageningen University & Research

Co-promotors

Dr Theo. A.J. van der Lee
Senior scientist, Biointeractions and Plant Health Group
Wageningen University & Research

Dr Jack H. Vossen
Senior scientist, Plant Breeding Group
Wageningen University & Research

Other members

Prof. Dr Thomas Debener, Leibniz University Hannover, Germany
Prof. Dr Francine P.M. Govers, Wageningen University & Research
Dr Kurt Heungens, Institute for Agricultural and Fisheries Research, Belgium
Prof. Dr Martijn Rep, University of Amsterdam

This research was performed under auspices of the graduate school of Experimental Plant Sciences (EPS)

From metagenome to gene

Identification of the first *Synchytrium endobioticum* effector
through comparative genomics

Bartholomeus Theodorus Leonardus Henricus
van de Vossenber

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Monday 1 July 2019
at 4 p.m. in the Aula.

B.T.L.H. van de Vossenberg

From metagenome to gene; identification of the first *Synchytrium endobioticum* effector through comparative genomics

290 pages

PhD thesis, Wageningen University, Wageningen, the Netherlands (2019)

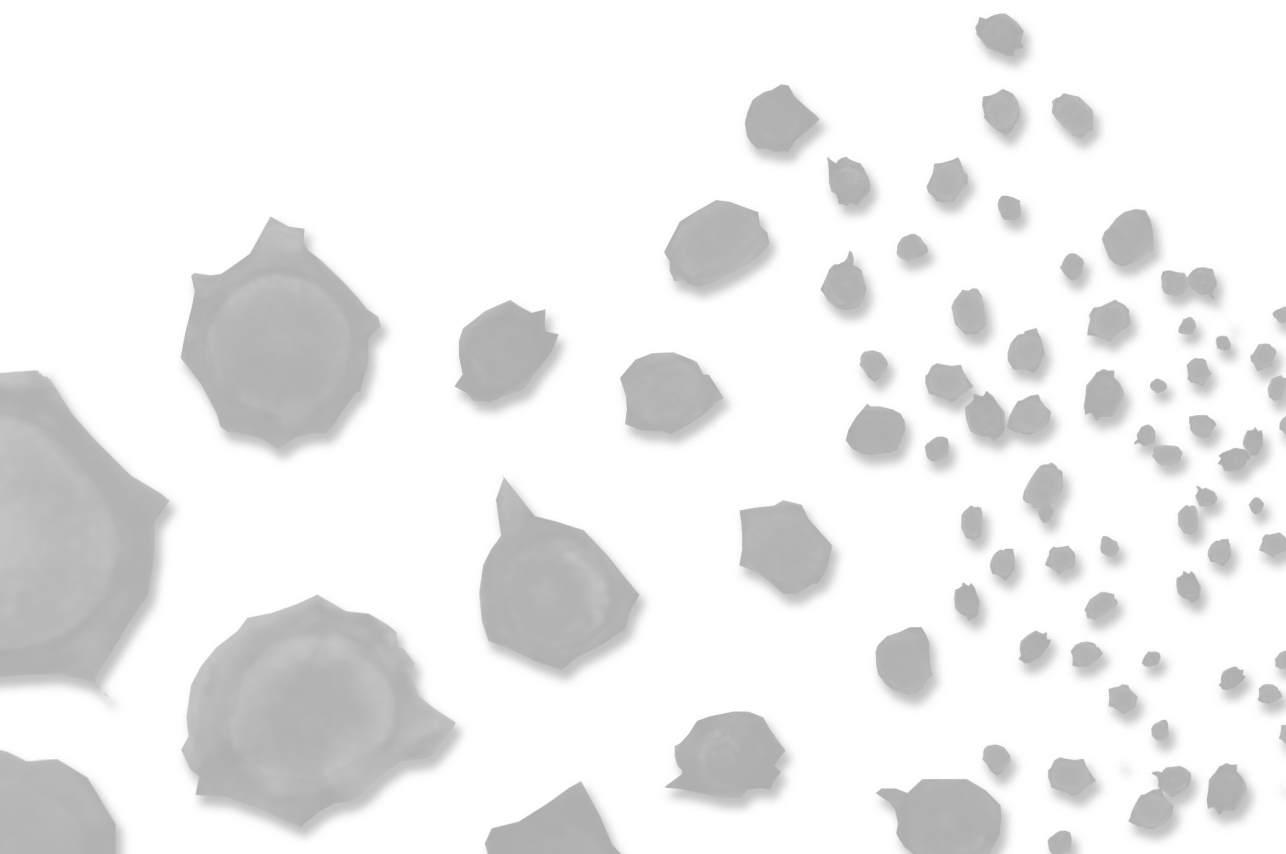
With references, with summary in English

DOI: <https://doi.org/10.18174/476563>

ISBN: 978-94-6343-964-0

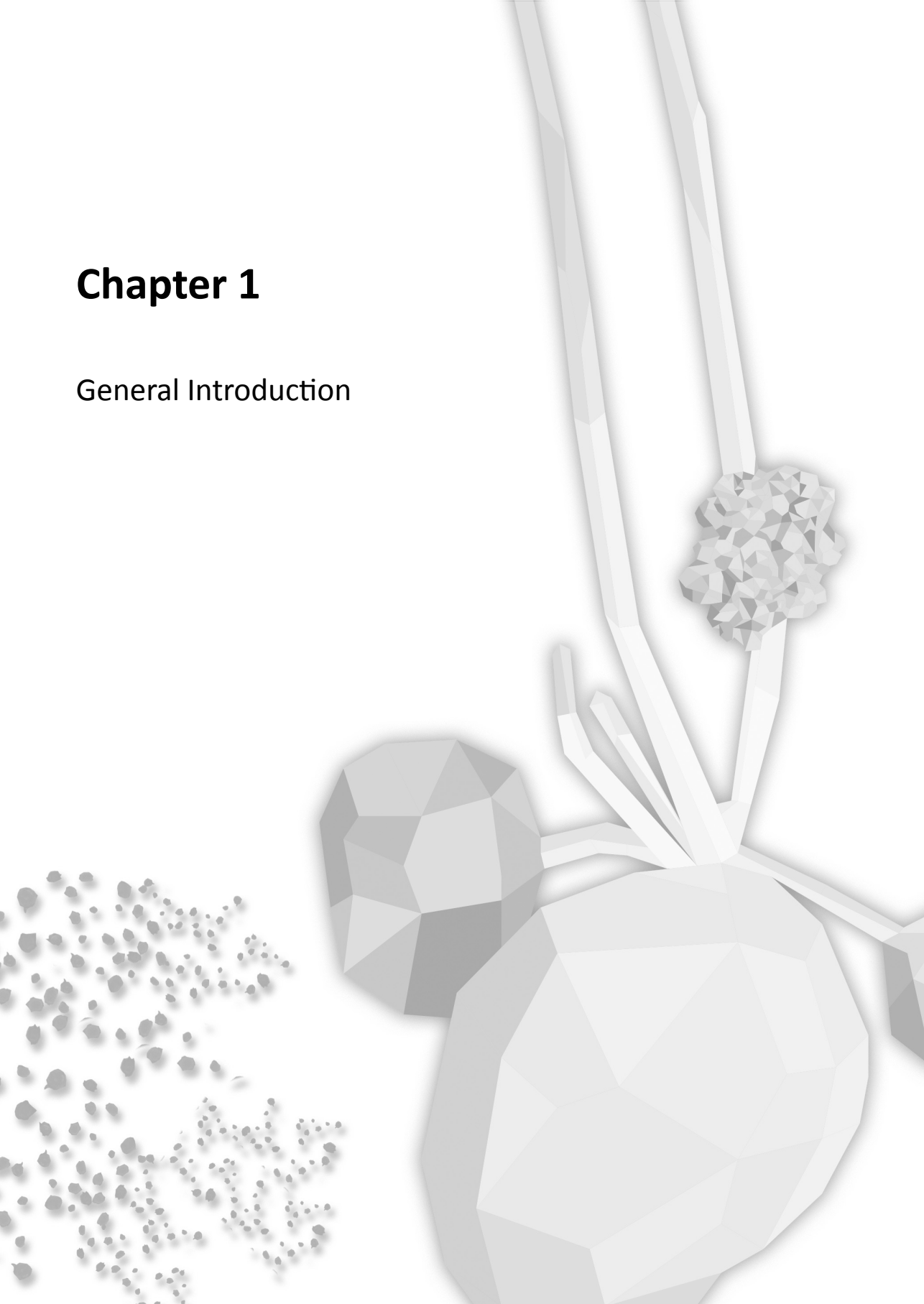
Contents

Chapter 1.	General Introduction	7
Chapter 2.	Comparative genomics of chytrid fungi reveal insights into the obligate biotrophic and pathogenic lifestyle of <i>Synchytrium endobioticum</i>	41
Chapter 3.	The linear mitochondrial genome of the quarantine chytrid <i>Synchytrium endobioticum</i> ; insights into the evolution and recent history of an obligate biotrophic plant pathogen	121
Chapter 4.	An alternative bioassay for <i>Synchytrium endobioticum</i> demonstrates the expression of potato wart resistance in aboveground plant parts	165
Chapter 5.	The <i>Synchytrium endobioticum</i> AvrSen1 triggers a Hypersensitive Response in Sen1 potatoes while natural variants evade detection	203
Chapter 6.	General Discussion	241
	Summary	263
	Acknowledgements	269
	About the author	275
	Publications	279
	Training and education statement	285



Chapter 1

General Introduction



Chapter 1

Introduction

In 1887, professor Károly Schilberszky of the Budapest University had a public appeal published in the Austro-Hungarian weekly agricultural newspaper *Gazdasági Lapok* to send in potatoes infected with powdery scab. Although he did not succeed in getting powdery scab infected potatoes, he did get diseased potato tubers with unusual symptoms late 1888 from Mr. Franz von Jattka. These tubers came from two potato fields in Hornyán (nowadays Horňany in the Slovak Republic), and he later learned that the entire yield of these fields was affected by the disease. Initial studies were without success, but with new infected tubers sent in 1896, Schilberszky could study the pathogen and produce a description of the causal agent of potato wart disease: *Synchytrium endobioticum* (Schilberszky) Percival [1, 2].

Plant pathogens and invertebrates harmful to plants or plant products can have great social and economic impact, and are a continuous threat to food security [3]. One of the best known examples of the impact plant pests can have is the great Irish potato famine (1845 - 1849) which was caused by the oomycete *Phytophthora infestans*. In two consecutive years, potato crops were destroyed resulting in a famine in which 1.5 million people died, and another million emigrated from the island [4]. Human activity has been tightly linked to invasions of plant pests [5], and this is also believed to be the case for *S. endobioticum*. It is now generally assumed that in the aftermath of the Irish potato famine, *S. endobioticum* was brought into Europe with the import of new seed potatoes [6]. Currently the obligate biotrophic Chytridiomycota (chytrid) fungus is regarded one of the most important quarantine pests for cultivated potato worldwide [6, 7], and is included on the United States Department of Agriculture (USDA) and Department of Health and Human Services (HHS) Select Agent list [8].

Much of our current knowledge on potato wart disease, its causal agent and the discovery of resistant potato varieties is the result of research performed in the first decades of the 20th century. With the discovery of new pathotypes, studies from the mid-1900s focused on bioassays for pathotype identification, selection of new resistant breeding material, and many new pathotypes were reported. In the last quarter of the twentieth century, different aspects of the pathogenesis of *S. endobioticum* were studied and electron microscopy was used to verify early observations. Finally, in the last two decades, potato wart disease research shifted more towards (molecular) diagnostics.

Chapter 1

In this introduction, over a century of potato wart research is reviewed, placing our current knowledge in historic perspective. Also, general aspects of molecular plant-pathogen interactions are introduced and a model for the molecular interaction between *S. endobioticum* and its host is proposed. Finally, an overview of the thesis chapters is presented.

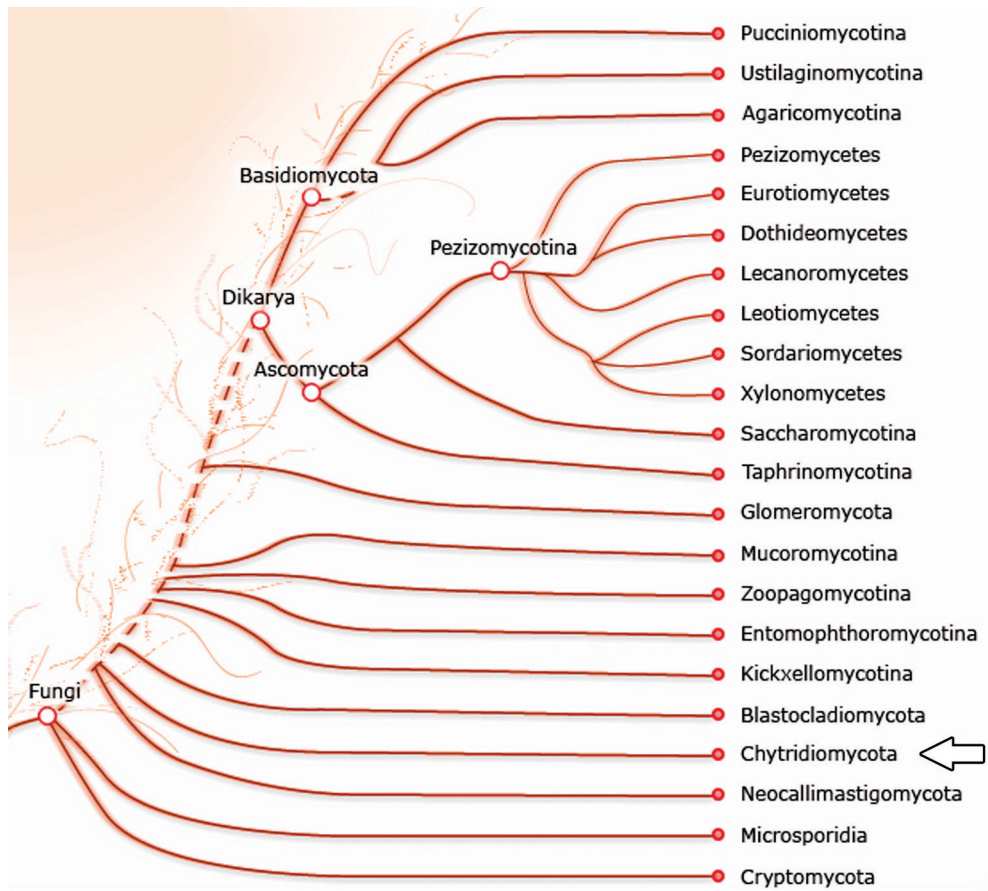


Figure 1. Fungal tree of life. Phylogenetic relationship of the major fungal phyla as treated by Grigoriev *et al.* [25]. This representation is supported by the genome-wide proteome tree of the fungal kingdom as determined by Choi and Kim [26], except for Cryptomycota which are treated as sister-clade to fungi by Grigoriev *et al.* but cluster as a basal lineage in fungi in the analysis of Choi and Kim. Chytridiomycota (arrow) is one of the basal lineages in the fungal kingdom.

Ethiology and taxonomy

S. endobioticum belongs to Chytridiomycota which is a basal fungal lineage that arose 1,000 to 1,600 million years ago [9]. The basal position of Chytridiomycota in the fungal kingdom is illustrated in Figure 1. This ancient lineage of true fungi is characterized with motile flagellated spores, which are called zoospores, and an absence of hyphae or mycelium. Chytrids inhabit aquatic and moist terrestrial environments, and are described as free-living saprophytes or as species being pathogens to plants, algae or invertebrates [10, 11]. Examples of the latter are *Batrachochytrium dendrobatidis* and *Batrachochytrium salamandrivorans* that have been causing mass mortalities in amphibian populations [12].

In the original description by Schilberszky, the *S. endobioticum* was placed in a newly erected genus in the order Chytridiales and was first named *Chrysophylyctis endobiotica* [1]. Fourteen years after the first description, Percival proposed to transfer the species to the genus *Synchytrium* based on the morphological and cytological resemblance to other members in the genus, and renamed it *S. endobioticum* [13]. Additional microscopic studies performed by Curtis [14] later supported the transfer of the species to this genus.

The genus *Synchytrium*, erected in 1865 with the obligate biotrophic *Synchytrium taraxaci* pathogen on dandelion as type species, contains approximately 200 species from aquatic and terrestrial habitats. Previously, all *Synchytrium* spp. were believed to be parasitic species for algae, mosses, ferns and flowering plants [15]. This view changed with the recent description of the first saprobic member of the genus, i.e. *Synchytrium microbalum* JEL517, which was isolated from an acidic pond in Hancock County, Maine, USA [16]. Species within the genus are assigned to one of six subgenera based on variations in the life cycle and morphological features. Karling (1964) placed *S. endobioticum* in the subgenus *Mesochytrium* based on the observations of Curtis that resting spores give rise to zoospores directly. However, with the work of Kole [17], which was later confirmed by Sharma and Cammack, Lange and Olson and Hampson [18-20], evidence was provided that resting spores function as a prosorus, i.e. a vesicle in which sporangia are formed that hold the zoospores. These observations supported the transfer of *S. endobioticum* to the subgenus *Microsynchytrium*.

There is strong support for *Synchytrium* being a monophyletic clade within the division Chytridiomycota from ribosomal DNA (rDNA) based phylogenies [21, 22]. Species of the genus *Synchytrium* are currently classified in the order Chytridiales. However, this placement was questioned following light and electron microscopy studies of the zoospore ultrastructure

[23, 24]. The molecular phylogeny of chytrid species based on the rDNA operon (18S, 28S and 5.8S) [21] showed that the order Chytridiales is polyphyletic, but 18S rDNA sequences could not confirm or refute *Synchytrium* within the order Chytridiales [22].

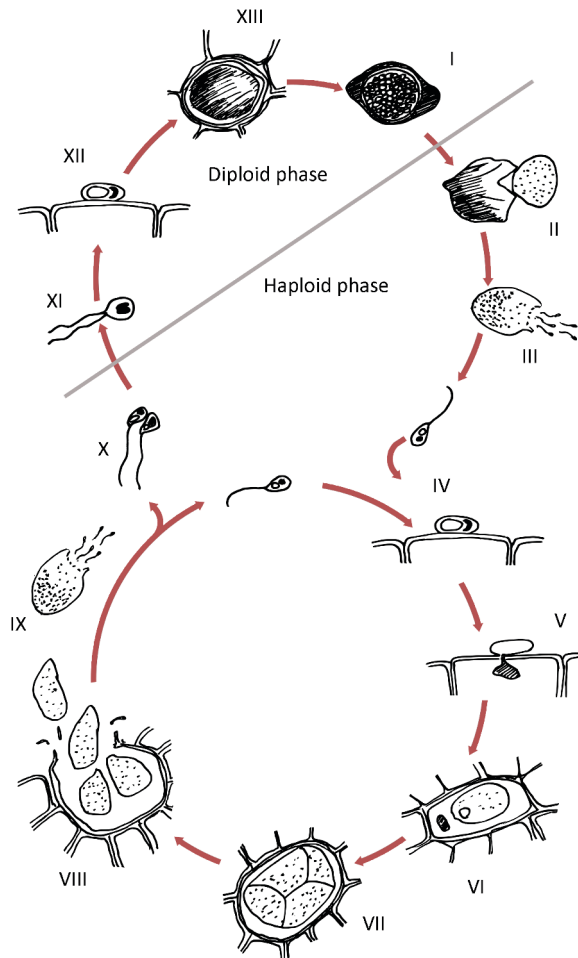


Figure 2. Life cycle of *S. endobioticum*. The long-cycled life cycle of *S. endobioticum* is based on the observations of Curtis [14] and Kole [17] with: **I** resting spores, **II** germination of resting spore and release of sorus, **III** zoospore release, **IV** encystment and **V** penetration of host cell, **VI** developing fungal thallus or prosorus, **VII** sporangium with multiple zoosporangia: summer spore, **VIII** rupture of host cell and release of zoosporangia, **IX** zoospore release from sorus, **X** plasmogamy: fusion of isogametes, **XI** karyogamy: zygote (*sensu* Curtis), **XII** encystment and infection of host by zygote, and **XIII** development of resting sporangium. When and if karyogamy and meiosis take place have been subject of debate, and different authors studying the subject have not yet been able to provide a conclusive answer.

Life cycle and pathogenesis

The species of *Synchytrium* vary considerably in the complexity of their life cycles, and *S. endobioticum* belongs to the long-cycled species [15]. The life cycle of *S. endobioticum* was described in detail by Curtis in 1921 [14], and her observations are largely unchallenged today (Fig. 2).

During the resting period, while the warted host tissue decays and resting sporangia are released in the soil, a new inner wall layer is formed inside the resting spore. Upon germination, the outer spore wall ruptures and the newly formed sorus is released which gives rise to numerous (200 - 300) uninucleate haploid zoospores [19]. These zoospores are approximately 3 μm in diameter with a single whiplash flagellum that is 17 μm long [27]. Released zoospores encyst on the host cell wall within an hour, after which the fungal thallus penetrates the host cell wall leaving the cyst wall outside the host cell [28]. It is not known if the fungal thallus is invaginated by the host plasma membrane, similar to haustoria in rust fungi and powdery mildew [29], or if it resides directly in the host cytoplasm. An individual host cell can be infected by a single or multiple zoospores [30]. The host nucleus enlarges, becomes irregular in shape, and is often found closely appressed to the developing fungal thallus. The fungal body from the time of entry into the host cell to that of the cleavage into sporangia is called prosorus [14]. Numerous smaller daughter nuclei are formed followed by the separation of the sporangium (i.e. summer spore) in 3 to 7 zoosporangia which are covered by a single wall [30]. Curtis reported that during this repeated mitosis five minute chromosomes could be observed [14]. At this stage, brown discoloration of the host tissue can be observed with light microscopy. Hypertrophic growth is induced in cells surrounding the infected host cell. Simultaneously, hyperplasia increases the amount of meristematic tissue which enhances the chance of re-infection of the host [6]. Enlarged host cells burst open and release the individual zoosporangia, which in turn release the zoospores.

These zoospores can re-infect the host tissue and start the disease cycle anew giving rise to new zoosporangia. The zoospores may also act as isogametes, fusing to form a diploid biflagellate zygote (*sensu* Curtis) which can infect the host in a similar fashion as the haploid zoospores [14]. However, instead of forming zoosporangia, host cells infected with the zygotes give rise to the thick-walled resting sporangia. The resting spore is sometimes referred to as winter spore, measures 35 to 80 μm in diameter, and can remain viable and infectious in infected soils at depths of 50 cm for decades [31]. The thick outer wall of the resting spore

Chapter 1

is made up of a chitin-protein complex and consist of the host's dead protoplasm and cell wall which is united with the fungal wall to create a firm composite structure [32, 33]. Even under favorable conditions, only small percentages (~10% at maximum) of sporangia are induced to germinate [17]. Where infections with zoospores induce hypertrophic growth and hyperplasia in the host, this is not the case for host cells infected with zygotes [15].

Curtis described the occurrence of karyogamy after fusion of the isogametes and meiotic divisions before resting spore germination, but this has been subject of debate [15]. It remains elusive if the *S. endobioticum* lifecycle reflects a sexual or a parasexual cycle. In both cases isogametes fuse (plasmogamy) after which the nuclei fuse (karyogamy) to form a diploid nucleus. Where a sexual cycle is highly coordinated and involves meiotic divisions, parasexual cycles are uncoordinated and do not involve meiosis but mitotic recombination and haploidization [34]. Electron microscopy studies have not been able to provide an ultrastructural description of gamete fusion or meiotic divisions, and sexual recombination remains elusive [30].

In resistant potato varieties, zoospores infect the host cell equally well as in susceptible varieties. However, the pathogen dies within a few hours after infection as a result of pathogen-induced defense necrosis [15, 35].

Occurrence of new pathotypes and nomenclature

Potato varieties resistant to potato wart disease have been reported as early as 1909, and by 1919 immunity trials performed in the United Kingdom, Germany and the United States resulted in the publication of a list with 64 potato varieties that were resistant to the disease [36]. In 1931, Köhler, stipulated that the pathogen has a uniform make-up and racial differentiation was not known for the pathogen [6]. However, in the 1930s, resistant potato varieties from the United Kingdom grown in Canada were found to be very susceptible to the local strain of *S. endobioticum* [37], and in the decades afterwards it became clear that different pathotypes of *S. endobioticum* exist. The initial *S. endobioticum* pathotype is nowadays known as pathotype 1(D1).

It was not until 1941 that scientists described wart development on formerly resistant potato varieties in Gießübel and Silberhütte [38, 39]. The pathotypes identified in these towns are nowadays known as 2(G1) and 3(SB) respectively [40]. In the following years new *S. endobioticum* pathotypes were defined in Germany, Poland, the Czech Republic, the

Ukraine and Canada, based on their response on different differential panels with resistant and susceptible potato varieties. Among these were the German pathotypes 6(O1) (found in Olpe, 1952), 8(F1) (Fulda, 1954), and 18(T1) (Trannroda, 1978) [40]. Pathotypes 1(D1), 2(G1), 6(O1), 8(F1) and 18(T1) are regarded to be most widespread and of major importance [41]. The latest two formally described pathotypes are 38(N1) from Nevsehir (Turkey, 2009) and 39(P1) from Piekienik (Poland, 2015) [42, 43]. Although not officially published in a peer-reviewed journal, pathotype 40(BN1) has been reported from Poland [44]. Collectively non-pathotype 1(D1) isolates are referred to as “higher pathotypes”. However, this does not imply that higher pathotypes descended from pathotype 1(D1).

After the end of the second world war, and shortly after the report of pathotype 2(G1), the iron curtain ascended in Europe. On each side of the iron curtain different naming conventions were implemented to identify the new pathotypes. In the east, new pathotypes were named after the locality where it was found, but in the west sequential Arabic numbers were used to define new pathotypes. As a consequence, the new pathotypes found in Gießübel and Silberhütte were named G1 and SB in the east, and 2 and 3 in the west respectively. Baayen *et al.* (2006) proposed a harmonization of the pathotype naming system, which is still used today. Under this system both naming systems are combined, for instance 2(G1) for the pathotype identified in Gießübel [40].

Disease symptoms and host plants

The majority of *Synchytrium* species induce distinct local cellular reactions, and because of it they are sometimes referred to as gall-inducing chytrids. The formation of galls or warts is the result of cell enlargement (hypertrophy), cell multiplication (hyperplasia) or a combination of both. Where often the host is not seriously affected by infection with *Synchytrium* spp., this is not the case for *S. endobioticum* [15]. The typical malformation caused by *S. endobioticum* has initially inspired the use of various names to describe the disease, such as potato tumor, potato cancer, black scab, black wart, and cauliflower disease [6]. Since 1908 the term “wart disease” has come into common use and has been used in official documents since then [36]. Disease symptoms on potato tubers are characterized by the warty cauliflower-like malformations induced by the pathogen (Fig. 3A). Percival (1910) noted that these warts can vary strongly in size, ranging from “the size of a small pie to the size of a hen’s egg”, and could be as large or even larger than the tuber from which they emerged [13]. Schilbersky’s original description was based on infected tubers,

but meristematic tissue on stolons, eyes, sprouts, and inflorescences can be infected while the root system seems to be immune [15, 28]. Where the warts formed on belowground plant parts are compact protuberances, malformations on aboveground affected plants parts seem to follow the morphology of the infected plant organ (Fig. 3B). Belowground warts are white to brown, whereas aboveground malformed tissues are green. Both types of wart turn black upon decay of the host tissue which coincides with enhanced resting spore formation. Decomposition of the warted host tissue allows release of the spores into the surrounding soil [13].

Hosts that can be infected with *S. endobioticum* are not restricted to susceptible *Solanum tuberosum* varieties. Other Solanaceous species such as *Solanum lycopersicum* (tomato), *Solanum dulcamara* (bittersweet), *Solanum nigrum* (black nightshade) have been reported as alternative hosts under experimental conditions [45, 46]. In *S. dulcamara* and *S. nigrum*, very little to no hypertrophy was observed, but the sporangia of *S. endobioticum* were found to be present in the infected tissue. Cotton (1916) hypothesized that these alternative hosts, which are found frequently bordering plots with potato cultivation, could act as “carriers” of the disease promoting the survival of the pathogen when only resistant potato varieties are cultivated [45].

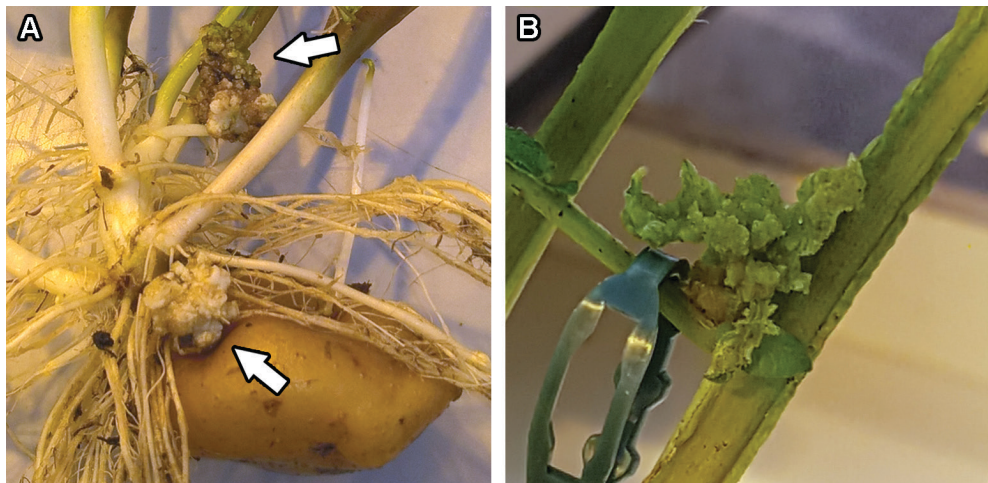


Figure 3. Potato wart disease symptoms on aboveground and belowground plant parts. Subterranean warts are ethiolated (A, lower arrow), whereas the warted tissue exposed to light turns green (A, top arrow). Wart formation on aboveground plant parts (B) is green and largely follows the morphology of the infected plant organ.

Martin (1923) studied the response to *S. endobioticum* in additional alternative hosts and observed warted outgrowth on *Solanum nodiflorum* (syn. *Solanum americanum*, American nightshade) and *Solanum villosum* (hairy nightshade). In *Nicandra physalodes* (apple-of-Peru), no disease symptoms were observed but *S. endobioticum* was found to be present in various life stages [47]. More recent, Hampson (1976) tested the reaction of 123 tomato cultivars to *S. endobioticum* pathotypes 2(G1) and 8(F1) and all of the tested combinations resulted in infection of the host [46, 48]. Where in potato the root system is not affected by the pathogen, roots are attacked in tomato [49]. Similar to *S. dulcamara* and *S. nigrum*, no hypertrophic growth of the infected tissue is observed in *S. lycopersicum* [46]. Other hosts that were found susceptible to *S. endobioticum*, on account of spore formation, are *Solanum pseudocapsicum* (Jerusalem Cherry), *Physalis franchetii* (Chinese lantern), *Schizanthus* sp. (butterfly flower), *Datura metel* var. *fastuosa* (Angel's Trumpet) [48]. Occurrence of *S. endobioticum* on wild Solanaceous species has been reported from Mexico, but this report has not been confirmed by the Mexican authorities [7]. The role of the alternative hosts to act as reservoirs for *S. endobioticum* and their potential to extend the longevity of the pathogen under field conditions was not studied.

Distribution and dispersal

Currently, *S. endobioticum* has been reported from all continents where potato is cultivated [50] (Fig. 4). Potato wart disease is believed to originate from the Andean region, where it co-evolved with its Solanaceous hosts [51], and early Peruvian art objects of potatoes have been suggested to depict wart-like malformations [6]. The disease was most likely introduced in Europe after the Irish potato famine (1845 - 1849) when growers were in need of new seed potatoes [6].

The first description of the pathogen in Europe by Schilzbersky was based on infected plants that were grown locally from seed tubers imported from England. The first official report from the United Kingdom dates from 1901, but based on oral reports the pathogen was believed to be present as early as 1876 [36]. Hampson regarded spread from the United Kingdom to the European main land, Canada and (via Ireland) to South Africa and New Zealand the most likely scenario [6]. In the United Kingdom, the losses as a result of the disease were very serious, and in some localities potato cultivation had to be practically abandoned. With the discovery of resistant potato varieties in the early 1900s, potato cultivation in these areas was resumed, and the demand for high-class resistant seed

potatoes far exceeded the supply [45, 52]. In the decades after description of the pathogen, disease reports came in from different countries describing the first occurrence of potato wart disease on their territory: e.g. Germany (1908), Ireland (1908) Canada (1909), Sweden (1912), the Netherlands (1915), Czech Republic (1915), Poland (1917) [6, 7, 36, 40, 50]. In 1952, potato wart disease was introduced in India from Danish seed potatoes [53].

Both world wars played a role in spreading the disease across the European continent. At the end of the first world war, there was a shortage of ware and seed potatoes, and potatoes from infected areas in the United Kingdom were used as seed in districts that were untouched by the disease [36]. Similarly, potato wart disease is believed to be introduced in the former Union of Soviet Socialist Republic (USSR) with the movement of German troops at the onset of the second world war [54]. After the second world war, trade with infected seed potatoes has often been suggested as the cause of spread to other countries. More recently, the disease was reported for the first time in Turkey (2003) [55], Bulgaria (2004) [56], and Georgia (2009) [57]. Also, recently the disease re-occurred in Denmark (2014) where it was believed to be eradicated since 1989 [58].

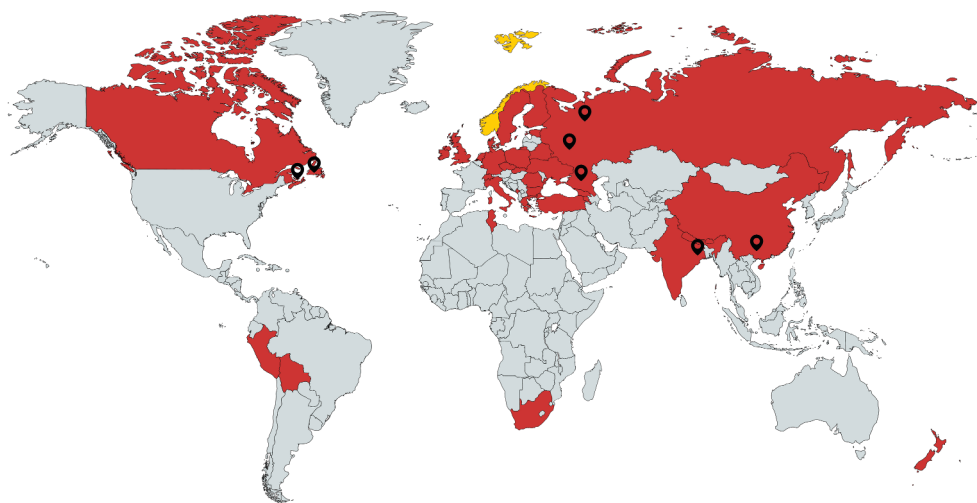


Figure 4. World-wide distribution of *S. endobioticum*. Countries from which the potato wart disease pathogen have been reported to be present [50] are shown in red. In several countries, *S. endobioticum* has a limited distribution compared to the geographical area they represent (black markers): i.e. in Canada the distribution is restricted to Newfoundland and Prince Edward Island; in the Russian Federation it is limited to the Central, North-western and Southern federal districts; the province of West Bengal in India; and the province of Guizhou in the People's Republic of China. In Norway, the pathogen status is "transient, under eradication" (orange).

S. endobioticum has limited capacity to spread by natural means. Zoospores are too short-lived to have any significant contribution to dissemination, and have been suggested to be able to migrate a distance of up to 5 cm on their own means [59]. Dispersal of resting spores by water currents [28] and wind [60, 61] have been reported, accounting for local spread. Typically, the pathogen is present in only restricted parts of the infested fields. Spread of the disease by human activity was recognized early on, and the movement of infected seed potatoes, soil and manure were regarded to have the biggest contribution to local and international dissemination of the pathogen [36]. Hampson [6] referred to partial resistant varieties that could carry inconspicuous warts with resting spores are particularly liable to disseminate the pathogen. Also, the floor area of cars, vans and trucks being operated in Newfoundland (Canada), where the pathogen is widespread, were demonstrated to harbor infectious resting spores [62, 63], and the spread of the disease was related to transportation lines such as, ports, roads and railways [6].

Another pathway for the dispersal of *S. endobioticum* is organic waste from potato processing industries, when infected tubers are processed. Organic waste has to be sanitized, for instance by composting or pasteurization, before it can be used as fertilizer in agriculture, horticulture or in private gardens [64]. However, while these techniques have been proven effective on the potato cyst nematode *Globodera rostochiensis*, they have little to no effect on the resting sporangia of *S. endobioticum* [65]. Even more, compostation of infected potato material is part of the Spieckermann bioassay to prepare the inoculum [66].

Potato wart disease in the Netherlands

The first report of Potato wart disease in the Netherlands dates from 1915, and was based on the finding of diseased potatoes in a private garden in the city of Winschoten in the province of Groningen (Fig. 5). Occurrence of the disease was reported by Mr. Uil who was a teacher at an agriculture school [67, 68]. How the pathogen got introduced into the country is not known, but the disease was believed to be present in Winschoten at least 8 years before the first official report. The infected potatoes were grown in a poor part of the city, just outside the city limits, where people often relied on the produce from their own gardens as food supply. Further investigation showed that the pathogen occurred in several gardens in the vicinity of the first finding, but also in two neighboring villages. In following years, potato wart was found in a small number of municipalities in a 15 kilometer radius that contained the first outbreak location. In 1922 however, numerous new outbreaks were detected in the

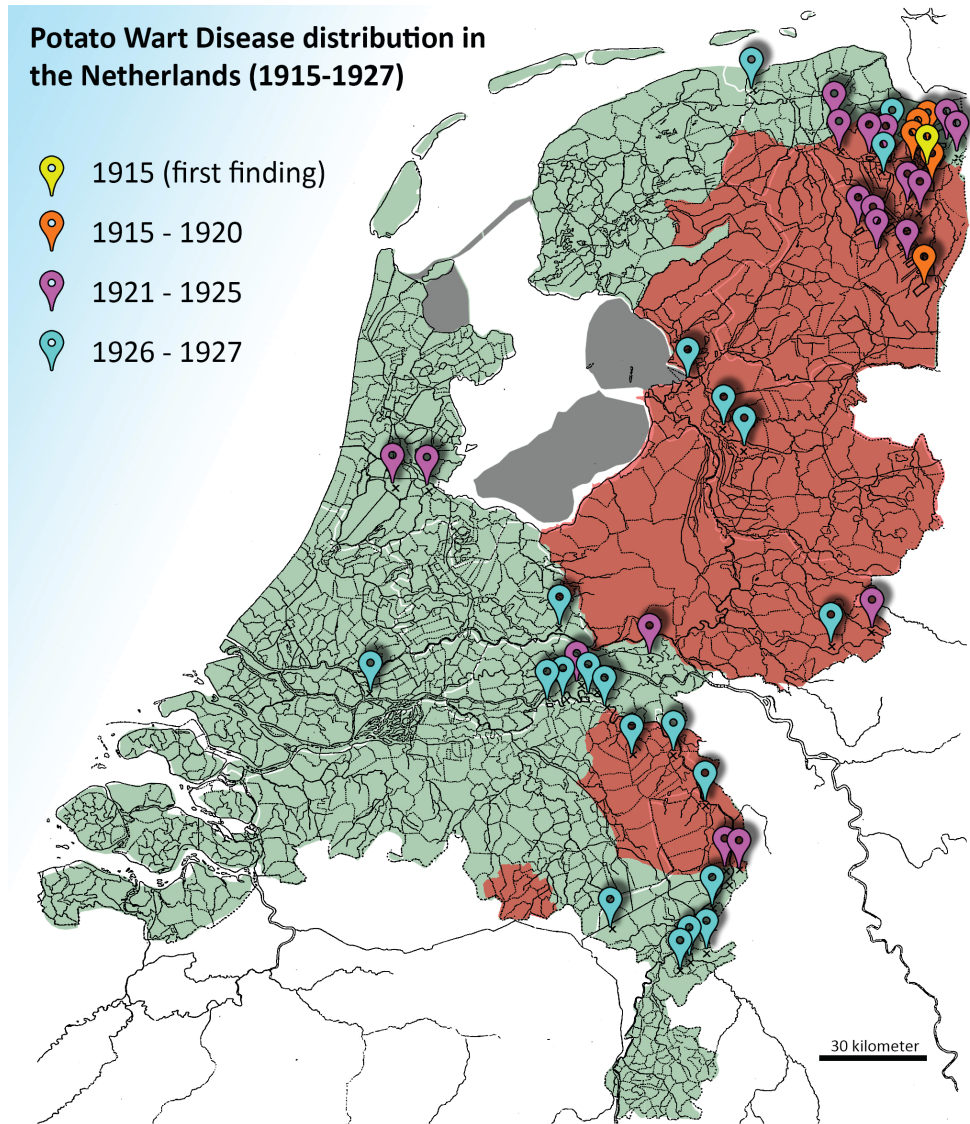


Figure 5. Historical map of the Netherlands (1928) with the distribution of potato wart disease in the period 1915 - 1927. Occurrence of *S. endobioticum* is based on [67]. Land mass as it existed in 1928 is shown in green, and reclaimed land is shown in grey. Markers are shown for the municipalities in which potato wart disease outbreaks were reported in the twelve years after the first finding. The markers' color reflect the period in which the disease was reported. Areas in which increased vigilance for potato wart disease occurrence is currently enforced by the Dutch authorities are shown in red.

north-eastern part of the Netherland, and also the first findings in the provinces Gelderland and Limburg were reported. The rapid expansion of outbreak locations in this year was attributed to the rainy weather in July and August of that year [67]. In the years 1926 and 1927, another rapid expansion of the pathogen was observed following the course of the river Meuse in the provinces Limburg, Noord-Brabant and Gelderland, and to less extent following the flow of the river IJssel. This spread followed unprecedented discharges of the rivers Rhine and Meuse which resulted in the breaking of dykes and subsequent flooding of large polder areas along the rivers IJssel and Meuse [69]. Based on the time and location of the reported cases in flooded areas it now seems likely that resting spores from previously infested fields upstream of the river Meuse were transported to downstream areas where they settled on soils used for potato cultivation. Weis [28] noted that periodic flooding of potato fields followed by drying and aeration were favorable conditions for disease occurrence. In many instances the source of the outbreaks could not be traced, and only in some cases where fields were used for potato cultivation for the first time, seed potatoes originating from infested areas were considered to be the source of infection [67].

Strict phytosanitary measures were implemented prohibiting the cultivation of susceptible potato cultivars in the infested regions, and by 1970 potato wart disease seemed to have been eradicated from the Netherlands [40]. In 1973 however, the disease was rediscovered in the province of Groningen on a potato variety known to be resistant to pathotype 1(D1). The isolate was identified as pathotype 2(G1) and marks the first non-pathotype 1(D1) isolate that was detected in the Netherlands. Short after, additional pathotype 2(G1) outbreaks were reported from the provinces Groningen and Drenthe, including several municipalities that were among the first ones to be infected in the beginning of the 20th century. In 1990 pathotype 6(O1) was found to be present in the north-eastern part of the Netherlands, followed by the first finding of pathotype 18(T1) in the same region in 2001. Also in the south of the Netherlands, *S. endobioticum* infected potatoes reoccurred after 1990. In contrast to the north-eastern parts of the country, only pathotype 1(D1) isolates of the pathogen were found in the southern parts of the Netherlands [40]. The latest expansion of infected areas in the Netherlands is represented by an infested field in Luyksgestel, close to the Belgian border, which was found in 2011.

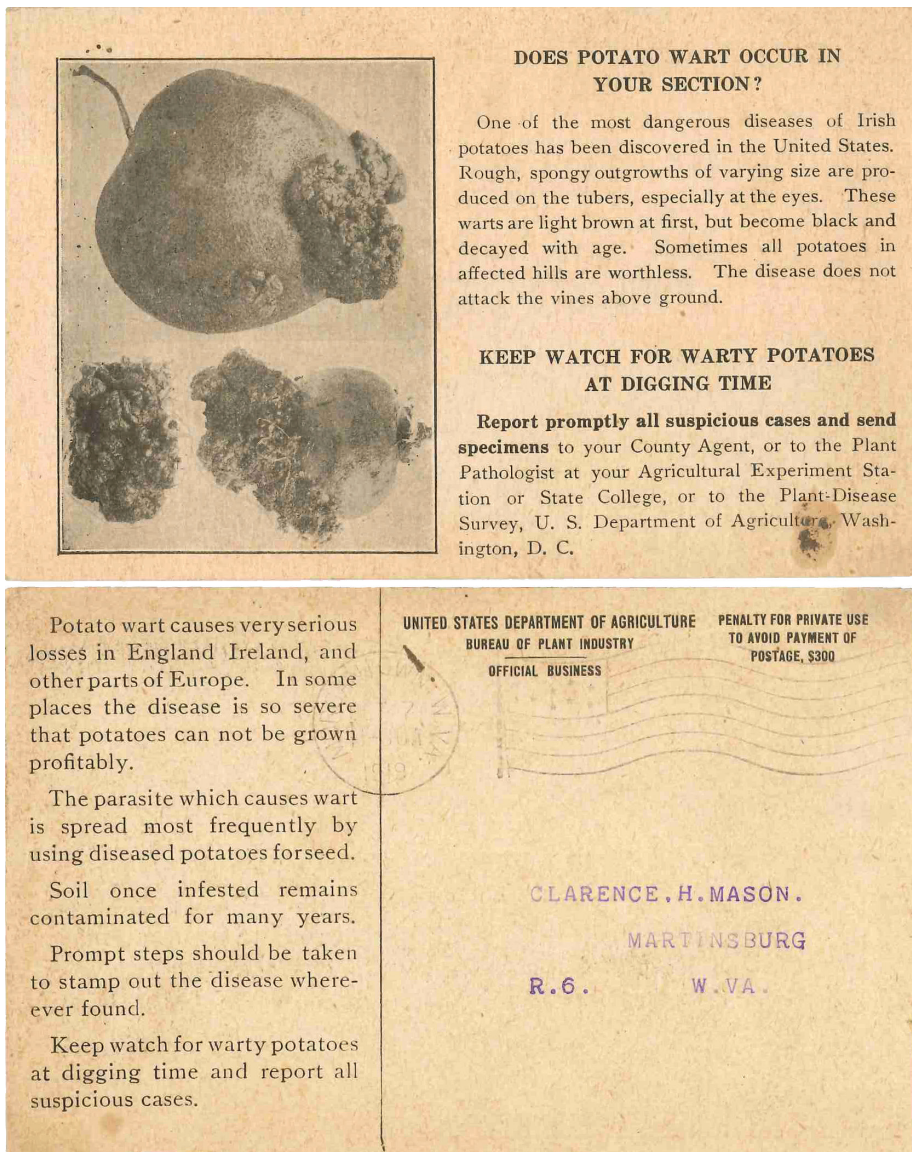


Figure 6. Postcard issued by the United States department of agriculture (USDA, 1919). Since the initial findings of *S. endobioticum* infestations, local and national authorities used different strategies to inform the public about the dangers associated with potato wart disease.

Control and phytosanitary measures

An impressive number of over 600 different treatments including application of more than 120 control agents have been used with the aim to reduce or eradicate the pathogen from infested soils [70]. These treatments were unsuccessful in most cases, and if they did have success to eradicate the pathogen, the control agents used were so toxic that they rendered the treated soils barren. Research for chemical treatments were largely abandoned from the 1960s as phytosanitary measures had proven themselves to be effective to minimize the spread of the disease [70]. In the 1990s, the use of crushed crabshell was shown to suppress wart disease in infested fields [71]. Also, crop rotation and intercropping potato with other crops like maize were reported to have a positive effect on the reduction of viable resting spores in soils [7]. Informing farmers and home-growers about the risks of the disease was regarded of utmost importance, and since the rapid spread of the disease in the early 1900s, national and local authorities started to produce information leaflets and postcards (Fig. 6).

The lack of chemical control agents, the impact of the pathogen (complete loss of tuber yields from infected field have been reported), and the longevity of resting spores have led to the introduction of pathogen and pest acts worldwide. In the European Union, strict phytosanitary measures are in place to prevent the introduction and spread of the pathogen [72]. The use of resistant potato varieties is central in these measures. Fields in which the pathogen is found are placed under official control (i.e. “scheduled”). No potato production or the production of plants for planting is allowed for a period of twenty years. A safety zone, or buffer zone, is established surrounding the infested plot in which cultivation of only resistant potato varieties is allowed. In addition, countries have to publish on resistant potato varieties and incidence of the pathogen. The European and Mediterranean Plant Protection Organization (EPPO) created several documents describing monitoring programs, measures to contain the pest and to prevent further spread, and procedures for de-scheduling plots previously infested by *S. endobioticum* [73, 74].

In the Netherlands, phytosanitary measures are enforced by the National Plant Protection Organization (NPPO) which is part of the Netherlands food and consumer product safety authority (NVWA). The NVWA has established large zones of increased vigilance around scheduled plots and buffer zones to prevent the spread of the pathogen (Fig. 5). In these areas cultivation of only slightly susceptible or resistant potato varieties are allowed [75].

Diagnosics and pathotype identification

Reliable detection of the pathogen in soils and grouping of *S. endobioticum* isolates into pathotypes is of major importance and fundamental to phytosanitary measures to restrict the spread of the disease [41]. Pathotype grouping is performed with tuber-based bioassay described by Spieckermann [76] or Glynne and Lemmerzahl [77-79] in combination with a standardized set of differential potato varieties. In Spieckermann assays, compost containing resting spores is used as inoculum, whereas in the Glynne-Lemmerzahl assays, fresh wart cuttings, which mainly contain summer sporangia, serve as inoculum. Even though the Spieckermann and Glynne-Lemmerzahl bioassays are routinely applied for pathotyping and resistance screening, these assays remain challenging and even after several attempts to harmonize assays and scoring, differences in the interpretation of phenotypes and potato varieties used hamper comparison of the results obtained by different institutes [40].

With the advances in molecular biology, the reliability and reproducibility of pathogen detection in host materials and soils were increased. The first (conventional) PCR detection test for *S. endobioticum* was described in 2004 by Niepold and Stachewiz [80], who amplified the partial fungal ribosomal DNA from infected potato tissue using the generic primer ITS4 combined with a species specific forward primer. Shortly after, another conventional PCR was described which was used to detect resting spores in soils [81]. In addition, real-time PCR tests based on the ITS1, ITS2 and 18S small ribosomal subunit were described offering increased sensitivity for detection of the pathogen in soils and the host matrix [22, 82]. There are currently no molecular tests available that allow identification of the different pathotypes. However, a real-time PCR test was described by Bonants *et al.* targeting a pathotype 1(D1) associated single nucleotide polymorphism (SNP), which was identified from a Complexity Reduction of Polymorphic Sequences (CRoPS™) analysis [83]. This test is based on association mapping and is the first attempt of molecular *S. endobioticum* pathotype 1(D1) identification using real-time PCR. Validation of this test showed that all isolates that contain this SNP are pathotype 1(D1), however, the associated SNP is absent in some isolates that produce a pathotype 1(D1) phenotype with the tuber-based bioassays. These molecular tests, together with information on pathotyping with the Glynne-Lemmerzahl bioassay are presented in the EPPO diagnostic standard for *S. endobioticum* [41]. Recently, international collaborative projects were organized to further harmonize methods for pathotyping and molecular species detection, and to generate validation data in an inter-laboratory setting [84, 85].

Genomics

With the advances in whole-genome sequencing, new opportunities arise for studies of the non-culturable obligate biotrophic *S. endobioticum*. Gagnon *et al.* [86] developed microsatellite markers based on small simple repeats (SSRs) mined from the *S. endobioticum* MB42 pathotype 1(D1) genome to study the intra-species genetic variation of 22 isolates. With the exception of most of the Canadian isolates in their study, clustering of the *S. endobioticum* isolates did not reflect geographic origin or pathotype identity [86]. In a separate study, SSR markers resulted in clustering of *S. endobioticum* isolates in three main groups which separated pathotypes 8(F1) and 18(T1) from other pathotypes [87]. Low levels of sequence diversity between *S. endobioticum* isolates representing different pathotypes is a common feature that was noted by several authors working with genome-wide sequence data [83, 86, 87]. Even though these authors exploited genome-wide sequence data to address different research questions, no genome assemblies with structural and functional annotations for *S. endobioticum* were available before the work presented in this thesis.

In an inventory from 2017, more than 3,000 fungi were in completed or ongoing genome projects and the largest category of sequenced genomes constituted pathogenic fungi and fungi of medical importance. Almost half of those were plant pathogens. In contrast, only five genomes of the entire chytrid phylum had been sequenced despite their ubiquitous nature [88].

Functionally annotated genome sequences enable the *in silico* analysis of molecular features underlying fungal lifestyles (e.g. biotrophic, necrotrophic, hemibiotrophic), pathogenicity, genome plasticity, and secondary metabolites. For many pathosystems, classical protein-based assays were unsuccessful in the identification of virulence factors (effectors), but genome-wide sequencing rapidly changed the playing field [89]. Effector proteins are mostly secreted proteins that suppress host defense mechanisms and facilitate infection. Several general characteristics have been attributed to these proteins, which are used for *in silico* detection (see further). Also, the presence of cell wall degrading enzymes (CWDE) in plant pathogens has been extensively studied from genome sequences, and the diversity of these enzymes appears to increase with the host range [90]. In several obligate biotrophic and symbiotic fungi the CWDE content was reported to be reduced or even completely absent, and it is suggested that metabolism-related genes have a potential predictive power to elucidate fungal lifestyle and infection strategies [88]. Analysis of the genomes of

Chapter 1

pathogenic fungi points toward a role of transposable elements (TEs) in genome plasticity and evolution. In several fungal pathogens differences in the number and nature of these TEs were found in the core chromosomes compared to dispensable supernumerary chromosomes. In *Fusarium poae*, RIP (repeat-induced point mutation) protects core chromosomes against repeat expansions. Interestingly, this defense mechanism against TEs was absent on supernumerary chromosomes, and as a result these genomic regions provide opportunities for duplication and diversification [91].

With the emergence of environmental DNA sequencing (metagenomics) opportunities arise to study genomes in their environmental context. Metagenomic studies allow for biodiversity detection in environmental samples, and are used to assess the presence of pathogens in vectors and hosts [e.g. 92, 93]. An increasing diversity of undescribed (in particular) early diverging fungi is uncovered from metagenomic studies, which are often known only from their genetic material and are referred to as “dark matter fungi” [94, 95]. In addition, community-level gene catalogs are constructed from metagenomes and used to elucidate environment-dependent (e.g. soil versus rhizosphere) differences [96]. In case for the obligate biotrophic *S. endobioticum*, resting spores are the purest form of the pathogen that can be obtained in quantities large enough for sequencing. These resting spores represent a metagenome as they contain both the organism of interest, but also the host and other associated soilborne micro-organisms.

For *S. endobioticum*, genomic features linked to obligate biotrophy and pathogenicity will improve our knowledge of the pathogen, and they may hold the key to more targeted and efficient control strategies, either based on novel control agents or improved resistance against the pathogen.

Potato wart disease resistance

Following the rapid spread of the disease in the United Kingdom, technical inspectors were appointed in 1907 to assess the extent of the disease. In the following year, Gough learned from one of his interviews that the potato variety “Snowdrop” remained unaffected when cultivated in infested plots [36]. Classic breeding for resistance began around a century ago and was very successful in selecting varieties resistant to pathotype 1(D1) of the disease. The use of resistant varieties and strict phytosanitary measures were so successful that breeding for potato wart resistance lost priority until the discovery of new pathotypes that

caused disease in previously resistant varieties [97]. There is still a necessity to develop new potato varieties that combine desired agronomic traits with resistance against the “higher” pathotypes. This is illustrated by the 2018 list of resistant potato cultivars which is released annually by the Dutch government [98]. In this release, 450 potato varieties are listed providing resistance against pathotype 1(D1), whereas only 19 varieties providing resistance to 2(G1), 29 with 6(O1) resistance and 15 with 18(T1) resistance are listed. Only five cultivars on this list are resistant to all of these pathotypes.

The first locus providing resistance to *S. endobioticum* was identified by Hehl *et al.* [99], and mapped to the distal end of the long arm of chromosome XI. The identified single dominant gene provides resistance to pathotype 1(D1) isolates and is referred to as *Sen1*. The genomic region in which *Sen1* resides is regarded as a “hot spot” for genes conferring resistance to several pathogens, such as the nematodes *Globodera pallida* and *Meloidogyne chitwoodi*, the bacterium *Erwinia carotovora*, the oomycete *Phytophthora infestans* and several potato viruses [7]. A second independent dominant gene for pathotype 1(D1) resistance (*Sen1-4*), which genetically mapped to the long arm of chromosome IV, was identified by Brugmans *et al.* [100]. These studies demonstrate that resistance to pathotype 1(D1) can be achieved by a single gene, but that there are different sources for pathotype 1(D1) resistance [100].

Genomic regions linked to potato wart resistance were identified from quantitative resistance locus (QRL) analyses providing resistance to (combinations of) the pathotypes of major importance, i.e. 1(D1), 2(G1), 6(O1) and 18(T1), on chromosomes 1, 2, 6, 7, 8, 9, 10 and 11 [101, 102]. In both studies the resistance to the three “higher” pathotypes 2(G1), 6(O1) and (18T1) was found to be highly correlated. In contrast to the QRLs with limited phenotypic effects on resistance against virulent pathotypes, a gene (*Sen2*) providing broad resistance to pathotypes 1(D1), 2(Ch1), 2(G1), 3(M1), 6(O1), 8(F1), 18(T1) and 39(P1) was identified. Almost all variance (95 - 98% depending on the pathotype) observed in the resistance of progeny of the mapping population was ascribed to the presence of this gene. Also localized on chromosome 11, *Sen2* was estimated to be approximately 32 Mbp distant from the *Sen1* locus [103].

Molecular pathogen-host interactions

Throughout the evolutionary history of terrestrial flora, that spans over 450 million years [104], plants have been exposed to diverse microbial agents establishing beneficial or parasitic interactions. The genomes of plants and the micro-organisms aiming to colonize them have been shaped by millions of years of co-evolution [105].

Fungi that attempt to colonize plants are recognized by the plant innate immune system and trigger host defenses. The initial defense responses are triggered by conserved indispensable molecules exposed by microbes, and characteristic for whole classes of micro-organisms. These molecules are referred to as pathogen-associated molecular patterns (PAMPs), or microbe-associated molecular patterns (MAMPs) as they are not restricted to plant pathogens [89]. One of the best studied PAMPs is the bacterial flagellin protein that is the main building block in eubacterial flagella [106]. In fungi, the cell wall component chitin and acts as a PAMP and is recognized by membrane-bound pattern recognition receptors (PRRs). The defense response triggered by PRRs is called PAMP-triggered immunity (PTI), and leads to molecular and physical adjustments that arrest pathogen proliferation mediated by the production of reactive-oxygen species, rapid changes in hormone biosynthesis and cell wall reinforcement by callose deposition [107]. PTI is considered to be effective to non-adapted microbes, and is generally regarded as part of the first layer of the plant defense system [89, 106].

To successfully colonize the plant host, invading micro-organisms have to avoid triggering PTI, suppress it, or cope with the deleterious effects of PTI. To achieve this, effector proteins are secreted that shield the fungus, suppress the host immune response or manipulate the host cell physiology [108]. Currently, close to 90 fungal and oomycete effectors with known function have been described [109]. Apart from shielding pathogen chitin or interfering host proteins interacting with chitin, host proteases and the salicylic acid and jasmonic signaling pathways (involved in plant immunity) are among the known effector targets [108, 110]. Opposed to the highly conserved PAMPs and MAMPs, effectors are very diverse, often specialized towards a specific host, and species or strain specific. The complex effector repertoire of plant pathogenic fungi is believed to influence the fungal lifestyle and host specialization [89].

Plants have responded to effector-induced suppression of PTI by evolving highly specific resistance (R) proteins. These intracellular receptors belong to the conserved nucleotide binding and leucine-rich repeat containing (NLR) superfamily, and can interact directly with pathogen effectors or indirectly involving effector targets or structural mimics of such targets [89, 111]. Upon recognition of effector proteins by the cognate R proteins, the second layer of the plant defense system, called effector-triggered immunity (ETI), is activated. Effects of ETI are the induction of pathogenesis-related genes and localized cell death at the site of infection. This process, that is the result of a specific interaction between the pathogen effector and the plant R protein following the gene-for-gene concept defined by Flor [112], is associated with a local hypersensitive response (HR). Genes coding for effectors, that are recognized by R proteins resulting in HR, are referred to as avirulence (*Avr*) genes.

Access to sequenced and annotated genomes have spurred the identification of new effector genes in fungi with biotrophic, necrotrophic and hemibiotrophic lifestyles, and inspired the creation of a number of computational tools that can be used in the detection of effector candidates [113]. Effector proteins are often secreted by the conventional endoplasmic reticulum - Golgi apparatus route, which requires the presence a N-terminal secretion signal peptide. Presence of a signal peptide and absence of transmembrane domains are strong prerequisites for candidate effectors. Frequently, additional criteria have been attributed to effector proteins such as molecular weight, size, and cysteine-richness. These characteristics are regarded particularly useful for effectors that have to survive the harsh conditions in the apoplast [89]. However, these criteria do not apply to all effector proteins identified so far [114]. As is the case for absence of predicted function or functional domains which is sometimes used as criterion in effector prediction. Some secreted proteins are translocated to the host cytoplasm. In oomycetes the largest group of translocated effectors contain a RxLR motif, and this motif was previously suggested to play an active role in the uptake of effector proteins. However, in the *P. infestans* AVR3a protein the RxLR is cleaved before secretion [115]. Still effector candidacy is sometimes based on specific motifs such as the RxLR motif or Crinkler (CRN) motifs [116]. The field studying effectors using genome-wide approaches is referred to as effectomics [117].

We hypothesize that *S. endobioticum*, similar to other fungal pathosystems, has evolved effectors suppressing or avoiding recognition by PRRs triggering PTI. These effectors are believed to be generally present among different pathotypes as the fungus has been reported to penetrate the host cell even in incompatible interactions [35]. *S. endobioticum*

grows intracellular in the host cell, where the fungal thallus migrates to the host nucleus which enlarges after fungal penetration. Hypertrophy and hyperplasia are induced by the pathogen as it takes control of the host in this intimate relationship. We believe these host reactions are the result of fungal effectors that are secreted into the host cytoplasm after penetration of the host cell. In addition, it is likely that the phenotypical differences obtained with *S. endobioticum* pathotypes in potato varieties with the dominant *Sen1* and *Sen2* resistance genes are based on a gene-for-gene interaction with a single specific effector triggering HR (Fig. 7).

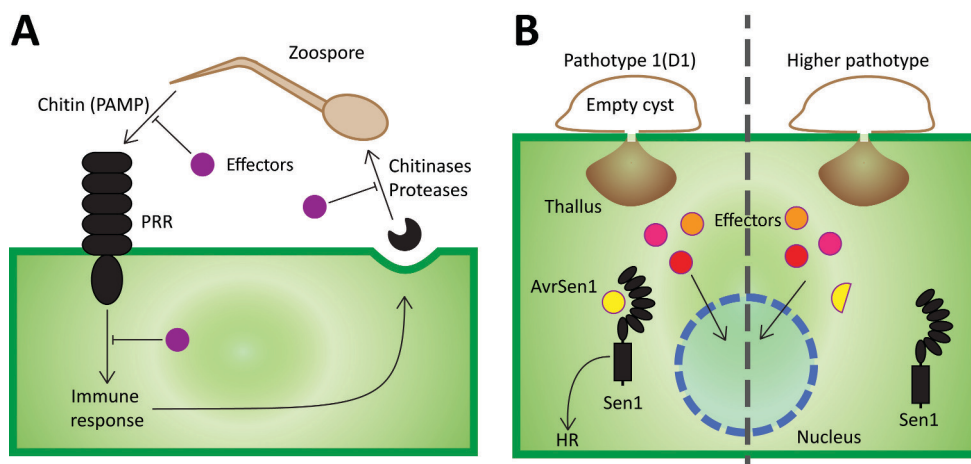


Figure 7. Hypothesized model for co-evolution of *S. endobioticum* effectors and the plant immune system. A. outside the host cell Host cells detect chitin in the fungal cell wall with pattern recognition receptors (PRR), triggering an immune response involving (among others) secretion of chitinases and proteases. The pathogen secretes effectors to either shield chitin, cytoplasmic blocking the PRR-induced signal avoiding triggering of immune responses, or apoplastic inhibition of chitinases and proteases released by the host cell in response to PAMP recognition by PRRs. **B. Inside the host cell** After penetration, the fungal thallus secretes effector proteins to manipulate the host cell and induces hypertrophy and hyperplasia. In an incompatible interaction (left of the dotted line, illustrated with a pathotype 1(D1) isolate attempting to colonize a *Sen1* plant), effector proteins are recognized by their cognate R protein. Here, the plant *Sen1* is illustrated as a NLR that upon recognition of the cognate Avr (*AvrSen1*) triggers a hypersensitive response (HR). In the compatible interaction (right of the dotted line, illustrated with colonization of a *Sen1* plant by a higher pathotype), the *AvrSen1* protein is either absent or it is structurally altered thusly avoiding recognition by *Sen1* and subsequent triggering of a HR.

From metagenome to gene: thesis outline

In this thesis, results from integrated functional genomic and effectoromic studies are presented. We set out to assemble and functionally annotate the *S. endobioticum* genome from a metagenomic sample to gain insights into genomic features underlying the obligate biotrophic and pathogenic lifestyle of the pathogen. As currently used bioassays oversimplify the *S. endobioticum* diversity, a population study using the complete mitochondrial genome was performed. Next, to determine if plant *R* genes were expressed in aboveground plant parts, a requirement for agroinfiltration experiments, an alternative aboveground bioassay was developed. Results from chapters 2, 3 and 4 were combined in pursuit of the *AvrSen1* gene which is hypothesized to be recognized by *Sen1* potato plants. Identifying *S. endobioticum* Avrs are of great interest as they potentially can replace time-consuming and labor-intensive bioassays used for pathotyping and resistance screening. Additionally, they can be instrumental in identifying plant *R* genes coding for wart disease resistance.

In **chapter 2** the genome sequences of two *S. endobioticum* isolates representing different pathotypes from different geographic origins are presented. Resting spores were the purest form of the obligate biotrophic pathogen that could be used to generate genomic sequences, and they represent a metagenomic sample containing sequences from the pathogen, its host and other soilborne biological agents. We used a comparative functional genomics approach exploiting knowledge acquired from other fungal taxa and compared this to several saprobic and pathogenic representatives of Chytridiomycota. Observations linked to chytrid biology, obligate biotrophy, genome plasticity and pathogenicity are reported. Our study underlines the high diversity in chytrids compared to the well-studied Ascomycota and Basidiomycota, and reflects characteristic biological differences between the phyla. Moreover, it shows commonalities in genomic features among pathogenic fungi. *S. endobioticum* is the most basal lineage among the fungal plant pathogens and is the first available genome of an obligate soilborne fungal species.

To identify potential linkages between genotypes, geographical origins and pathotypes, the *S. endobioticum* mitochondrial genomes of 30 isolates were sequenced, assembled and analyzed (**chapter 3**). We applied a comparative approach using chytrid mitogenomes to determine taxonomical relationships and to gain insights into the evolution and recent history of introductions of *S. endobioticum*. The mitochondrial genome is often assumed to be a small and circular molecular, but the mitochondrial genome of *S. endobioticum* was

Chapter 1

found to be linear with terminal inverted repeats. Mitochondrial genomic variation shows that *S. endobioticum* has likely been introduced into Europe multiple times, that several pathotypes emerged independently in different genetic backgrounds, and that isolates represent communities of different genotypes. We also demonstrated that the use of semi-resistant potato cultivars can trigger a rapid shift in virulence which coincides with a shift in the mitochondrial haplotype and is likely the result of disruptive selection in the original mix of genotypes in the *S. endobioticum* isolate.

Chapter 4 describes the development of an alternative bioassay in which aboveground plant parts are inoculated, compared to the currently used bioassays in which etiolated tuber sprouts are inoculated. Recognition of *S. endobioticum* pathotypes by differentially resistant potato varieties was identical in aboveground plant parts and the tuber-based bioassays, which suggests that *S. endobioticum* resistance genes are expressed both in etiolated “below-ground” sprouts and green aboveground organs. Strongly reduced microbial contamination was observed in the aboveground bioassay which could be an important advantage of this bioassay to study this obligate biotrophic plant-pathogen interaction. As wart resistance is active in both below and above ground organs, the aboveground bioassay can potentially speed up screening for *S. endobioticum* resistance in potato breeding programs as it omits the requirement for tuber formation. In addition, possibilities arise to express *S. endobioticum* effectors in potato leaves through agroinfiltration, thereby providing additional phenotyping tools for research and breeding.

A comparative genomic approach to identify candidate effectors, using the annotated *S. endobioticum* pathotype 1(D1) genome as reference, is reported in **chapter 5**. We aimed to identify the effector that is recognized by the potato *Sen1 R* gene product, resulting in pathotype 1(D1) resistance. The cognate effector of the *Sen1* resistance gene, referred to as *AvrSen1*, was believed to be present in pathotype 1(D1) but absent in all other pathotypes. The *S. endobioticum* specific secretome was screened in search for effectors following this specific pattern of presence and absence using fourteen *S. endobioticum* isolates representing six different pathotypes. We discovered a single *AvrSen1* candidate and expression of this candidate in potato plants showed a specific hypersensitive response in *Sen1* plants. Natural allelic variants of the *AvrSen1* gene found in other pathotypes did not trigger this specific hypersensitive response. The *AvrSen1* gene is the first *Avr* gene identified in Chytridiomycota.

In **chapter 6** the results of this thesis are summarized and discussed with a particular focus on the formation of pathotypes and pathotype diversity. Outcomes of our comparative functional genomics are linked to observations from early research on potato wart disease, and the importance of the identification of the *AvrSen1* gene and its observed within-isolate variation, are discussed. Finally, opportunities and challenges for future research are presented.

References

1. Schilberszky, K., Ein neuer Schorfparasit der Kartoffelknollen. Ber. Deut. Botan. Ges., 1896. 14: p. 36-37.
2. Schilberszky, K., Die gesamtbiologie des kartoffel-krebses. 1930, Freising, München: F. P. Datterer & cie.
3. MacLeod, A., *et al.*, Evolution of the international regulation of plant pests and challenges for future plant health. Food Security, 2010. 2(1): p. 49-70.
4. Woodham Smith, C., The great hunger, Ireland 1845-9. 1962, London: H. Hamilton. 509 p.
5. Santini, A., *et al.*, Tracing the role of human civilization in the globalization of plant pathogens. The ISME Journal, 2018. 12(3): p. 647-652.
6. Hampson, M.C., History, biology and control of potato wart disease in Canada. Canadian Journal of Plant Pathology, 1993. 15(4): p. 223-244.
7. Obidiegwu, J.E., K. Flath, and C. Gebhardt, Managing potato wart: a review of present research status and future perspective. Theor Appl Genet, 2014. 127(4): p. 763-80.
8. Federal Select Agent Program, Select Agents and Toxins List. 2018 [cited 2018 25 April]; Available from: www.selectagents.gov/selectagentsandtoxinslist.html.
9. Heckman, D.S., *et al.*, Molecular Evidence for the Early Colonization of Land by Fungi and Plants. Science, 2001. 293(5532): p. 1129.
10. Sparrow, F.K., Aquatic Phycomycetes. 2nd ed. 1960, Ann Arbor: The university of Michigan Press.
11. Barr, D.J.S., Chytridiomycota, in Systematics and Evolution: Part A. 2001, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 93-112.
12. Van Rooij, P., *et al.*, Amphibian chytridiomycosis: a review with focus on fungus-host interactions. Veterinary research, 2015. 46: p. 137-137.
13. Percival, J., Potato wart disease: the life history and cytology of *Synchytrium endobioticum*. Zentralblatt für Bakteriologie, Parasitenkunde, Infektionskrankheiten und Hygiene., 1910. 2(25): p. 440-447.
14. Curtis, K.M., IX.— The life-history and cytology of *Synchytrium endobioticum* (schilb.), perc., the cause of wart disease in potato. Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character, 1921. 210(372-381): p. 409.
15. Karling, J.S., *Synchytrium*. 1 ed. 1964: Academic Press, New York.
16. Longcore, J.E., D.R. Simmons, and P.M. Letcher, *Synchytrium microbalum* sp. nov. is a saprobic species in a lineage of parasites. Fungal Biology, 2016. 120(9): p. 1156-1164.
17. Kole, A.P., Resting-spore germination in *Synchytrium endobioticum*. Netherlands Journal of Plant Pathology, 1965. 71(3): p. 72-78.

Chapter 1

18. Sharma, R. and T.L.R.H. Cammack, Spore germination and taxonomy of *Synchytrium endobioticum* and *S. succisae*. Transactions of the British Mycological Society, 1976. 66(1): p. 137-147.
19. Lange, L. and L.W. Olson, Germination and parasitism of the resting sporangia of *Synchytrium endobioticum*. Protoplasma, 1981. 106(1): p. 69-82.
20. Hampson, M.C., Sequence of events in the germination of the resting spore of *Synchytrium endobioticum*, European pathotype 2, the causal agent of potato wart disease. Canadian Journal of Botany, 1986. 64(9): p. 2144-2150.
21. James, T.Y., *et al.*, A molecular phylogeny of the flagellated fungi (Chytridiomycota) and description of a new phylum (Blastocladiomycota). Mycologia, 2006. 98(6): p. 860-71.
22. Smith, D.S., *et al.*, Phylogeny of the genus *Synchytrium* and the development of TaqMan PCR assay for sensitive detection of *Synchytrium endobioticum* in soil. Phytopathology, 2014. 104(4): p. 422-32.
23. Lange, L. and L.W. Olson, The zoospore of *Synchytrium endobioticum*. Botany, 1978. 56(10): p. 1229-1239.
24. Barr, D.J.S., An outline for the reclassification of the Chytridiales, and for a new order, the Spizellomycetales. Canadian Journal of Botany, 1980. 58(22): p. 2380-2394.
25. Grigoriev, I.V., *et al.*, MycoCosm portal: gearing up for 1000 fungal genomes. Nucleic Acids Res, 2014. 42(Database issue): p. D699-704.
26. Choi, J. and S.-H. Kim, A genome Tree of Life for the Fungi kingdom. Proceedings of the National Academy of Sciences, 2017. 114(35): p. 9391-9396.
27. Lange, L. and L.W. Olson, The zoospore of *Synchytrium endobioticum*. Canadian Journal of Botany, 1978. 56(10): p. 1229-1239.
28. Weiss, F., The conditions of infection in potato wart. American Journal of Botany, 1925. 12(7): p. 413-443.
29. Szabo, L.J. and W.R. Bushnell, Hidden robbers: The role of fungal haustoria in parasitism of plants. Proceedings of the National Academy of Sciences, 2001. 98(14): p. 7654.
30. Lange, L. and L.W. Olson, Development of the zoosporangia of *Synchytrium endobioticum*, the causal agent of potato wart disease. Protoplasma, 1981. 106(1): p. 97-108.
31. Przetakiewicz, J., The Viability of Winter Sporangia of *Synchytrium endobioticum* (Schilb.) Perc. from Poland. American Journal of Potato Research, 2015. 92(6): p. 704-708.
32. Lange, L. and L.W. Olson, Development of the resting sporangia of *Synchytrium endobioticum*, the causal agent of potato wart disease. Protoplasma, 1981. 106(1): p. 83-95.
33. Bal, A.K., A.M. Murphy, and M.C. Hampson, Ultrastructure and Chemical Analysis of the Resting Sporangium Wall of *Synchytrium endobioticum*. Canadian Journal of Plant Pathology, 1981. 3(2): p. 86-89.
34. Sidhu, G.S., Sexual and parasexual variability in soil fungi with special reference to *Fusarium moniliforme*. Phytopathology, 1983. 73(6): p. 952-956.
35. Cartwright, K., On the Nature of the Resistance of the Potato to Wart Disease. Annals of Botany, 1926. 40(158): p. 391-395.
36. Gough, G.C., Wart disease of potatoes (*Synchytrium endobioticum* Perc.). A study of its history, distribution and the discovery of immunity. Jour. Roy. Hort. Soc., 1919. 45: p. 301-312.
37. Olsen, O.A., Potato wart investigations in Newfoundland. Canadian Plant Diseases Survey, 1961. 41: p. 148-155.
38. Braun, H.C., Biologische Spezialisierung bei *Synchytrium endobioticum* (Schilb.). Zeitschrift für Pflanzenkrankheiten (Pflanzenpathologie) und Pflanzenschutz, 1942. 52(11): p. 481-486.

39. Blattny, C., Preliminary note on the races of *Synchytrium endobioticum*. Annales de l'Academie tchecoslovaque d'agriculture, 1942. 17: p. 40-46.
40. Baayen, R.P., *et al.*, History of potato wart disease in Europe - a proposal for harmonisation in defining pathotypes. European Journal of Plant Pathology, 2006. 116(1): p. 21-31.
41. OEPP/EPP0, PM 7/28 (2) *Synchytrium endobioticum*. EPP0 Bulletin, 2017. 47(3): p. 420-440.
42. Çakır, E., *et al.*, Identification of pathotypes of *Synchytrium endobioticum* found in infested fields in Turkey. EPP0 Bulletin, 2009. 39(2): p. 175-178.
43. Przetakiewicz, J., First Report of New Pathotype 39(P1) of *Synchytrium endobioticum* Causing Potato Wart Disease in Poland. Plant Disease, 2015. 99(2): p. 285-285.
44. Anonymous, Final report of an audit carried out in Poland from 04 November to 14 November 2014 in order to evaluate the plant health controls applied in the potato sector. 2014, European Commission.
45. Cotton, A.D.C., Host Plants of *Synchytrium endobioticum*. Studies from the Pathological Laboratory: IV. Royal Botanic Gardens, 1916. 1916(10): p. 272-275.
46. Hampson, M.C., Infection of additional hosts of *Synchytrium endobioticum*, the causal agent of potato wart disease: I. Tomato. Canadian Plant Disease Survey, 1976. 56: p. 93-94.
47. Martin, M.S., Additional hosts of *Synchytrium endobioticum* (Schilb.) Perc. Annals of Applied Biology, 1929. 16(3): p. 422-429.
48. Hampson, M.C., Infection of additional hosts of *Synchytrium endobioticum*, the causal agent of potato wart disease: 2. Tomato, tobacco and species of Capsicastrum, Datura, Physalis and Schizanthus. Canadian Plant Disease Survey, 1979. 59(1): p. 3-6.
49. Hampson, M.C. and N.F. Haard, Pathogenesis of *Synchytrium endobioticum*: 1. Infection responses in potato and tomato. Canadian Journal of Plant Pathology, 1980. 2(3): p. 143-147.
50. EPP0. European and Mediterranean Plant Protection Organization (EPP0) global database. 2018 [cited 2018 12 September]; Available from: <https://gd.eppo.int>.
51. Przetakiewicz, J., Assessment of the resistance of potato cultivars to *Synchytrium endobioticum* (Schilb.) Per. in Poland. EPP0 Bulletin, 2008. 38(2): p. 211-215.
52. Moore, W.C., The Breakdown of Immunity from Potato Wart Disease. Outlook on Agriculture, 1957. 1(6): p. 240-243.
53. Phadtare, S.G., P.R. Singh, and G.S. Shekhawat, Wart disease of potato in Darjeeling hills. Technical Bulletin. Vol. 19. 1990, Shimla, India: Central Potato Research Institute. 79.
54. Hampson, M., Research on potato wart disease in the USSR-a literature review (1955-1977). Canadian Plant Disease Survey (Canada), 1979.
55. Çakır, E., First report of Potato Wart Disease caused by *Synchytrium endobioticum* in Turkey. New Disease Reports, 2005. 11(12).
56. Dimitrova, L., *et al.*, Occurrence of potato wart disease (*Synchytrium endobioticum*) in Bulgaria: identification of pathotype(s) present. EPP0 Bulletin, 2011. 41(2): p. 195-202.
57. Gorgiladze, L., *et al.*, First report of *Synchytrium endobioticum* causing potato wart in Georgia. New Disease Reports, 2014. 30(4).
58. Anonymous, *Synchytrium endobioticum* found in Denmark. EPP0 Reporting Service, 2014. 11(206).
59. Franc, G., Potato wart. Online. APSnet Features. doi, 2007. 10.
60. Jösting, X., Verhütung und bekämpfung des kartoffelkrebses. Deutsch. Landwirtsch. Presse, 1909. 36(88): p. 941.

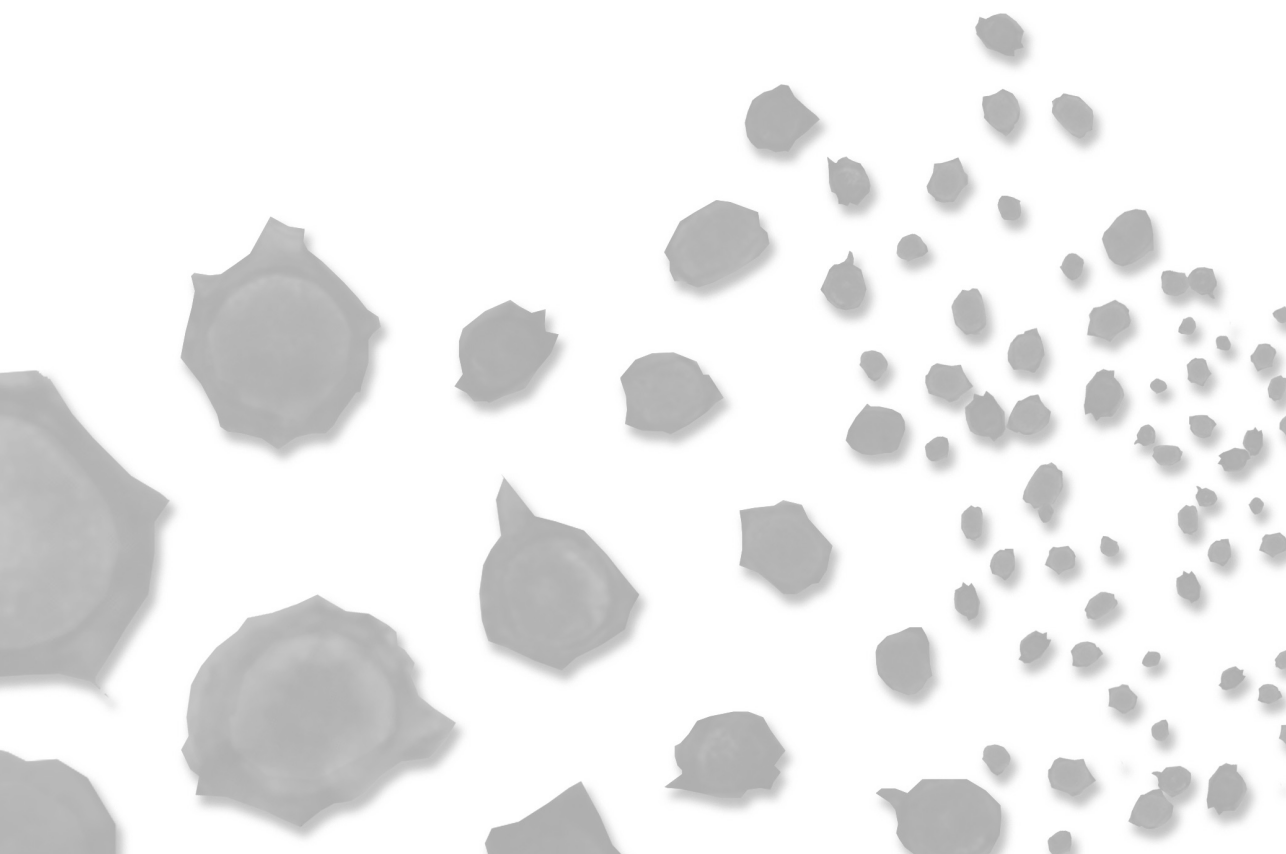
Chapter 1

61. Hampson, M.C., A qualitative assessment of wind dispersal of resting spores of *Synchytrium endobioticum*, the causal agent of wart disease of potato. *Plant Disease*, 1996. 80(7): p. 779-782.
62. Hampson, M.C. and S.L. Wood, Detection of infective resting spores of *Synchytrium endobioticum* in vehicles. *Canadian Journal of Plant Pathology*, 1997. 19(1): p. 57-59.
63. Hampson, M.C., S.L. Wood, and J.W. Coombes, Detection of resting spores of *Synchytrium endobioticum* in soil from vehicles at Port-aux-Basques, Newfoundland. *Canadian Journal of Plant Pathology*, 1996. 18(1): p. 59-63.
64. OEPP/EPPO, Guidelines for the management of plant health risks of biowaste of plant origin. *EPPO Bulletin*, 2008. 38(1): p. 4-9.
65. Steinmöller, S., *et al.*, Effects of sanitation processes on survival of *Synchytrium endobioticum* and *Globodera rostochiensis*. *European Journal of Plant Pathology*, 2012. 133(3): p. 753-763.
66. OEPP/EPPO, PM 7/28 (1) *Synchytrium endobioticum*. *EPPO Bulletin*, 2004. 34(2): p. 213-218.
67. Anonymous, De aardappelwratziekte. 2nd ed. Verslagen en mededeelingen van den plantenziektenkundige dienst te Wageningen. Vol. 16. 1928, Wageningen, the Netherlands: H. Veenman & Zonen. 24.
68. Anonymous, Black Scab (wart disease) in the Netherlands. 1st ed. Verslagen en mededeelingen van den plantenziektenkundige dienst te Wageningen. Vol. 16a. 1921, Wageningen, the Netherlands: H. Veenman & Zonen.
69. Nienhuis, P.H., Environmental history of the Rhine-Meuse Delta: an ecological story on evolving human-environmental relations coping with climate change and sea-level rise. 2008: Springer Science & Business Media.
70. Hampson, M.C., Control of potato wart disease through the application of chemical soil treatments: a historical review of early studies (1909-1928). *EPPO Bulletin*, 1988. 18(1): p. 153-161.
71. Hampson, M.C. and J.W. Coombes, Reduction of potato wart disease with crushed crabshell: suppression or eradication? *Canadian Journal of Plant Pathology*, 1995. 17(1): p. 69-74.
72. Anonymous, Council Directive 69/464/EEC of 8 December 1969 on control of Potato Wart Disease. 1969.
73. OEPP/EPPO, PM 9/5 (2) *Synchytrium endobioticum*. *EPPO Bulletin*, 2017. 47(3): p. 511-512.
74. OEPP/EPPO, PM 3/59 (3) *Synchytrium endobioticum*: descheduling of previously infested plots. *EPPO Bulletin*, 2017. 47(3): p. 366-368.
75. Warenautoriteit, N.V.-e. Teeltvoorschrift wratziekte. [cited 2019 22 January]; Available from: <https://www.nvwa.nl/onderwerpen/teeltvoorschriften-akkerbouw-en-tuinbouw/teeltvoorschrift-wratziekte>.
76. Spieckermann, A.K., P., Testing potatoes for wart resistance. *Deutsche Landwirtschaftliche Presse*, 1924. 51: p. 114-115.
77. Glynne, M.D., Infection experiments with wart disease of potatoes. *Synchytrium endobioticum* (Schilb.) Perc. *Annals of Applied Biology*, 1925. 12(1): p. 34-60.
78. Lemmerz, J., Neues vereinfachtes Infektionsverfahren zur Prüfung von Kartoffelsorten auf Krebsfestigkeit. *Der Züchter*, 1930. 2: p. 288-297.
79. Noble, M.G., M.D., Wart disease of potatoes. *FAO Plant. Protection Bulletin*, 1970. 18: p. 125-135.
80. Niepold, F. and H. Stachewicz, PCR-detection of *Synchytrium endobioticum* (Schub.) Perc./PCR-Nachweis von *Synchytrium endobioticum* (Schilb.) Perc. *Journal of Plant Diseases and Protection*, 2004. 111(4): p. 313-321.
81. van den Boogert, P.H.J.F., *et al.*, Development of PCR-based Detection Methods for the Quarantine Phytopathogen *Synchytrium endobioticum*, Causal Agent of Potato Wart Disease. *European Journal of Plant Pathology*, 2005. 113(1): p. 47-57.

82. van Gent-Pelzer, M.P.E., M. Krijger, and P.J.M. Bonants, Improved real-time PCR assay for detection of the quarantine potato pathogen, *Synchytrium endobioticum*, in zonal centrifuge extracts from soil and in plants. *European Journal of Plant Pathology*, 2010. 126(1): p. 129-133.
83. Bonants, P.J.M., *et al.*, A real-time TaqMan PCR assay to discriminate between pathotype 1 (D1) and non-pathotype 1 (D1) isolates of *Synchytrium endobioticum*. *European Journal of Plant Pathology*, 2015. 143(3): p. 495-506.
84. Flath, K., *et al.*, Interlaboratory tests for resistance to *Synchytrium endobioticum* in potato by the Glynne-Lemmerzähl method. *EPPO Bulletin*, 2014. 44(3): p. 510-517.
85. van de Vossenbergh, B.T.L.H., *et al.*, Euphresco Sendo: An international laboratory comparison study of molecular tests for *Synchytrium endobioticum* detection and identification. *European Journal of Plant Pathology*, 2018. 151(3): p. 757-766.
86. Gagnon, M.C., *et al.*, Development of Polymorphic Microsatellite Loci for Potato Wart from Next-Generation Sequence Data. *Phytopathology*, 2016. 106(6): p. 636-44.
87. Busse, F., *et al.*, Genomic and Transcriptomic Resources for Marker Development in *Synchytrium endobioticum*, an Elusive but Severe Potato Pathogen. *Phytopathology*, 2016. 107(3): p. 322-328.
88. Aylward, J., *et al.*, A plant pathology perspective of fungal genome sequencing. *IMA fungus*, 2017. 8(1): p. 1-15.
89. Lo Presti, L., *et al.*, Fungal effectors and plant susceptibility. *Annu Rev Plant Biol*, 2015. 66: p. 513-45.
90. Kubicek, C.P., T.L. Starr, and N.L. Glass, Plant Cell Wall-Degrading Enzymes and Their Secretion in Plant-Pathogenic Fungi. *Annual Review of Phytopathology*, 2014. 52(1): p. 427-451.
91. Vanheule, A., *et al.*, Living apart together: crosstalk between the core and supernumerary genomes in a fungal plant pathogen. *BMC Genomics*, 2016. 17: p. 670.
92. Miller, K.E., *et al.*, Metabarcoding of fungal communities associated with bark beetles. *Ecology and evolution*, 2016. 6(6): p. 1590-1600.
93. Roossinck, M.J., Metagenomics of plant and fungal viruses reveals an abundance of persistent lifestyles. *Frontiers in microbiology*, 2015. 5: p. 767-767.
94. Grossart, H.-P., *et al.*, Discovery of dark matter fungi in aquatic ecosystems demands a reappraisal of the phylogeny and ecology of zoospore fungi. *Fungal Ecology*, 2016. 19: p. 28-38.
95. Berbee, M.L., T.Y. James, and C. Strullu-Derrien, Early Diverging Fungi: Diversity and Impact at the Dawn of Terrestrial Life. *Annu Rev Microbiol*, 2017. 71: p. 41-60.
96. Ofaim, S., *et al.*, Analysis of Microbial Functions in the Rhizosphere Using a Metabolic-Network Based Framework for Metagenomics Interpretation. *Frontiers in Microbiology*, 2017. 8(1606).
97. Obidiegwu, J.E., *et al.*, Genomic architecture of potato resistance to *Synchytrium endobioticum* disentangled using SSR markers and the 8.3k SolCAP SNP genotyping array. *BMC Genet*, 2015. 16: p. 38.
98. Government, D., Besluit van de Minister van Landbouw, Natuur en Voedselkwaliteit van 20 februari 2018, nr. trcvwa/2018/1406, houdende aanwijzing en bekendmaking van de aardappelen behorende tot resistente rassen, bedoeld in artikel 3, tweede en derde lid, van het Besluit bestrijding wratziekte 1973. *Staatscourant*, 2018(11941): p. 1-10.
99. Hehl, R., *et al.*, TMV resistance gene N homologues are linked to *Synchytrium endobioticum* resistance in potato. *Theoretical and Applied Genetics*, 1999. 98(3): p. 379-386.
100. Brugmans, B., *et al.*, Exploitation of a marker dense linkage map of potato for positional cloning of a wart disease resistance gene. *Theor Appl Genet*, 2006. 112(2): p. 269-77.

Chapter 1

101. Ballvora, A., *et al.*, Multiple alleles for resistance and susceptibility modulate the defense response in the interaction of tetraploid potato (*Solanum tuberosum*) with *Synchytrium endobioticum* pathotypes 1, 2, 6 and 18. *Theor Appl Genet*, 2011. 123(8): p. 1281-92.
102. Groth, J., *et al.*, Molecular characterisation of resistance against potato wart races 1, 2, 6 and 18 in a tetraploid population of potato (*Solanum tuberosum* subsp. *tuberosum*). *J Appl Genet*, 2013. 54(2): p. 169-78.
103. Plich, J., *et al.*, Novel gene *Sen2* conferring broad-spectrum resistance to *Synchytrium endobioticum* mapped to potato chromosome XI. *Theoretical and Applied Genetics*, 2018. 131(11): p. 2321-2331.
104. Delwiche, Charles F. and Endymion D. Cooper, The Evolutionary Origin of a Terrestrial Flora. *Current Biology*, 2015. 25(19): p. R899-R910.
105. Carella, P., E. Evangelisti, and S. Schornack, Sticking to it: phytopathogen effector molecules may converge on evolutionarily conserved host targets in green plants. *Curr Opin Plant Biol*, 2018. 44: p. 175-180.
106. Zipfel, C., Pattern-recognition receptors in plant innate immunity. *Curr Opin Immunol*, 2008. 20(1): p. 10-6.
107. Toruño, T.Y., I. Stergiopoulos, and G. Coaker, Plant-Pathogen Effectors: Cellular Probes Interfering with Plant Defenses in Spatial and Temporal Manners. *Annual Review of Phytopathology*, 2016. 54(1): p. 419-441.
108. Asai, S. and K. Shirasu, Plant cells under siege: plant immune system versus pathogen effectors. *Curr Opin Plant Biol*, 2015. 28: p. 1-8.
109. Selin, C., *et al.*, Elucidating the Role of Effectors in Plant-Fungal Interactions: Progress and Challenges. *Frontiers in Microbiology*, 2016. 7(600).
110. Wang, Y. and Y. Wang, Trick or Treat: Microbial Pathogens Evolved Apoplastic Effectors Modulating Plant Susceptibility to Infection. *Mol Plant Microbe Interact*, 2018. 31(1): p. 6-12.
111. Jones, J.D., R.E. Vance, and J.L. Dangl, Intracellular innate immune surveillance devices in plants and animals. *Science*, 2016. 354(6316).
112. Zurbriggen, M.D., N. Carrillo, and M.-R. Hajirezaei, ROS signaling in the hypersensitive response: when, where and what for? *Plant signaling & behavior*, 2010. 5(4): p. 393-396.
113. Sonah, H., R.K. Deshmukh, and R.R. Belanger, Computational Prediction of Effector Proteins in Fungi: Opportunities and Challenges. *Front Plant Sci*, 2016. 7: p. 126.
114. Djamei, A., *et al.*, Metabolic priming by a secreted fungal effector. *Nature*, 2011. 478(7369): p. 395-8.
115. Wawra, S., *et al.*, The RxLR Motif of the Host Targeting Effector *AVR3a* of *Phytophthora infestans* Is cleaved before Secretion. *Plant Cell*, 2017. 29(6): p. 1184-1195.
116. Stam, R., *et al.*, Identification and Characterisation CRN Effectors in *Phytophthora capsici* Shows Modularity and Functional Diversity. *PLoS One*, 2013. 8(3): p. e59517.
117. Du, J. and V.G. Vleeshouwers, The do's and don'ts of effectoromics. *Methods Mol Biol*, 2014. 1127: p. 257-68.

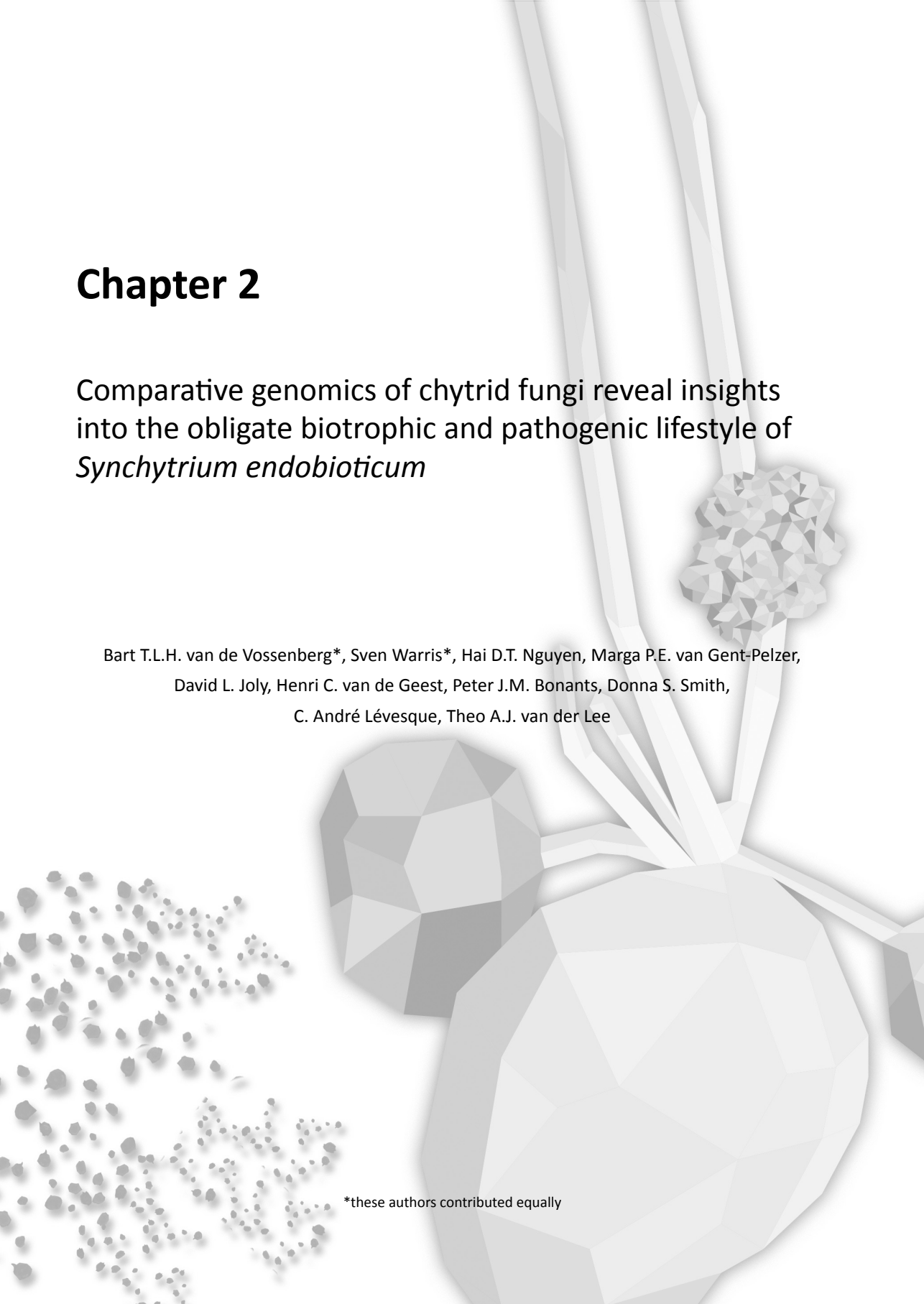


Chapter 2

Comparative genomics of chytrid fungi reveal insights into the obligate biotrophic and pathogenic lifestyle of *Synchytrium endobioticum*

Bart T.L.H. van de Vossenberg*, Sven Warris*, Hai D.T. Nguyen, Marga P.E. van Gent-Pelzer, David L. Joly, Henri C. van de Geest, Peter J.M. Bonants, Donna S. Smith, C. André Lévesque, Theo A.J. van der Lee

*these authors contributed equally



Chapter 2

This chapter is currently under review at Nature Scientific Reports with additional supplementary materials. Slight formatting changes were made to facilitate thesis consistency.

Abstract

Synchytrium endobioticum, the species causing potato wart disease, has a world-wide quarantine status and is included on the HHS and USDA Select Agent list. This chytrid species is the most basal lineage among the fungal plant pathogens and is the first available genome of an obligate soilborne fungal species.

We have chosen a functional genomics approach exploiting knowledge acquired from other fungal taxa and compared this to several saprobic and pathogenic representatives of Chytridiomycota (chytrids). Observations linked to chytrid biology, obligate biotrophy, genome plasticity and pathogenicity are reported. Essential purine and pyrimidine pathway genes were found uniquely absent in *S. endobioticum*, suggesting that *S. endobioticum* relies on scavenging of purine and pyrimidine bases from its host for survival. Chytrid genomes were small, gene dense and intron-rich, but not protected for genome duplications by repeat induced point mutation. Both pathogenic chytrids *Batrachomyces dendrobatidis* and *S. endobioticum* contained the largest amounts of repeats, which is often linked to pathogenicity. We identified *S. endobioticum* specific candidate effectors that are associated with repeat-rich regions. They share a highly conserved motif, and show isolate specific duplications. A reduced set of cell wall degrading enzymes, and LysM protein expansions were found in *S. endobioticum*, which may prevent triggering plant defense responses as a result of damage- or pathogen-associated molecular patterns.

Our study underlines the high diversity in chytrids compared to the well-studied Ascomycota and Basidiomycota, and reflects characteristic biological differences between the phyla. Moreover, it shows commonalities in genomic features among pathogenic fungi.

Background

Species of the Chytridiomycota, also called chytrids, represent a basal lineage of fungi that arose in the Mesoproterozoic Era about 1,000 to 1,600 million years ago [1]. Most are described as free-living saprophytes, inhabiting aquatic and terrestrial environments, with some species being notorious pathogens. The knowledge of the role of chytrids in ecosystems is sparse and fragmented [2], but recent studies indicate they may dominate ecological niches with low nutrient availability such as periglacial soils [3]. Despite being ubiquitous in nature, only of a few genomes of the entire chytrid phylum have been sequenced and studied so far. Here we present the first comprehensive functional comparative genomic study of fungi belonging to Chytridiomycota.

There are approximately 1,500 formally described chytrid species which is likely an underestimation as the relatively well-studied genus *Synchytrium* alone contains over 200 described species [4]. Most of the *Synchytrium* species are obligate biotrophic plant pathogens but recently the first saprobic free-living *Synchytrium* species, *Synchytrium microbalum*, was isolated and reported [5]. One of best studied species in the genus is *Synchytrium endobioticum* (Schilb.) Percival, the causal agent of potato wart disease, which is believed to originate from the Andes. It spread to the United Kingdom in the late 19th century when breeders were in search of new potato varieties following the great Irish potato famine. From Europe potato wart spread and today it is reported from all continents with potato cultivation. Isolates are currently grouped as pathotypes (races) based on their virulence on a differential set of potato varieties, reviewed by [6].

Like other chytrids, *S. endobioticum* does not produce hyphae or a mycelium often considered to be characteristic for fungi, but rather sporangia with motile zoospores. Upon infection, host cells are triggered to grow uncontrollably resulting in warted tissues, rendering tubers unmarketable [8]. Thin-walled, short-lived summer sporangia are formed in infected potato tissue and give rise to zoospores that can infect epidermal cells of potato in growing sprout or stolon tissues. The zoospores have been reported to conjugate possibly resulting in a (para)sexual cycle generating thick-walled resting sporangia that are released into the soil from decomposing warted potato tissue. Zoospores released from these resting sporangia can start a new infection cycle on potato tubers [7]. Because of the lack of effective chemical treatments and the longevity of resting spores, potato wart disease is one of the most important quarantine diseases on cultivated potato [9, 10].

So far, studies on *S. endobioticum* focused on its life cycle, epidemiology and management of the pest, and more recently, molecular tools for detection, identification and characterization of isolates have been published, reviewed in [6]. However, still very little is known about the molecular mechanisms underlying the obligate biotrophic or pathogenic lifestyle of this pathogen. To identify genes linked to obligate biotrophy, cell wall degradation, pathogenicity and sexuality of chytrid fungi, we assembled the genomes of two *S. endobioticum* isolates from different geographic origins and of different pathotypes, and compared those to nine culturable chytrid isolates, including the amphibian pathogen *Batrachochytrium dendrobatidis*, and six fungal organisms from other phyla.

Results

Genome assembly and annotation We independently sequenced and assembled the genomes of *S. endobioticum* pathotype 1(D1) isolate MB42 and pathotype 6(O1) isolate LEV6574, originating from the Netherlands and Canada respectively (Fig. 1). Although we purified the sporangia from infected potatoes as much as possible by filtration and centrifugation, the obligate biotrophic nature of *S. endobioticum* implies that our assembled sequences contained genomic information from *S. endobioticum*, the potato host, as well as microbial contaminants. Identification of *S. endobioticum* specific sequences was achieved using a comparative read-mapping approach which resulted in assembly sizes of 21.48 and 23.21 megabases (Mb), for MB42 and LEV6574 respectively (Figs. S1-S3). Both genomes were highly similar with 98.71% average nucleotide identity (Fig. S4) and encoded 8,031 and 8,671 protein coding genes for MB42 and LEV6574 respectively. To conduct additional comparative analyses, we sequenced, assembled and annotated the genomes of four culturable chytrid species: *Chytridium confervae* CBS 675.73, *Powellomyces hirtus* CBS 809.83, *Spizellomyces palustris* (= *Phlyctochytrium palustre*) CBS 455.65, and *Synchytrium microbalum* JEL517. Species identity was confirmed using rDNA sequences (NCBI: MH660417- MH660420). Additional data from five publicly available chytrid genomes were included, namely, *Batrachochytrium dendrobatidis* JAM81 and JEL423, *Gonapodya prolifera* JEL478, *Spizellomyces punctatus* DAOM BR117 and *Homolaphlyctis polyrhiza* JEL142 (Tables S1 and S2). Based on conserved eukaryotic genes, we obtained 84.1% to 97.2% annotation completeness for the chytrid genomes analyzed based on conserved core fungal single-copy genes [11] (Table S3).

Synchytrium endobioticum

Isolate MB42
 Pathotype 1(D1)
 21.48 Mb
 786 contigs
 8,031 genes

Synchytrium endobioticum

Isolate LEV6574
 Pathotype 6(O1)
 23.21 Mb
 1,285 contigs
 8,671 genes

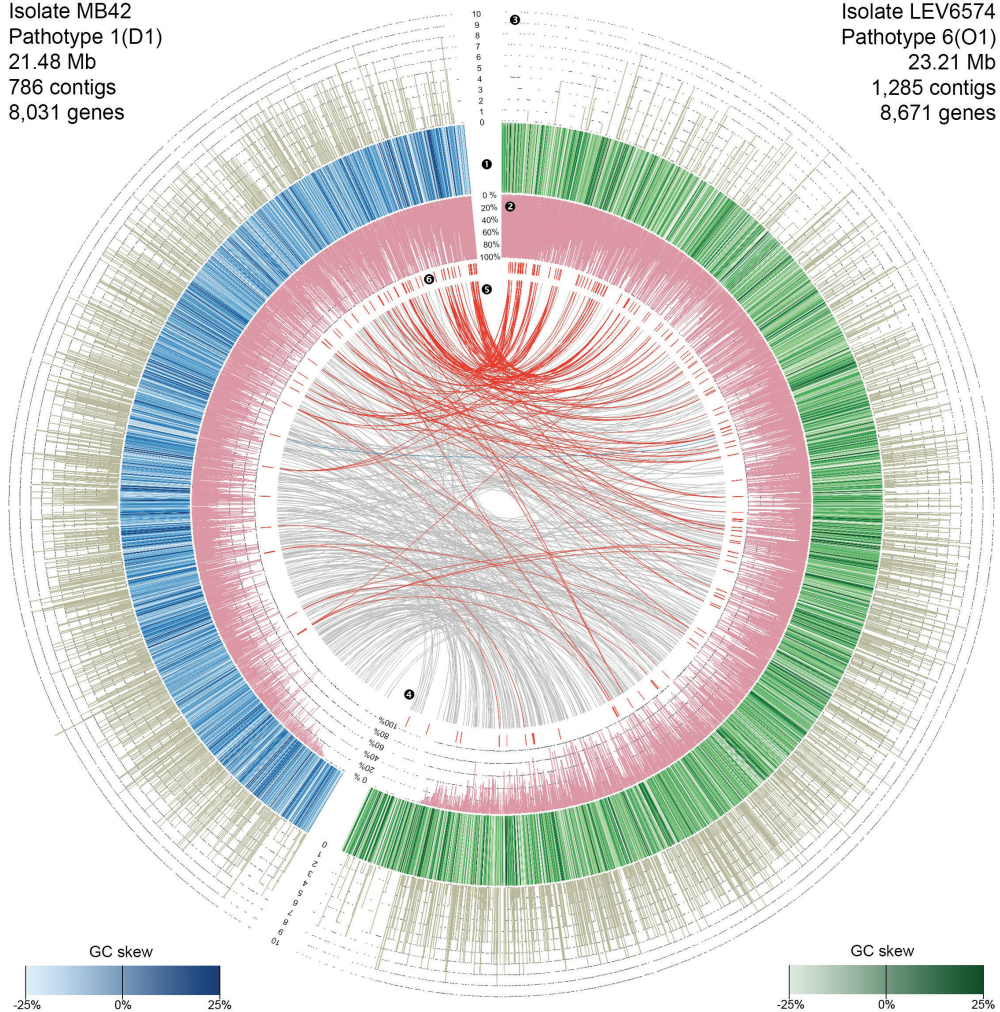


Figure 1. Circos plot representing the genomes of *S. endobioticum* isolates MB42 (1, blue) and LEV6574 (1, green). Different shades of green and blue used indicate the GC skew $((G-C)/(G+C))$ for the respective scaffolds and contigs. Purple annotations (2) represent the repeat content given in percentage of 1kb windows. Number of GO-terms predicted for each gene in the *S. endobioticum* genomes are shown in 3. Internal links in grey (4) represent bidirectional best hits between both isolates of single copy orthologs from the chytrid core genes. Red internal links (5) represent COG membership of candidate effector genes indicated with 6. The contigs and scaffolds of each genome are sorted by total repeat content. Key observations are: both genomes are gene dense and genes across the genomes are annotated (3). No obvious GC-content clustering (1) is observed, but both genomes contain many repeats (2). The LEV6574 genome is larger on account of these repetitive sequences. Candidate effector genes (6) are strongly associated to genomic regions with high repeat content (3).

Phylogenomics

To verify and further specify the phylogenetic position of the chytrid species analyzed, we performed a maximum likelihood phylogenetic analysis, with 47 fungal isolates representing the major fungal phyla (Table S4). Based on a trimmed amino acid alignment containing 54,485 positions extracted from 192 conserved eukaryotic genes as defined by Spatafora *et al.* [12]. This concatenated tree (Fig. 2) shows 100% bootstrap support for all major phyla, as treated by Spatafora *et al.* [12], except for Mucoromycota. There is strong support for the three classes Monoblepharidomycetes, Neocallimastigomycetes and Chytridiomycetes. The ASTRAL greedy consensus tree (Fig. S5), inferred from the combined phylogeny from each individual alignment, showed similar support for the major phyla, again with the exception of the Mucoromycota, and with a few inconsistencies in the backbone of the tree above the class level. The topologies of both trees are identical within the Chytridiomycetes confirming the placement of *S. endobioticum* within this class. The genus *Synchytrium* is currently classified in the order Chytridiales. However, our phylogenomic analysis shows that *Synchytrium* species form a distinct clade separated from members of the Chytridiales, and supports the potential transfer of the genus *Synchytrium* to Synchytriales, an order previously erected by Doweld [13] and emended by Longcore *et al.* [5]. Our data further underlines the high diversity in the Chytridiomycota compared to the well-studied Ascomycota and Basidiomycota.

Chytrid biology and obligate biotrophy

To identify genomic features in *S. endobioticum* that could explain its dependency on the host for survival, we performed a search of all predicted proteins against the Kyoto Encyclopedia of Genes and Genomes (KEGG). Functional annotations of *S. endobioticum* genes were compared with those of nine other chytrid species, and representative Ascomycota and Basidiomycota of which three are axenically culturable (*Saccharomyces cerevisiae*, *Neurospora crassa*, and *Cryptococcus neoformans*), and two are obligate biotrophic (*Melampsora larici-populina*, and *Puccinia graminis* f.sp. *tritici*), and one is obligate biotrophic in its diploid pathogenic phase (*Ustilago maydis*). We further performed a functional classification of genes using a Gene Ontology (GO-)term analysis. Central in the latter analysis were orthologous genes shared by all chytrid species as well as those unique to, or absent in *S. endobioticum*. Additionally we performed an analysis of carbohydrate-active enzymes (CAZymes) in *S. endobioticum*, as the ability to use different types of carbohydrates is an important feature of plant pathogens.

Clusters of orthologous genes On average, 87.9% of all proteins were assigned to one of 11,462 clusters of orthologous genes (COG) with more than one member. Both pathogenic species, *S. endobioticum* and *B. dendrobatidis*, have the highest number of species specific COGs compared to other chytrid species with similar gene content (Fig. 3; Table S5). In total 1,848 COGs were shared by all chytrid species, referred to as “chytrid core”, containing 694 single copy orthologs (SCOs). The second and third largest intersections of COGs were specific to both pathogenic species: 1,413 COGs with 1,179 SCOs for *S. endobioticum*; and 1,335

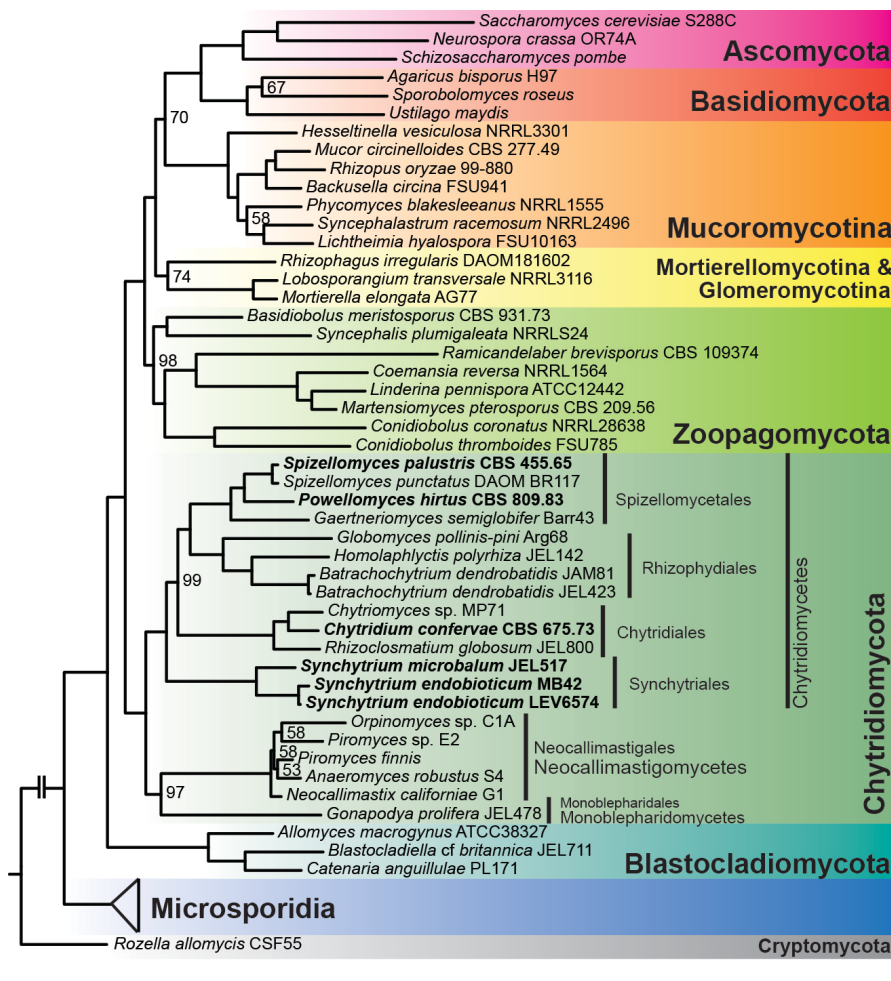


Figure 2. Phylogenomic maximum likelihood phylogenetic analysis of 47 fungal isolates. The analysis is based on a trimmed amino acid alignment containing 54,485 positions extracted from 192 genes. Genomic sequences assembled and annotated in this study are highlighted in bold. Bootstrap values other than 100% are shown on the nodes.

COGs with 1,066 SCOs for *B. dendrobatidis*. We identified 76 COGs present in all culturable chytrid species, which were absent in *S. endobioticum*. For both *S. endobioticum* isolates, a similar percentage of all genes were assigned to COGs or the chytrid core. However, the LEV6574 genome had a significantly higher percentage of genes assigned to the *S. endobioticum* specific COGs (LEV6574: 22.7%, MB42: 14.6%, two-sample binomial test, $p < 0.001$).

Functional annotation GO-terms identified in chytrid species were compared to those identified in the six fungal organisms from other phyla. Terms linked to the presence of whiplash flagella were found exclusively in the chytrid genomes, i.e. “motile cilium”; “cilium”; “ciliary basal body”; “ciliary transition zone”; “MKS complex”; “outer dynein arm”; and “axonemal dynein complex”. Linked to cell motility, the biological processes term “cilium movement” was found exclusively in the chytrid genomes (Fig. S6, Table S6). Similarly, the term “chitinase activity” was found exclusively in all studied chytrid species. To further understand the lifestyle of *S. endobioticum*, a level 2 GO-term analysis was performed on

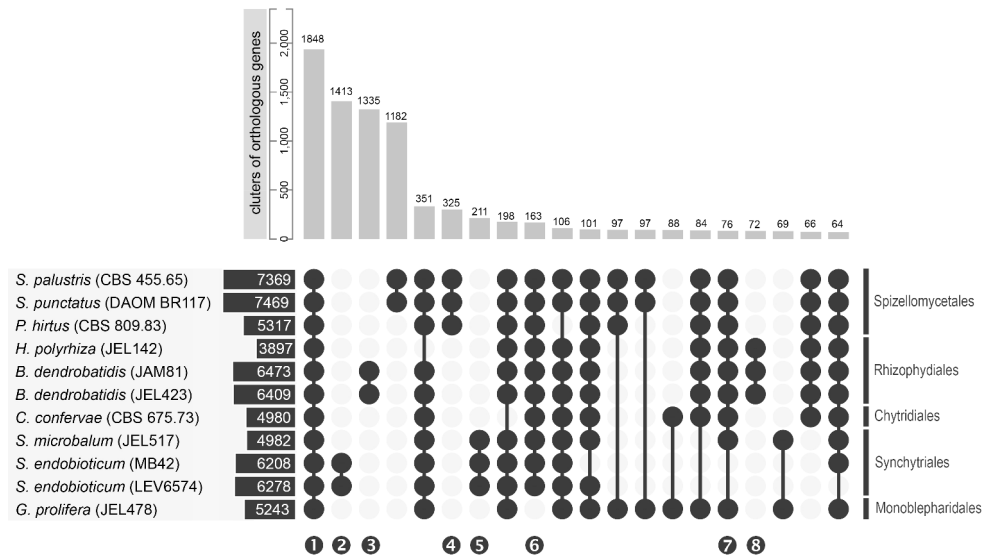


Figure 3. Protein families (COGs) predicted for the chytrid species analyzed (sorted based on taxonomical classification). Horizontal bars indicate the number of COGs to which protein sequences are assigned for the respective isolates. Vertical bars indicate the number of COGs in intersecting selections. The number of COGs for selected intersections, indicated with connected dark grey circles, are shown sorted from those with the most abundant to the least abundant ones for intersections with >60 COGs in the intersection. Selected intersections are highlighted: **1** shared by all chytrid species (chytrid core), **2** *S. endobioticum* specific, **3** *B. dendrobatidis* specific, **4,5,8** order specific, **6** Chytridiomycetes (class) specific, and **7** present in all chytrids but absent in *S. endobioticum*.

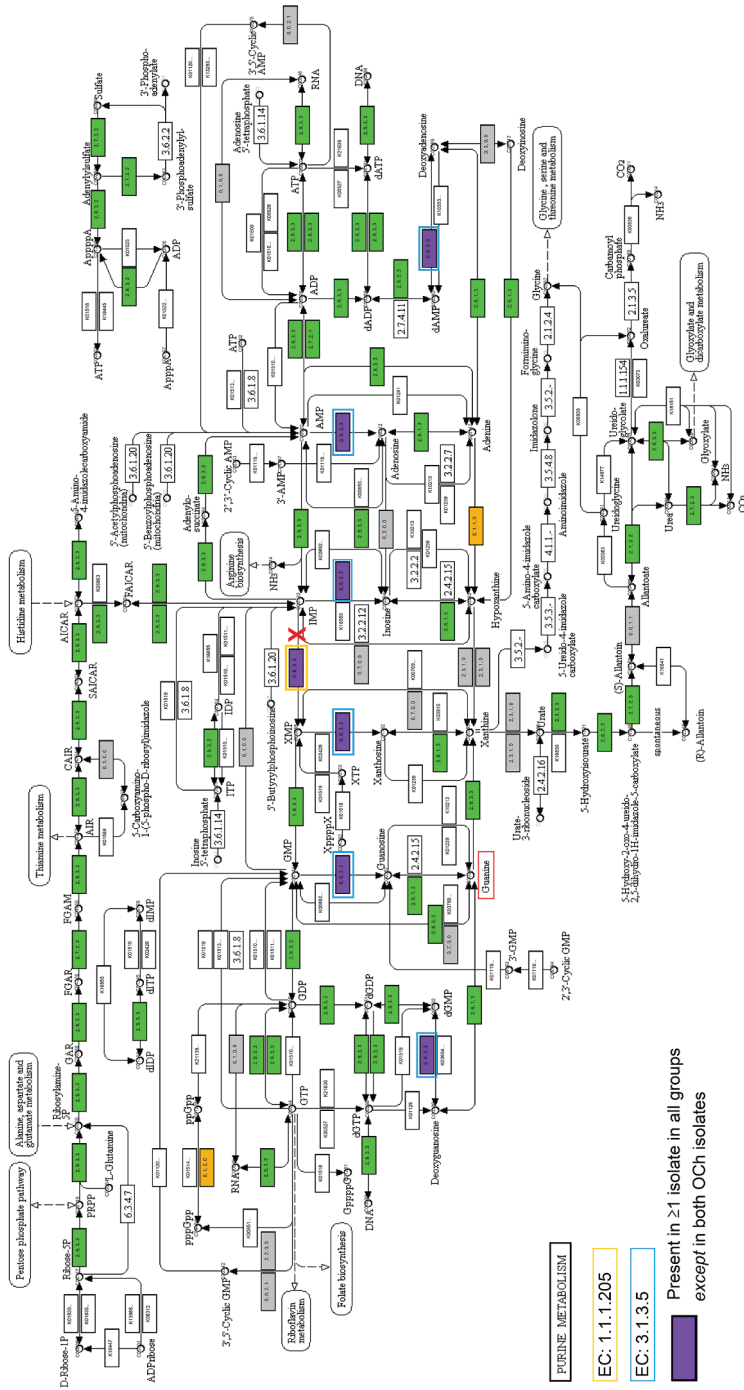


Figure 4. KEGG reference pathway 230: Purine metabolism. Numbers (a, b, c, d) in the enzymatic steps indicate the number of isolates in a given group for which the corresponding gene was detected (*S. endobioticum*; culturable chytrids; culturable higher fungi; obligate biotrophic higher fungi). Colors indicate which groups are represented for a given enzymatic step: present in ≥1 isolate in all groups except (purple); present in ≥1 isolate in all groups (green); present in ≥1 obligate biotrophs (orange), but not in the obligate biotrophs (white). The red “X” indicates the pathway for *de novo* Guanine triphosphate (GTP) synthesis that is blocked in *S. endobioticum*. Boxed in orange is inosine-5'-monophosphate dehydrogenase (EC:1.1.1.205), and boxed in blue is guanine that could be scavenged by *S. endobioticum* from its host which were both found to be present in all genomes analyzed except in *S. endobioticum*. Boxed in red is guanine that could be scavenged by *S. endobioticum* from its host allowing synthesis of GTP.

the COG intersections “*S. endobioticum* specific genes”, and “genes present in culturable chytrid species but absent in *S. endobioticum*” (Fig. S7). Compared to the genes in the COG intersection “chytrid core”, *S. endobioticum* specific genes show an increase of genes associated with signal transduction. In the underlying terms, no specific terms unique to *S. endobioticum* were found, but for some functions, an increased number of proteins with that particular function was observed, namely, “small GTPase mediated signal transduction” and “G protein coupled receptor signaling pathway”. In the COG intersection “genes present in all chytrid species but absent in *S. endobioticum*”, GO-terms for “methylation”, “regulation of biological quality”, and “cellular component organization” were more abundant compared to the chytrid core. However, most of the lower-level terms that were found in these genes were also found in genes belonging to the chytrid core, indicating that the other chytrid species simply have more genes with the same functionality. One term was an exception: “regulation of mitochondrial membrane permeability involved in programmed necrotic cell death”, which was specifically absent in *S. endobioticum*.

KEGG pathways We benchmarked the analysis using the six fungal organisms from other phyla with known biochemical pathways included in the KEGG database, demonstrating an overall accuracy of 84.5% (Table S7, Supplementary file 4), and subsequently we compared these to the functional annotation of the chytrid genomes. InterProScan function assignment was found to be accurate with a high percentage of specificity (average = 93%). However, we found that the KEGG pathway analysis was particularly prone to false negatives (Figs. S8 and S9) resulting in incomplete function prediction. Therefore, absence of specific genes was verified with tblastn. On average, 7.3% of the genes in the chytrid genomes were assigned to a KEGG pathway, compared to 5.9% in the fungal organisms with known KEGG pathways. For both *S. endobioticum* isolates, 6.2% of the gene models were assigned to a KEGG pathway (Table S8). Enzymes that were absent in *S. endobioticum* but were present in the majority of proteomes of the other species analyzed fell in 26 of 72 reference biochemical pathways. After tblastn verification, 22 enzymes were found to be absent in 18 biochemical pathways (Table S9; Supplementary file 3). The absence of two enzymes in the purine metabolism pathway, i.e. inosine-5'-monophosphate dehydrogenase (IMPDH, EC:1.1.1.205) and pyrimidine 5'-nucleotidase (EC:3.1.3.5), could provide further insight on the obligate biotrophic lifestyle of *S. endobioticum* (Fig. 4). The latter enzyme also plays a role in the pyrimidine pathway. Unique to the chytrid species is uracil phosphoribosyltransferase (EC:2.4.2.9) which catalyzes the reaction from UMP to uracil in the pyrimidine pathway (Supplementary file 3, Pyrimidine pathway).

CAZymes Altogether 246 different CAZyme modules were detected in the chytrid species and the six fungal organisms from other phyla. Six families were only found in chytrids (GT92, GH46, GT14, PL14, AA1 subfamily 1, GH5 subfamily 27); a single family (GT67) was found to be unique to *S. endobioticum*; and 29 families were absent in *S. endobioticum* but present in all species analyzed (Table S10). The Carbohydrate-Binding Module Family 35 (CBM35) family is present in all obligate biotrophic species analyzed but missing in the culturable species. In *S. endobioticum*, genes with CBM50, also known as LysM, were found to be enriched compared to the other species included in the analysis (means: 22 to 7, T-test: $p < 0.001$). Similarly, CBM18 was found enriched in *B. dendrobatidis* (means: 83 to 8; T-test: $p < 0.001$). Cluster analysis based on all assigned CAZyme modules grouped the two *S. endobioticum* isolates and also the closely related species *S. microbalum*, *P. hirtus*, *S. punctatus* and *S. palustris* group into a clade (Fig. S10). The two *B. dendrobatidis* isolates, causal agents of amphibian chytridiomycosis, also clustered together but were clearly an outgroup compared to the other fifteen species. The three other obligate biotrophic species (*U. maydis*, *M. larici-populina*, and *P. graminis* f. sp. *tritici*) grouped together but were distant from the two *S. endobioticum* isolates.

Genome plasticity

Genome plasticity is an important feature of fungal plant pathogens, and is often facilitated by repeats. We therefore studied the repeat content of all chytrid genomes, and analyzed if the repeat induced point mutation (RIP) mechanism [14], which protects fungal genomes against duplications and transposable elements, was active in chytrid species. As a sexual cycle has great impact on genome dynamics, we further performed an *in silico* analysis for the potential of sexual reproduction in chytrids.

Repeat content and transposable elements Together with the amphibian pathogen *B. dendrobatidis*, *S. endobioticum* displayed the highest percentage of repetitive sequences in its genome (Table S12). The repeat content could be split in small simple repeats (SSR) and complex repeats. In total, 399 complex repeat families were identified of which 15 were shared by all chytrid genomes analyzed. For the culturable chytrid species the repeat repertoire mainly consisted of small simple repeats (64.1 to 84.4%). In contrast, the repeat repertoire of both pathogenic species mainly consists of complex repeats (77.6 to 89.9%). Over a third of all repeat families was unique to the two *S. endobioticum* isolates, and *S. endobioticum* has significantly the most diverse repeat repertoire (ANOVA, $p < 0.001$)

compared to other chytrid species (Fig. S11). Both Class I (retro) and Class II (DNA) transposable elements were identified in the two *S. endobioticum* genomes. Most repeats could not be classified to any known element. For the elements that could be classified, Gypsy and Copia retrotransposons were most abundant. These are also the dominant transposable elements in other chytrid species. Unique to *S. endobioticum* is the non-LTR retrotransposon Tad1 (Fig. 5). In the pathogenic *B. dendrobatidis*, repeats were clustered in repeat-rich compartments. In contrast, the repeats were found to be dispersed over the entire genome in both *S. endobioticum* isolates (Fig. 1).

RIP activity RIP activity can be observed by a bimodal distribution of GC content as a result of CpA to TpA transitions. Such a distribution was not observed for both *S. endobioticum* genomes, nor for any other chytrid genome indicating that no apparent RIP affected regions were present in these genomes (Fig. S12). To further determine if repeat families in *S. endobioticum* were specifically affected by RIP, the 10 most abundant species specific repeat families matching the prerequisites for RIP (400 bp repeats [15] with $\geq 80\%$ sequence similarity [16] in *N. crassa*) were subjected to the alignment based RipCal analysis.

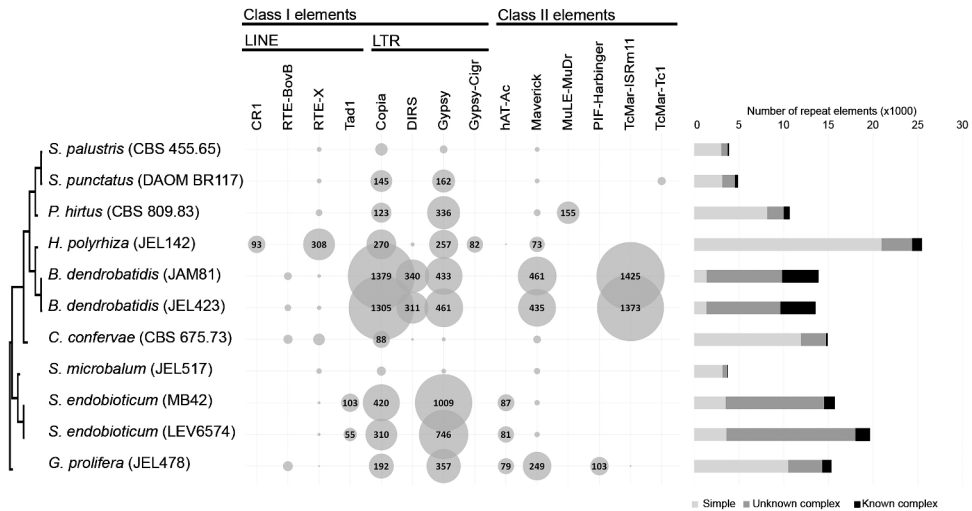


Figure 5. Repeat and transposable element (TE) content in chytrid species analyzed. Chytrid isolates are sorted based on taxonomical classification, and spheres indicate the number of class I and II TEs were identified in the individual isolates. Horizontal bars indicate the number of repeat elements per genome, which is further differentiated as small simple repeats (light grey), unknown complex repeats (dark grey), and complex repeats with TE identity (black).

Chapter 2

For some repeat families, the copy number of (near) complete gene models represented was similar for both isolates, while for others expansions in one of the isolates were observed (Figs. S13 and S14). RIP signatures were not detected in any of the analyzed alignments. In addition, the protein essential for RIP, RID (RIP defective) [17], could not be detected in the *S. endobioticum* isolates using blastp or tblastn. RID exhibits methyltransferase activity and belongs to the protein family of C-5 cytosine methyltransferases. All C-5 cytosine methyltransferases contain a motif (motif VI) with two conserved amino acids (NV: asparagine-valine), except for RID and the gene associated with Methylation Induced Premeiotically (MIP: another fungal defense mechanism against TEs) Masc1 [18]. In the latter two proteins NV is replaced with either QT or ET. Thirty-seven C-5 cytosine methyltransferase proteins were identified in the chytrid species analyzed. The motif VI sequences of these proteins did not possess the consecutive QT or ET amino acid sequences associated with RID and Masc1 (Fig. S15), which confirmed that RID is not present in the analyzed chytrids.

Sexual cycle GO-terms linked to meiosis were identified in all chytrid genomes (Fig. S16). Subsequently, we investigated if their genomes possess the genes required for meiosis referred to as the meiotic toolbox. A set of 86 genes required for successful meiotic recombination in *S. cerevisiae* and *C. neoformans* was used to identify orthologs in Chytridiomycota. Of this set, 31 are regarded core meiotic proteins with ten considered as meiosis specific proteins [19]. Homologs were found in one or more chytrid species for 90% of the meiotic core proteins (Table S11). Verification by tblastn was performed for the meiotic core and meiosis-specific proteins to determine if they were truly absent. Of the ten meiosis specific genes, only six were detected in all chytrid species analyzed, and three were detected in all but one or two isolates. One (*Rec8*) could not be detected using the *S. cerevisiae* protein, while it was detected in six of the chytrid fungi analyzed using the *Allomyces macrogynus* homolog. In contrast, the mitotic homolog of *Rec8* (*Rad21*) was detected in all chytrid species analyzed. For the remaining 21 meiotic core genes, homologs were identified in all chytrid species for 16 genes. Similar to the findings of Halary *et al.* [19], we were unable to identify a homolog of *Scs3* in any of the chytrid species. For the remaining four genes (*Mlh2*, *Mlh3*, *Mus81*, *Rad51*), a patchy distribution was obtained with the detection of homologs in four or more chytrid species. Clearly some genes from the meiotic toolbox are absent, however, the consequences for meiosis are yet unclear.

Pathogenicity

Plant pathogens facilitate host colonization by manipulating the plant or its microbiome by secreting effector proteins which are typically species or even isolate specific and may contain specific motifs, reviewed by Selin *et al.* [20]. We analyzed the secretomes of both *S. endobioticum* isolates to identify candidate effector protein that could act as elicitors during host infection.

Candidate effectors For both pathogenic species, *S. endobioticum* and *B. dendrobatidis*, the majority of secreted proteins is species specific. Screening of the *S. endobioticum* secretome, one particular 15 amino acid motif was found: i.e. RAYHxxVFExLKxLF, which we refer to as the RAYH-motif (Fig. 6). Proteins with the RAYH-motif were found exclusively in *S. endobioticum* with 148 and 75 proteins for the isolates LEV6574 and MB42 respectively. The majority of these proteins had one occurrence of this motif, but 41 and 28 proteins for the respective isolates had two (Fig. S17; Table S14). Significant expansions were found between the two isolates (Fig. 1; Fig. S18). While many orthologous groups were found to be expanded in LEV6574 this was rarely the case for MB42. However, read coverage on these candidate effector genes indicates a higher copy number variation for MB42, suggesting that in some cases these genes were not resolved in the MB42 genome assembly (Fig. S19).

The RAYH-genes appear to be overrepresented in regions that show depletion of chytrid core genes (Fig. 7). Moreover, they were significantly closer to terminal ends of genomic scaffolds compared to chytrid core SCOs (Fig. S20), which is possibly a consequence of their association with repeats that break up the assembly. We observed several (near) identical members of these candidate effectors, and an alignment of the genes coding for RAYH-proteins revealed a conserved architecture (Fig. S21) with some gene models showing discrepancies. Interestingly, most genes that were not predicted to be secreted were among those showing additional or missing amino acids in the N-terminal region, which could indicate errors in the gene predictions or pseudogenized genes.

Apart from conservation in some structural features (intron-exon boundaries) and the motif, the proteins were highly variable. RAYH-proteins were found to be separated within 70 orthologous groups (41 when excluding singletons). This could indicate that they may have different functions but share a similar localization or translocation mechanism. We therefore further analyzed the possible localization of the proteins and compared proteins with this motif with chytrid core SCOs. Proteins carrying a mitochondrial targeting peptide (mTP)



Figure 6. Sequence logo of the *S. endobioticum* specific RAYHxxVFExLKxLF sequence motif. The motif is referred to as the “RAYH-motif” based on the first four conserved amino acids, found in 148 LEV6574 and 75 MB42 candidate effector proteins.

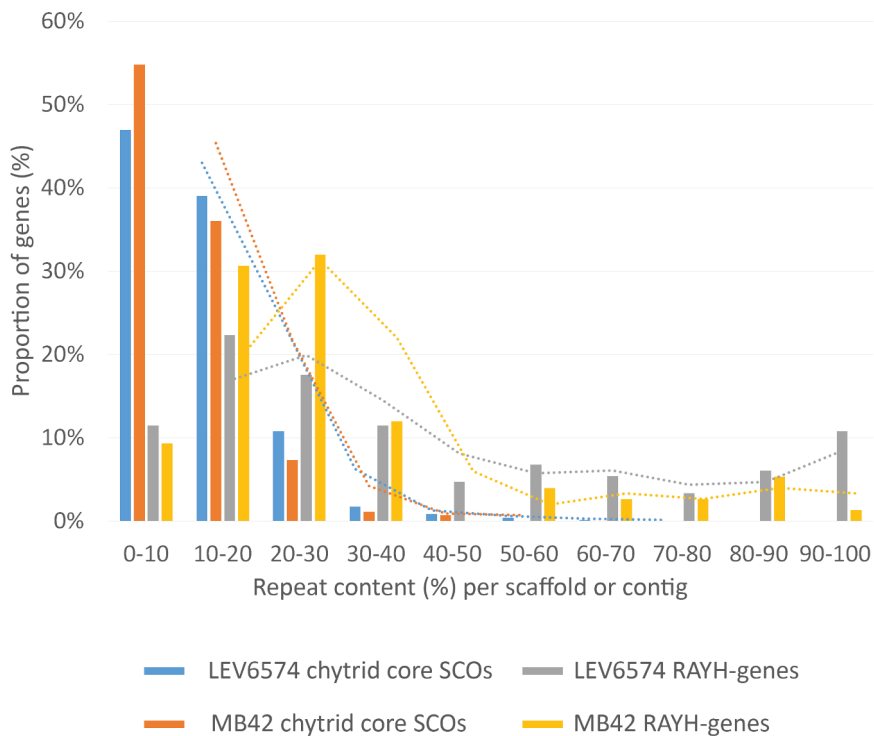


Figure 7. Occurrence of *S. endobioticum* chytid core SCO genes and RAYH-genes in function of scaffold repeat content percentages. The proportion of candidate effector or SCO genes are compared against the scaffolds where repeats occur. The *S. endobioticum* SCO chytid core genes are mainly found on scaffolds and contigs with 0 to 30% repeat content. Candidate effector genes are found mainly on scaffolds and contigs with 0 to 40% repeat content but also on genomic sequences with up to 100% repeat content.

and/or a nuclear localization signal (NLS) were found to be overrepresented in secreted proteins having the RAYH-motif. In LEV6574, mTP was found in 39.1% of secreted proteins carrying the motif, versus 12.6% when considering the whole secretome. Similarly, a NLS was found in 43.4% of secreted proteins harboring the motif, versus 24.1% of the whole secretome. In many cases, more than one targeting peptide or motif was present (e.g. 16.3% of motif-containing proteins had both mTP and NLS, compared to 3.6% for the whole secretome). No significant differences were observed when comparing the occurrence of the chloroplastic transit peptide.

Discussion

We assembled and annotated the genomes of two *S. endobioticum* isolates representing different pathotypes from different continents. As this is an obligate pathogen, the DNA and RNA had to be extracted from resting spores purified from infected potatoes. Apart from the fungal pathogen and its host, these nucleic acid extracts also represented other microbial soil organisms. The comparative read-mapping approach used to identify sequences of the target species was successful in removing contaminant sequences and resulted in genome sizes with a gene content similar to other chytrid fungi. A high level of annotation completeness was obtained as indicated by the BUSCO scores obtained from the predicted proteins. From our phylogenomic analysis, we get very strong support for placement of the *Synchytrium* species in the order Synchytriales.

Genes shared by all chytrid species (chytrid core), specific to *S. endobioticum* and present in all chytrid species except *S. endobioticum* were identified. We postulate that the chytrid core genes reflect shared biological features and processes of species in the Chytridiomycota. On the other hand, genes specific to *S. endobioticum* are hypothesized to hold the key to pathogenicity while genes absent in *S. endobioticum* but present in other chytrids could potentially explain the obligate biotrophic lifestyle of the pathogen. Functional annotation and analysis of these genes resulted in several important observations.

Chytrid biology and obligate biotrophy

Motility Our functional analysis reflects the presence of flagella and associated motility in Chytridiomycota, which are absent in Ascomycota and Basidiomycota. Motility and sensing of the environment using G coupled receptors reflects a fundamental difference between Chytridiomycota and Ascomycota and Basidiomycota. The lack of a self-directed motile life

stage makes ascomycetes and basidiomycetes dependent on passive spore dispersal or physical growth, resulting in sometimes extreme body sizes such as found for *Armillaria mellea* [21]. The motile life stage of Chytridiomycota may allow them to be more selective and grow more efficiently (energy-wise) in a targeted (pathogenic) lifestyle or under low nutrient conditions. This could perhaps explain why chytrids dominate periglacial soils [3] or are well adapted for a pathogenic lifestyle in which a self-directed motile, search life stage could be of critical importance. Associated with movement is taxis, the directed movement of a motile cell in response to an external stimulus. Both pathogenic chytrids, *S. endobioticum* and *B. dendrobatidis* as well as *H. polyrhiza* possess genes linked to chemotaxis. A distinct phylogenetic group, the oomycetes, share a self-directed motile life stage and include mostly plant and animal pathogens, such as *Phytophthora* and *Saprolegnia* species.

Obligate biotrophy *S. endobioticum* lacks genes encoding the enzymes IMPDH and pyrimidine 5'-nucleotidase as opposed to chytrids with a saprophytic lifestyle (including *S. microbalum*) where these enzymes are present. These genes encode enzymes that are essential in the purine and pyrimidine pathways. IMPDH irreversibly catalyzes the conversion of inosine monophosphate (IMP) to xanthosine monophosphate (XMP), which serves as an intermediate for the nucleotide guanosine monophosphate (GMP). This indicates that *S. endobioticum* is not able to synthesize Guanosine 5'-triphosphate (GTP) from IMP *de novo*, and therefore would rely on scavenging guanine to synthesize GMP which is converted to GDP and GTP. Guanine nucleosides are crucial prerequisites for many cellular functions including transmembrane and intracellular signaling, DNA replication, and RNA and protein synthesis. This is demonstrated by null mutants of *Saccharomyces cerevisiae* for all four IMPDH orthologs. The simultaneous deletion of these genes was lethal unless growth media were supplemented with guanine [22]. Lack of the pyrimidine 5'-nucleotidase restricts the catalysis of phosphorylytic cleavage of ribonucleotide monophosphates GMP, AMP, IMP and XMP to their respective nucleosides guanosine, adenosine, inosine and xanthosine in the purine pathway and of UMP, CMP and TMP to uridine, cytidine and thymidine in the pyrimidine pathway. To create the before mentioned nucleosides, we hypothesize that *S. endobioticum* relies on scavenging of the purine and pyrimidine bases from its host for survival, as was previously described for pathogenic Mycobacteria [23] which also cannot be cultured axenically.

CAzymes The major polysaccharides present in plant cell walls are cellulose, hemicellulose and pectin. The CAzymes that are involved in degradation of these molecules are referred to as cell wall degrading enzymes (CWDE). We split the CAZyme families into several functional categories following Kubicek *et al.* [24]. Overall, *S. endobioticum* has the ability to process complex sugars, which could include cellulose and starch. It can degrade only some hemicellulose because it is lacking the alpha-galactosidases, beta-glucuronidases and alpha-arabinosidases. As for pectin degradation, it is not as clear because only a single family (GH28) with one copy was found in both *S. endobioticum* isolates. We noted however that there were discrepancies between the CAzyme analysis results, and results linked to CAzymes obtained with KEGG and GO-terms (Fig. S14). For its energy needs, *S. endobioticum* has the ability to process starch and fructose and it is therefore not directly dependent on the degradation of cellulose (Supplementary file 3, starch and sucrose metabolism pathway).

Genome plasticity

Repeats, TE, and RIP The *S. endobioticum* assemblies are relatively fragmented compared to other genomes assembled in this study which is a direct result of the repeat content of the *S. endobioticum* genome. Both *S. endobioticum* isolates, together with *B. dendrobatidis*, have the highest repeat content of all chytrid species analyzed. For both species the repeat content mainly comprises complex repeats, opposed to small simple repeats which are the predominant form of repeats in other chytrid species. The repeat repertoire and TE activity of filamentous fungi have been linked to pathogenic lifestyle in many cases [25]. In *U. maydis* for instance, repetitive compartments containing TEs were significantly associated with virulence gene clusters [26]. In the pathogenic *B. dendrobatidis*, repeats are clustered together in repeat rich compartments. However, in both *S. endobioticum* isolates the repeats are dispersed over the genome. The majority of the complex repeats could not be identified as either class I or II elements, suggesting that novel families of transposable elements could exist in Chytridiomycota. In the context of the other chytrids, the non-LTR retrotransposon Tad1, which was found to be active in a RIP-proficient isolate of *N. crassa* [27], was unique to *S. endobioticum*. The genomes of chytrids do not appear to be protected by RIP, other defense mechanisms against active TEs such as post-transcriptional gene silencing (quelling) [28] could play a role to restrict genome expansion of Chytridiomycota. Interestingly, the small genome of *S. endobioticum* contained many repeats and introns, and the condensed size of the genome is achieved by the combination of a relatively low number of genes and short intergenic spaces.

Sexual cycle A sexual cycle in Chytridiomycota was described from species in several genera, for instance *Synchytrium*, *Chytridium*, *Micromyces*, and *Rhizophydium* [29], but the occurrence of a sexual cycle is disputed for others including the best studied species *B. dendrobatidis* and *S. endobioticum*. The fusion of *S. endobioticum* zoospores, which were postulated to act as isogametes, and the formation of biflagellated zygote were observed as early as 1921 by Curtis [7]. Her observations on the fusion of zoospores were confirmed by several authors [30-34] but comprehensive electron microscopy studies were not able to provide ultrastructural evidence for karyogamy and the meiotic divisions that hallmark meiosis. Halary *et al.* [19] used the publically available genomes of *B. dendrobatidis* and *A. macrogynus* to define the presence of meiotic genes in Chytridiomycota. However, the latter species was previously included in Chytridiomycota but is now regarded to be a member of the Blastocladiomycota. This resulted in an inventory of chytrid meiotic genes which also included specific presence or absence of meiotic genes in Blastocladiomycota. The results from our study paint a more complete picture regarding the meiotic toolbox of chytrid species. However, because of the patchy presence and absence of some specific meiotic genes in the chytrid species analyzed, a conclusive answer regarding a sexual cycle in chytrids remains elusive. Another indicator of a cryptic sexual cycle, i.e. RIP activity [16], was not found in chytrids. Active Copia and Gypsy retrotransposons however, which were suggested to be a sign for a sexual cycle [35], were found in the chytrids analyzed.

Pathogenicity

Host recognition Our CAZyme analyses point to a reduced number of cell wall degrading hemicellulases in *S. endobioticum* compared to its saprobic sister species *S. microbalum*. Reduced CWDE content was previously found in the stealth like pathogen *Zymoseptoria tritici* (formerly known as *Mycosphaerella graminicola*) [36], and the biotrophic and symbiotic fungus *Lacaria bicolor* [37], which may reduce or prevent the triggering of defense responses as a result of damage-associated molecular patterns from cell wall degradation [38]. In *B. dendrobatidis*, genes encoding proteins belonging to the CBM18 family were found to be more abundant. Proteins from this family binds to the N-acetyl-D-glucosamine and sialic acid found in chitin and mucous membranes, respectively. In *S. endobioticum*, genes with CBM50, also known as LysM were found to be enriched. LysM containing proteins could be involved in chitin metabolism, or could have a function in the interference of chitin-triggered immunity, as was suggested for the endophytic fungus *Piriformospora indica* [39]. Also, the arbuscular mycorrhizal fungus *Rhizophagus irregularis* encodes LysM effectors

and candidate chitinases (which were also identified in the chytrid species analyzed) on its genome, and it is suggested that these mutualistic fungi remodel their cell wall during intracellular colonization [40]. Similar remodeling of cell walls could be a viable strategy for both pathogenic species *S. endobioticum* and *B. dendrobatidis* to avoid recognition by their host.

Candidate Effectors Effector proteins are expected to play a central role in the pathogenic lifestyle of *S. endobioticum* as they manipulate the host or its microbiome to facilitate infection. *S. endobioticum* has no mycelium and is intracellular. In some pathogenic groups, motifs such as the RxLR motif in oomycetes [41] have been identified in secreted proteins and are now routinely used to predict candidate effectors. We found a novel 15 amino acid motif, referred here as the RAYH-motif, and RAYH-proteins are regarded candidate effectors as they are species specific, secreted, associated with repeat rich regions and carry localization signals. The conservation in the gene architecture of genes carrying the motif is surprisingly high, and the number of members is numerous in the two *S. endobioticum* isolates.

Effectors from various pathogenic organisms are preferentially located in repeat-rich and gene-poor regions of the genome [42]. Although the RAYH-proteins were associated with repeat rich regions in general, no direct association with specific repeat content, as is the case in *Fusarium* species with two specific miniature inverted-repeat transposable elements (MITEs) [43], was observed. The specialization of pathogens is reflected by the fact that both pathogenic chytrid species have the highest species specific gene content. Also, the majority of the secretome was found to be species specific.

When intersecting RAYH-proteins with EffectorP, only a few members were predicted as candidate effectors (e.g. in LEV6574, only 3% of proteins with the motif were classified as effectors by EffectorP, while this was the case for 23% of the entire secretome). It is likely that the features of the chytrid RAYH-proteins are different from effector proteins from ascomycetes and basidiomycetes which were used to train the EffectorP tool. Also, the tool is optimized for recognition of putative apoplastic effectors. In the intracellular obligate biotrophic *S. endobioticum* we expect to find an abundance of cytoplasmic effectors. The RAYH-proteins could act as cytoplasmic effectors, which could explain why they were not detected by EffectorP.

Chapter 2

This conserved motif could be important in interacting with other proteins or lipids, to relocate these RAYH-motif proteins into or within the host cell. The motif was not found in other chytrid species including the saprobic sister species *S. microbalum*. *S. endobioticum*-specific secreted proteins with the RAYH-motif are diverse and carry further localization signals, such as mitochondrial targeting peptides or a nuclear localization signal. Mitochondrial targeting peptides could allow targeting to the plant mitochondria to enhance susceptibility, a strategy that is used by several intracellular bacteria (e.g. *Vibrio cholerae* [44], *Anaplasma phagocytophilum* [45] and *Legionella pneumophila* [46]). Besides energy production, mitochondria are also involved in cellular processes such as programmed cell death, calcium homeostasis, biosynthesis of amino acids, lipids and nucleotides, and innate immune signaling against viruses and bacteria [47]. Effectors with nuclear localization signals have been observed in bacteria and fungi [48, 49], and were shown to induce transcriptional reprogramming, reviewed in [50]. We hypothesize that for an obligate intracellular pathogen such as *S. endobioticum*, manipulating the potato host cell by targeting both the mitochondria and nucleus directly is a likely scenario.

Verification of the function of these proteins will require innovative approaches since no transformation protocols are described for *S. endobioticum* and that the current bioassays for pathotyping and virulence testing are cumbersome and time consuming.

Conclusions

This study is the first comprehensive comparative and functional genome analysis of chytrid fungi in general and *S. endobioticum* in particular, providing insights and understanding of fundamental biological questions in an economically important and a highly regulated plant pathogen. Particularly when working with obligate biotrophic organisms, the scientific community relies on the accuracy of computational tools for functional predictions to direct and design future experiments. The use of a redundancy in sequenced strains, combined with overlapping tools that are benchmarked, validated and verified, provide a robust basis for functional comparative studies of non-culturable organisms.

Two *S. endobioticum* isolates from different geographical locations representing different pathotypes were independently sequenced and functionally annotated. These were compared to the genomes of nine culturable chytrids, including the amphibian pathogen *B. dendrobatidis*, and six fungal organisms from other phyla of which three are axenically

culturable, and three are obligate biotrophic. Analysis of these genomes resulted in several important observations.

Chytrid lifestyle and obligate biotrophy Functional analysis reflects the presence of flagella and associated (directed) motility in Chytridiomycota, which are absent in Ascomycota and Basidiomycota. The motile life stage of chytrids may allow them to be more selective and grow more efficiently (energy-wise) in a targeted (pathogenic) lifestyle or under low nutrient conditions. *S. endobioticum* lacks essential genes in the purine and pyrimidine pathways as opposed to the other chytrids analyzed including the saprobic sister species of *S. endobioticum*. We hypothesize that *S. endobioticum* relies on scavenging of the purine and pyrimidine bases from its host for survival, as was shown for pathogenic Mycobacteria which also cannot be cultured axenically.

Genome plasticity Genome plasticity is often facilitated by repeats and mobile genomic elements, which have been associated with pathogens in many instances. Indeed, the pathogenic chytrid species *B. dendrobatidis* and *S. endobioticum* contain the largest amounts of complex repeats among chytrids. The chytrid genomes do not appear to be protected for genome duplications by RIP, yet they maintain their small genome size. As a sexual cycle has great impact on genome dynamics, we performed an *in silico* analysis for the potential of sexual reproduction in chytrids. Genes from the meiotic toolbox are partially present, and a conclusive answer regarding a sexual cycle in chytrids remains elusive. We do provide the most comprehensive inventory of meiosis related genes in chytrids, which will foster the research on a sexual cycle in chytrid species.

Pathogenicity Essential to pathogens is avoiding detection by the host, thus preventing triggering immune responses. A reduced set of cell wall degrading enzymes was found in *S. endobioticum* relative to its saprobic sister species, which may reduce or prevent the triggering of defense responses as a result of DAMPs from cell wall degradation. In addition, remodeling of cell walls has been reported from other pathogens to avoid recognition of PAMPs by the host. This could be a viable strategy for both pathogenic chytrids. Proteins containing LysM domains were found enriched in *S. endobioticum*, which could be involved in chitin metabolism, or could have a function in the interference of chitin-triggered immunity, as was shown for the arbuscular mycorrhizal fungus *Rhizophagus irregularis*. Effector proteins are expected to play a central role in the pathogenic lifestyle of *S. endobioticum* as they manipulate the host or its microbiome to facilitate infection. We identified a new class of

Chapter 2

candidate effectors specific to *S. endobioticum* that share a highly conserved RAYH-motif and gene architecture but also show isolate specific expansions and gene duplications.

In fungi, the focus for (genomic) research lies with representatives of Ascomycota and Basidiomycota. With this comparative study, we significantly increased the number of chytrid genomes and added the first economically important plant pathogen to the annotated Chytridiomycota genomes. They can now be included in comparative fungal genomic studies on a systematic basis as they provide a unique phylogenetic view and can shed new light on fundamental aspects of fungal biology.

Materials and Methods

Detailed descriptions on materials and methods used are provided in the supplementary materials and methods section. In brief, the genomes of two *S. endobioticum* isolates (MB42 and LEV6574) were sequenced from DNA and RNA extracted from resting spores. Additionally, cultures of *C. confervae* (CBS 675.73), *P. hirtus* (CBS 809.83), *S. palustris* (= *Phlyctochytrium palustre*) (CBS 455.65) and *S. microbalum* (JEL517) maintained on ARCH medium were sequenced. Sequence data was generated with one or more of the following techniques: Roche 454 GS-FLX Titanium, Illumina HiSeq, Illumina MiSeq, PacBio SMRT. A comparative read-mapping approach was used to identify *S. endobioticum* sequences from the metagenomic assemblies. Hybrid assemblies were obtained and RNAseq data was used in the structural annotation of the *S. endobioticum* genomes. Genomes were functionally annotated with InterProScan v5.1.6. [51] and data were deposited in NCBI GenBank under accession numbers QEAN000000000, QEAM000000000, QEAP000000000, QEAQ000000000, QEAR000000000, and QEAO000000000. Phylogenomic analyses were carried out using 192 Profile Hidden Markov Models (HMM) built from phylogenetically informative markers. The analysis and determination of CAZymes encoding genes were performed using a similar methodology described in [52]. For the analysis of meiotic toolbox genes, the methodology described in [19] was followed. RepeatModeler [53] was run on the individual genome sequences and consensus repeat models were created with 90% similarity to allow comparison of repeat content for the genomes under investigation. Presence of RIP was investigated with Occultercut [54] and RIPcal [55] and by determining the presence of the RID gene which is required for RIP. Protein motifs were identified in the LEV6574 secretome using MEME [56], which were used to screen the entire protein set of all chytrid isolates using MAST [57]. Analyses results of different tools were intersected and verified to assess

genomic features linked to the utilization energy sources, obligate biotrophy, cell wall degradation, pathogenicity and sexuality of chytrid fungi.

Acknowledgements

The authors thank Elio Schijlen and Bas te Lintel Hekkert (Wageningen University & Research: WUR) for generating the MB42 sequence data, David Carter (London Regional Genomics Centre) for sequencing the LEV6574 cDNA; Linda Bakker (WUR) for adapting and running Blobology; Jules Beekwilder (WUR) for reviewing and commenting on KEGG pathways; Julie Chapados and Kasia Dadej (Agriculture and Agri-Food Canada) for technical assistance in the sequencing of LEV6574; Genome Quebec for performing PacBio sequencing of LEV6574; Richard G.F. Visser and J.H. Vossen (Wageningen University and Research) for their support and advice on the project. We thank Averis seeds BV, Böhm-Nordkartoffel Agrarproduktion GmbH & Co. OHG, Danespo, HZPC Holland BV, C. Meijer BV, SaKa Pflanzenzucht GmbH & Co. KG and Teagasc for their involvement in the public-private partnership that was formed to fund this study.

Funding

Genome sequencing of LEV6574 was funded by Agriculture and Agri-Food (project J-000985) and Canadian Safety and Security Program (project CSSP 30vv01). Genome sequencing of MB42 and the following functional analysis was funded by Topsector Tuinbouw & Uitgangsmaterialen grant 1406-056 entitled “An integrated genomics and effectomics impulse for potato wart resistance management and breeding”.

Author contributions

BV, CL, HN, SW, TL designed parts of the study. CL, DS, MG, PB collected, prepared and pathotyped (where relevant) biological material. HG created the de novo assembly of M42, and HN created the assembly of LEV6574. BV, SW performed the structural annotation of both *S. endobioticum* genomes, in close collaboration with BV, HN. HN structurally annotated the genomes of the culturable chytrid. Functional annotations were performed by SW, and BV, HN, SW performed the validations. HN performed the phylogenomic analysis and, with BV, SW, analyzed the CAZymes. BV, SV performed the analyses for KEGG pathways, GO-terms, repeats and transposable elements. BV performed RIP and meiotic toolbox analyses. DJ, BV

Chapter 2

performed the candidate effector and motif analyses. CL, PB, TL acquired funding for the study. TL supervised BV and SW, and with CL discussed results and proposed solutions. BV, HN, SW, TL wrote the draft manuscript and all authors read, commented and approved the final manuscript.

Author affiliations

BV, HG, MG, PB, SW, TL: Wageningen University and Research, Droevendaalsesteeg 1, 6708 PB, Wageningen, the Netherlands; BV: Dutch National Plant Protection Organization, Geertjesweg 15, 6706 EA, Wageningen, the Netherlands; HN, AL: Agriculture and Agri-Food Canada, 960 Carling Avenue, Ottawa, Canada; DS: Canadian Food Inspection Agency, 93 Mount Edward Road, Charlottetown, Canada; DJ: Université de Moncton, 18 avenue Antonine-Maillet, Moncton, Canada.

References

1. Heckman, D.S., *et al.*, Molecular Evidence for the Early Colonization of Land by Fungi and Plants. *Science*, 2001. 293(5532): p. 1129.
2. Gleason, F.H., *et al.*, The ecology of chytrids in aquatic ecosystems: roles in food web dynamics. *Fungal Biology Reviews*, 2008. 22(1): p. 17-25.
3. Freeman, K.R., *et al.*, Evidence that chytrids dominate fungal communities in high-elevation soils. *Proceedings of the National Academy of Sciences*, 2009. 106(43): p. 18315.
4. Karling, J.S., *Synchytrium*. 1 ed. 1964: Academic Press, New York.
5. Longcore, J.E., D.R. Simmons, and P.M. Letcher, *Synchytrium microbalum sp. nov.* is a saprobic species in a lineage of parasites. *Fungal Biology*, 2016. 120(9): p. 1156-1164.
6. Obidiegwu, J.E., K. Flath, and C. Gebhardt, Managing potato wart: a review of present research status and future perspective. *Theor Appl Genet*, 2014. 127(4): p. 763-80.
7. Curtis, K.M., IX.— The life-history and cytology of *Synchytrium endobioticum* (schilb.), perc., the cause of wart disease in potato. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character*, 1921. 210(372-381): p. 409.
8. Hampson, M.C., History, biology and control of potato wart disease in Canada. *Canadian Journal of Plant Pathology*, 1993. 15(4): p. 223-244.
9. Laidlaw, W.M.R., A method for the detection of the resting sporangia of potato wart disease (*Synchytrium endobioticum*) in the soil of old outbreak sites. *Potato Research*, 1985. 28(2): p. 223-232.
10. Smith, I.M., *et al.*, Quarantine Pests for Europe - Data Sheets on quarantine pests for the European Union and for the European and Mediterranean Plant Protection Organization. 2 ed. 1997: CAB International, Wallingford, UK.
11. Simao, F.A., *et al.*, BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 2015. 31(19): p. 3210-3212.

12. Spatafora, J.W., *et al.*, A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia*, 2016. 108(5): p. 1028-1046.
13. Doweld, A.B. *Synchytriales* Doweld, ord.nov. IF550404. 2014 [cited 2018 23 November]; Available from: <http://www.indexfungorum.org/Publications/Index%20Fungorum%20no.92.pdf>.
14. Selker, E.U., *et al.*, Rearrangement of duplicated DNA in specialized cells of *Neurospora*. *Cell*, 1987. 51(5): p. 741-52.
15. Watters, M.K., *et al.*, Action of repeat-induced point mutation on both strands of a duplex and on tandem duplications of various sizes in *Neurospora*. *Genetics*, 1999. 153(2): p. 705-14.
16. Cambareri, E.B., M.J. Singer, and E.U. Selker, Recurrence of repeat-induced point mutation (RIP) in *Neurospora crassa*. *Genetics*, 1991. 127(4): p. 699-710.
17. Freitag, M., *et al.*, A cytosine methyltransferase homologue is essential for repeat-induced point mutation in *Neurospora crassa*. *Proc Natl Acad Sci U S A*, 2002. 99(13): p. 8802-7.
18. Gladyshev, E., Repeat-Induced Point Mutation (RIP) and Other Genome Defense Mechanisms in Fungi. *Microbiology spectrum*, 2017. 5(4): p. 10.1128/microbiolspec.FUNK-0042-2017.
19. Halary, S., *et al.*, Conserved meiotic machinery in *Glomus* spp., a putatively ancient asexual fungal lineage. *Genome Biol Evol*, 2011. 3: p. 950-8.
20. Selin, C., *et al.*, Elucidating the Role of Effectors in Plant-Fungal Interactions: Progress and Challenges. *Frontiers in Microbiology*, 2016. 7(600).
21. Smith, M.L., J.N. Bruhn, and J.B. Anderson, The fungus *Armillaria bulbosa* is among the largest and oldest living organisms. *Nature*, 1992. 356: p. 428.
22. Hyle, J.W., R.J. Shaw, and D. Reines, Functional distinctions between IMP dehydrogenase genes in providing mycophenolate resistance and guanine prototrophy to yeast. *J Biol Chem*, 2003. 278(31): p. 28470-8.
23. Wheeler, P.R., Biosynthesis and Scavenging of Purines by Pathogenic Mycobacteria Including *Mycobacterium leprae*. *Microbiology*, 1987. 133(11): p. 2999-3011.
24. Kubicek, C.P., T.L. Starr, and N.L. Glass, Plant Cell Wall-Degrading Enzymes and Their Secretion in Plant-Pathogenic Fungi. *Annual Review of Phytopathology*, 2014. 52(1): p. 427-451.
25. Sánchez-Vallet, A., *et al.*, The Genome Biology of Effector Gene Evolution in Filamentous Plant Pathogens. *Annual Review of Phytopathology*, 2016.
26. Dutheil, J.Y., *et al.*, A Tale of Genome Compartmentalization: The Evolution of Virulence Clusters in Smut Fungi. *Genome Biology and Evolution*, 2016. 8(3): p. 681-704.
27. Kinsey, J.A., *et al.*, The *Neurospora* transposon Tad is sensitive to repeat-induced point mutation (RIP). *Genetics*, 1994. 138(3): p. 657-64.
28. Fulci, V. and G. Macino, Quelling: post-transcriptional gene silencing guided by small RNAs in *Neurospora crassa*. *Curr Opin Microbiol*, 2007. 10(2): p. 199-203.
29. Sparrow, F.K., *Aquatic Phycomycetes*. 2nd ed. 1960, Ann Arbor: The university of Michigan Press.
30. Hampson, M.C., Sequence of events in the germination of the resting spore of *Synchytrium endobioticum*, European pathotype 2, the causal agent of potato wart disease. *Canadian Journal of Botany*, 1986. 64(9): p. 2144-2150.
31. Köhler, E., Zur Biologie und Cytologie von *Synchytrium endobioticum* (Schilb.) Perc. . *Phytopathol. Z.* , 1932. 4: p. 43-55.
32. Köhler, E., Zur Kenntnis der Sexualität bei *Synchytrium*. *Ber. deut. Botan. Ges.* , 1956. 69: p. 121-127.

Chapter 2

33. Lange, L. and L.W. Olson, Development of the resting sporangia of *Synchytrium endobioticum*, the causal agent of potato wart disease. *Protoplasma*, 1981. 106(1): p. 83-95.
34. Lange, L. and L.W. Olson, Germination and parasitism of the resting sporangia of *Synchytrium endobioticum*. *Protoplasma*, 1981. 106(1): p. 69-82.
35. Arkhipova, I. and M. Meselson, Transposable elements in sexual and ancient asexual taxa. *Proceedings of the National Academy of Sciences*, 2000. 97(26): p. 14473-14477.
36. Goodwin, S.B., *et al.*, Finished Genome of the Fungal Wheat Pathogen *Mycosphaerella graminicola* Reveals Dispensome Structure, Chromosome Plasticity, and Stealth Pathogenesis. *PLOS Genetics*, 2011. 7(6): p. e1002070.
37. Martin, F., *et al.*, The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature*, 2008. 452: p. 88.
38. Choi, H.W. and D.F. Klessig, DAMPs, MAMPs, and NAMPs in plant innate immunity. *BMC Plant Biology*, 2016. 16: p. 232.
39. de Jonge, R., *et al.*, Conserved fungal LysM effector *Ecp6* prevents chitin-triggered immunity in plants. *Science*, 2010. 329(5994): p. 953-5.
40. Sánchez-Vallet, A., J.R. Mesters, and B.P.H.J. Thomma, The battle for chitin recognition in plant-microbe interactions. *FEMS Microbiology Reviews*, 2015. 39(2): p. 171-183.
41. Boutemy, L.S., *et al.*, Structures of Phytophthora RXLR Effector Proteins: a conserved but adaptable fold underpins functional diversity. *The Journal of Biological Chemistry*, 2011. 286(41): p. 35834-35842.
42. Dong, S., S. Raffaele, and S. Kamoun, The two-speed genomes of filamentous pathogens: waltz with plants. *Current Opinion in Genetics & Development*, 2015. 35: p. 57-65.
43. Schmidt, S.M., *et al.*, MITEs in the promoters of effector genes allow prediction of novel virulence genes in *Fusarium oxysporum*. *BMC Genomics*, 2013. 14(1): p. 119.
44. Suzuki, M., O. Danilchanka, and J.J. Mekalanos, *Vibrio cholerae* T3SS effector *VopE* modulates mitochondrial dynamics and innate immune signaling by targeting Miro GTPases. *Cell host & microbe*, 2014. 16(5): p. 581-591.
45. Niu, H., *et al.*, *Anaplasma phagocytophilum* Ats-1 Is Imported into Host Cell Mitochondria and Interferes with Apoptosis Induction. *PLoS Pathogens*, 2010. 6(2): p. e1000774.
46. Dolezal, P., *et al.*, *Legionella pneumophila* Secretes a Mitochondrial Carrier Protein during Infection. *PLOS Pathogens*, 2012. 8(1): p. e1002459.
47. West, A.P., G.S. Shadel, and S. Ghosh, Mitochondria in innate immune responses. *Nat Rev Immunol*, 2011. 11(6): p. 389-402.
48. Vargas, W.A., *et al.*, A Fungal Effector With Host Nuclear Localization and DNA-Binding Properties Is Required for Maize Anthracnose Development. *Molecular Plant-Microbe Interactions*, 2015. 29(2): p. 83-95.
49. Canonne, J. and S. Rivas, Bacterial effectors target the plant cell nucleus to subvert host transcription. *Plant signaling & behavior*, 2012. 7(2): p. 217-221.
50. Sinclair, S.H., K.E. Rennoll-Bankert, and J.S. Dumler, Effector bottleneck: microbial reprogramming of parasitized host cell transcription by epigenetic remodeling of chromatin structure. *Frontiers in genetics*, 2014. 5: p. 274-274.
51. Jones, P., *et al.*, InterProScan 5: genome-scale protein function classification. *Bioinformatics*, 2014. 30(9): p. 1236-40.

52. Zerillo, M.M., *et al.*, Carbohydrate-Active Enzymes in Pythium and Their Role in Plant Cell Wall and Storage Polysaccharide Degradation. PLOS ONE, 2013. 8(9): p. e72572.
53. Smit, A.F.A. and R. Hubley. RepeatModeler Open-1.0. 2008 [cited 2018 21 September]; Available from: <http://www.repeatmasker.org>.
54. Testa, A.C., R.P. Oliver, and J.K. Hane, OcculterCut: A Comprehensive Survey of AT-Rich Regions in Fungal Genomes. Genome Biol Evol, 2016. 8(6): p. 2044-64.
55. Hane, J.K. and R.P. Oliver, RIPCAL: a tool for alignment-based analysis of repeat-induced point mutations in fungal genomic sequences. BMC Bioinformatics, 2008. 9(1): p. 478.
56. Bailey, T.L. and C. Elkan, Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol, 1994. 2: p. 28-36.
57. Bailey, T.L. and M. Gribskov, Combining evidence using p-values: application to sequence homology searches. Bioinformatics, 1998. 14(1): p. 48-54.

Supplementary material

Supplementary materials and methods

Supplementary figures

- Fig. S1 Normalization of average read coverage for ZOO selection
- Fig. S2 ZOO scores and selection of *S. endobioticum* scaffolds
- Fig. S3 Effect of the ZOO selection illustrated with Blobology
- Fig. S4 Mummerplot showing synteny between *S. endobioticum* pathotype 1(D1) isolate MB42 and pathotype 6(O1) isolate LEV6574
- Fig. S5 ASTRAL greedy consensus cladogram based on analyses of individual bootstrap trees for each of 192 conserved orthologous proteins
- Fig. S6 Cellular component GO-terms associated to flagella and movement, and their occurrence in chytrid species and higher fungi (Ascomycota and Basidiomycota) analyzed visualized with CytoScape
- Fig. S7 Predicted level 2 biological processes GO-term functions of genes shared by all chytrid species, genes present in all chytrids except in *S. endobioticum*, and genes unique to *S. endobioticum*
- Fig. S8 Example of false negatives in KEGG pathway predictions, illustrated by KEGG reference pathway 20: citrate cycle (TCA)
- Fig. S9 Detail KEGG reference pathway 20: citrate cycle (TCA) for six Ascomycota and Basidiomycota species with known KEGG pathways, focusing on aconitate hydratase (EC:4.2.1.3)
- Fig. S10. Cluster analysis based on assigned CAZyme modules
- Fig. S11 Repeat family content in chytrid species analyzed
- Fig. S12 GC-content and their frequencies in eleven chytrid genomes determined with Occultercut

Chapter 2

- Fig. S13 RIPcal analysis of “repeat family 36” containing (near) complete gene models of *S. endobioticum* isolates LEV6574 and MB42
- Fig. S14 RIPcal analysis of “repeat family 356” containing (near) complete gene models of *S. endobioticum* isolates LEV6574 and MB42
- Fig. S15 Detail of motif VI as defined by Gladyshev *et al.* (2017) in chytrid proteins with predicted C-5 cytosine methyltransferase activity
- Fig. S16 GO-terms associated with eleven meiosis specific proteins and their presence in the proteomes of chytrid species analyzed visualized with CytoScape
- Fig. S17 Annotation of the RAYH-motif on several LEV6574 and MB42 protein sequences
- Fig. S18 Maximum Likelihood cladogram of 193 *S. endobioticum* proteins carrying the RAYH-motif
- Fig. S19 Normalized average read coverage for candidate effectors and chytrid core SCOs
- Fig. S20 Comparison of gene and protein statistics for candidate effectors and chytrid core SCOs
- Fig. S21 MAFFT nucleotide alignment of the 223 candidate effector genes from *S. endobioticum* isolates MB42 and LEV6574 containing the RAYH-motif

Selected supplementary tables

- Table S1 Information on fungal isolates, NextGen sequence data, and genome sequences used in this study
- Table S2 Genome and annotation statistics of Chytridiomycota isolates included in this study
- Table S3 Benchmarking Universal Single Copy Orthologs (BUSCO) statistics
- Table S4 Genomes and download source used for phylogenomic analysis
- Table S7 Validation of KEGG pathway analysis using six Ascomycota and Basidiomycota organisms represented in the KEGG database

- Table S8 Protein sequences assigned to enzymatic steps in KEGG pathways for the eleven Chytridiomycota species analyzed
- Table S11 Presence of core meiotic genes as defined by Halary *et al.* (2011) in Chytridiomycota isolates analyzed
- Table S12 Repeat content in the Chytridiomycota species analyzed
- Table S13 RID protein sequences used for the detection of RID homologs in Chytridiomycota genomes

Several supplementary tables and supplementary files 3 and 4 are not included in this thesis on account of their size. These files are available online after publication of this chapter.

Chapter 2

Supplementary Materials and Methods

Fungal material and Next Generation Sequencing

Potato wart material containing *S. endobioticum* pathotype 6(O1) isolate LEV6574 (SeLEV6574) was collected on potato variety Russet Burbank in a field in St. Eleanors, Prince Edward Island, Canada in 2012. Potato wart material containing *S. endobioticum* pathotype 1(D1) isolate MB42 (SeMB42) was field collected in Langenboom, the Netherlands in 2002. This isolate is used as reference in the Dutch potato wart resistance testing program. Sampling, DNA extraction and Next Generation Sequencing of these isolates using one or more of three sequencing platforms, i.e. HiSeq 2500 (Illumina), MiSeq (Illumina), PacBio RSII (Pacific Biosciences), was previously described in [1]. In addition, four Roche 454 datasets from total genomic DNA of SeMB42 derived from purified resting spores as described previously [2], were produced. Total RNA was extracted from about 80 mg of purified LEV6574 spores using the Nucleospin RNA Plant kit (Machery-Nagel, Düren, Germany) following the manufacturer's instructions. Chilled spores were lysed in buffer RA1 (Machery-Nagel) containing 1 % beta-mercaptoethanol (Sigma-Aldrich) and 1 g of pre-chilled 2.0 mm zirconia beads (Biospec) for 1 min at room temperature. RNA was recovered from the lysate following the kit instructions and quantified using the Qubit RNA quantification system (ThermoFisher) followed by cDNA was preparation from 20 ng total RNA extracted from LEV6574 spores using the SMARTer PCR cDNA Synthesis Kit following the manufacturer's instructions (Clontech, Mountain View CA). The reverse transcription reaction was incubated for 110 min, and the cDNA was amplified for 20 cycles. The cDNA was sent to the London Regional Genomics Centre at the Western University (London, ON), where a Nextera XT (Illumina, San Diego CA) library was prepared and sequenced on a MiSeq (Illumina). For MB42, total RNA was extracted from about 1.5 gram of purified resting spores using RNeasy Plant mini kit (Qiagen) following manufacturer's instruction. Chilled spores were lysed in buffer RLC containing beta-mercaptoethanol and grinded droplet by droplet under liquid Nitrogen in a mortar using a pestle. Culturing, DNA extraction and Next Generation Sequencing of four culturable chytrid species, i.e. *Chytridium confervae* (CBS 675.73), *Powellomyces hirtus* (CBS 809.83), *Spizellomyces palustris* (= *Phlyctochytrium palustre*) (CBS 455.65) and *Synchytrium microbalum* (JEL517), was described in [1]. Sequencing information on materials used in this study are provided in table S1.

Genome assembly

For SeLEV6574, both MiSeq and PacBio CCS and CLR reads were quality trimmed with Trimmomatic v0.35 [3]. Bbmap v35 was used to merge overlaps in MiSeq reads. These resulting single end MiSeq reads were mapped to the potato (*Solanum tuberosum*) genome [4] with bbmap v35. Unmapped reads were collected and considered not belonging to the host and potentially belonging to the pathogen, *S. endobioticum*. Using the unmapped MiSeq reads combined with CCS and CLR PacBio reads, a hybrid genome assembly was performed with SPAdes v3.8.1 [5] (k=21, 33, 55, 77, 99, 127) with error correction and mismatch correction enabled. Using BLAST, the assembled contigs were compared to a collection of bacterial genomes and again to the potato genome. Contigs that were > 90% identical and possessed an e-value of < 1e-50 to bacteria or potato were reported. These matches were manually inspected. Contigs, where the match length to a bacteria and potato sequence divided by the total length of the contig were > 10%, were filtered out.

For SeMB42, host derived reads were identified by mapping against the potato genome assembly [4] with BWA v7.5 [6], and were removed from the datasets. A hybrid assembly using 454 and HiSeq data was performed with Celera v7.0. [7] following the authors' recommendations for hybrid assemblies with Illumina and 454 data. HiSeq reads were used for base correction, resolving mis-assemblies and gap-filling with three iterations of the Pilon v1.16 assembly improvement tool [8]. Further inspection of contaminant contigs and scaffolds for both isolates was performed using a comparative read-mapping approach referred to as "ZOO selection" (see below). Contigs smaller than 1kb were omitted from the assemblies.

Paired Illumina HiSeq data of chytrid species *C. confervae*, *P. hirtus*, *S. palustris*, and *S. microbalum* were used, after quality trimming (quality limit: 0.05, ambiguous limit: 2, min read length: 75), for *de novo* assembly with CLC genomics workbench v. 8.0.2 (word size: automatic, bubble size: automatic, scaffolding: on). Scaffolds were corrected using a read mapping approach (length fraction: 0.5, similarity fraction: 0.8) and consensus sequences were extracted for regions with $\geq 5x$ coverage.

ZOO selection

As *S. endobioticum* cannot be cultured, generating sequence data for this fungus free from its host and other contaminants is virtually impossible. ZOO selection, which is a comparative read mapping approach containing three separate elements, was performed

to identify *S. endobioticum* contigs and scaffolds. The following datasets were used in the procedure: A. ZOO_{mapping}: Fourteen HiSeq or MiSeq datasets generated for nine *S. endobioticum* isolates comprising six different pathotypes (i.e. 1(D1), 2(G1), 6(O1), 8(F1), 18(T1) and 38(Nevsehir)); B. ZOO_{RNAseq}: RNAseq data from *S. endobioticum* isolates MB42 and LEV6574; and C. ZOO_{heel-end}: HiSeq data from 24 potato heel end samples (healthy or infected with pathogens other than *S. endobioticum*) (table S1). For each reference isolate the procedure was as follows: HiSeq and MiSeq sequence data were stringently mapped to scaffolds and contigs in CLC genomics workbench (length fraction: 0.8, similarity fraction: 0.9) and average coverage scores per scaffold were determined. For the analysis of DNA and RNA derived *S. endobioticum* sequence data (ZOO_{mapping} and ZOO_{RNAseq}), normalisation of the data using median coverage values from regions with high correlation between the different isolates was performed per isolate (Fig. S1). As average read coverage of genomic scaffolds and contigs present in all *S. endobioticum* isolates are believed to have a uniform distribution, the median average coverage of the majority of scaffolds with a similar average coverage were used for normalization. For each scaffold-dataset combination, the absolute difference of the normalized coverage score to the median coverage value was determined. Per scaffold the sum of absolute differences for all datasets was calculated, and divided by the sum of normalized average read mapping coverage for all datasets. The resulting fractions were used to assign a ZOO_{mapping} and ZOO_{RNAseq} score per scaffold. For the heel-end datasets, average coverage was directly used to assign the ZOO_{heel-end} score. Total ZOO scores were calculated based on awards and penalties assigned to the individual ZOO scores: ZOO_{mapping} (<0.9: 10; ≥0.9 - 1: 8; ≥1: 0), ZOO_{RNAseq} (<1: 10; ≥1 - 10: 4; ≥10: 0), and ZOO_{heel-end} (0 - 0.1x: 0; ≥0.1 - 10x: -2; ≥10 - 100x: -5; ≥100 - 1,000x: -10; ≥1,000x: -20) (Fig. S2). Read mappings, blastn and blastx results were manually inspected for selected strong candidates, all potential candidates and selected weak candidates. The effect of the ZOO selection was visualized with Blobology [9] (Fig. S3).

Structural and Functional Genome Annotation

Respective *S. endobioticum* RNAseq reads were mapped to their respective genome assembly with tophat v2.1.1 [10]. The resulting BAM files were used as input for the BRAKER1 pipeline [11] for structural annotation of both *S. endobioticum* isolates. Genomic contigs of *C. confervae* CBS 675.73, *P. hirtus* CBS 809.83, *S. palustris* CBS 455.65, and *S. microbalum* JEL517, generated in this study, and the publically available *H. polyrhiza* genome were structurally annotated using the MAKER pipeline v2.31.8 [12]. Within the MAKER pipeline,

Chapter 2

the Genemark-ET v4.10 program [13, 14] was enabled and protein homology evidence from several closely related species, downloaded from JGI MycoCosm (*Batrachochytrium dendrobatidis* JAM81 v1.0, *Chytriomycetes* sp. MP 71 v1.0, *Gaertneriomyces semiglobifer* Barr 43 v1.0, *Globomyces pollinis-pini* Arg68 v1.0, *Gonapodya prolifera* [15], *Rhizoclostridium globosum* JEL800 Mondo [16], *Spizellomyces punctatus* DAOM BR117 v1.0 [17], *Anaeromyces robustus* S4 v1.0 [18], *Neocallimastix californiae* G1 v1.0 [18], *Piromyces finnis* v3.0 [18], *Piromyces* sp. E2 v1.0 [18], *Orpinomyces* sp C1A v1.0 [19]), as well as from *S. endobioticum* LEV6574 determined in this study, were provided to the pipeline for gene model prediction. Predicted gene models were validated using the NCBI tool tbl2asn and erroneous models were filtered out from the final annotations. Annotation completeness was assessed using version 2 of the fungal Benchmarking Universal Single Copy Orthologs (BUSCO) dataset [20]. Incomplete gene models and gene models with internal stop codons were removed from the dataset. For protein family classification, Gene Ontology (GO) analysis and the analysis of biochemical pathways, we used InterProScan v5.1.6 [21] to query multiple protein oriented databases such as Pfam v28.0, PANTHER v10.0, Phobius v1.01, SUPERFAMILY v1.75, PRINTS v42.0.

Genome statistics

Basic genome statistics were determined with QUAST v4.5 [22]. Obtained genome sizes, gene lengths and mean exons counts were compared to 172 fungal and Oomycete genome sizes published in Mohanta and Bae [23] with the addition of publically available Chytridiomycota genomes (*B. dendrobatidis* JAM81 and JEL423, *Gonapodya prolifera* JEL478, *Homolophlyctis polyrhiza* JEL142, and *Spizellomyces punctatus* DAOM BR117).

Orthologous genes and functional analysis

Orthologous genes shared between the two *S. endobioticum* isolates were identified using a Smith-Waterman alignment [24]. Inference of clusters of orthologous genes (COGs) was performed with OrthoFinder v1.1.4 [25] with default settings on protein datasets of both *S. endobioticum* and nine additional culturable chytrid species: *C. confervae* (CBS 675.73), *P. hirtus* (CBS 809.83), *S. palustris* (CBS 455.65), *S. microbalum* (JEL517), *B. dendrobatidis* JAM81 and JEL423, *G. prolifera* JEL478, *H. polyrhiza* JEL142, and *S. punctatus* DAOM BR117. Gene Ontologies (GO) and KEGG pathways for COGs specific to *S. endobioticum* and COGs absent in *S. endobioticum* but present in the culturable chytrid species were compared to COGs present in all isolates included (i.e. chytrid core COGs).

Phylogenetic Reconstruction

Phylogenomic analyses were carried out following similar methodology described in Spatafora *et al.* [26]. Briefly, each of the 192 Profile Hidden Markov Models (HMM), built from phylogenetically informative markers [27], was searched against the predicted proteome from 59 species (table S3) across kingdom Fungi with `hmmsearch` from the `hmmer3.1b` package [28]. For each marker, a cutoff of $1e-20$ was used to find the best scoring protein sequence in each species. Sequence alignment by profile HMM was carried out using `hmmalign`. These alignments were trimmed in two steps with `trimAl v1.4rev15` [29], first with options `-resoverlap 0.50 -seqoverlap 60`, and then with the `-automated1` option. The alignments were concatenated into a single super matrix alignment using the `catfasta2phym.pl` script (Johan A.A. Nylander - <https://github.com/nylander/catfasta2phym.pl>). A maximum likelihood phylogenetic analysis was performed using `RAXML v8.2.9` [30] with the fast bootstrap method (option `-f a`) and 100 bootstrap replicates. The `-m PROTCATAUTO` option was called to determine the best model of amino acid substitution. To assess potential conflict among markers, a phylogeny for each individual alignment was inferred with `RAXML` using the same parameters. The resulting trees were analyzed in `Astral v4.10.12` [31] to construct a greedy consensus tree of the 100 bootstrapped replicate trees.

KEGG pathway analysis

The KEGG pathway analysis is available from www.biorxiv.org/content/early/2018/07/16/369785 (last accessed: 13 February 2019). For the KEGG pathway comparison and InterProScan output validation, the results from the seventeen protein sets were processed using the `comparePathway.py` and `controlPathway.py` scripts (<https://github.com/swarris/sendo>). Both scripts use the KEGG API to associate EC to KO annotations and count the number of times an EC was predicted. For the comparison, included isolates were grouped in four categories: obligate biotrophic chytrids (OCh; i.e. *S. endobioticum*); culturable chytrids (CCh); culturable higher fungi (CHF); and obligate biotrophic higher fungi (OHF).

To benchmark the analysis, results from our pipeline were compared to the reference pathways in the KEGG database for the six fungal organisms with known pathways (CHF, OHF). The coloring in the resulting PDFs for the benchmark results were as follows: true positives: blue; false positives: red; false negative: light blue; and true negative: white.

Chapter 2

The counts for the reference pathways of the 6 control species were used to calculate the sensitivity, specificity and accuracy of the InterProScan predictions using the following formulas:

Sensitivity or true positive rate: $TPR=TP/(TP+FN)$

Specificity or true negative rate: $TNR=TN/(TN+FP)$

Accuracy: $ACC=(TP+TN)/(TP+TN+FP+FN)$

in which

True positive (TP): element in the reference pathway correctly predicted by InterProScan

True negative (TN): element in the not reference pathway and not predicted by InterProScan

False positive (FP): element in the not reference pathway but predicted by InterProScan. Note could also be a valid new prediction

False negative (FN): element in the reference pathway but not predicted by InterProScan.

Some elements in the pathways have no EC number record in the database. These cannot be mapped to the InterProScan output containing only EC numbers and are left out of the calculations

GO-term analysis

The GO-term analysis is available from www.biorxiv.org/content/early/2018/07/16/369785 (last accessed: 13 February 2019).

CAzymes

The analysis and determination of Carbohydrate-active enzymes (CAZymes)-encoding genes were performed using a similar methodology described in Zerillo *et al.* [32]. Briefly, the CAZymes for the 11 chytrids and control species were predicted using the online tool dbCAN [33], which searches for protein domains signatures search of CAZymes, via hidden Markov models constructed for each of the CAZy families, based on the CAZy (Carbohydrate-Active Enzyme) database classification [34], using default parameters. The CAZymes were categorized according to the type of reaction catalyzed (table S10): carbohydrate esterases (CE), glycoside hydrolases (GH), glycosyl transferases (GT), polysaccharide lyases (PL) and carbohydrate-binding modules (CBM), as described by [34]. Clustering based on the CAZyme results was done by counting the number of times a particular CAZyme was found in the

protein set. These counts were subsequently normalized by dividing the counts by the total number of proteins in the dataset. The resulting matrix was clustered using pvclust in R (<http://stat.sys.i.kyoto-u.ac.jp/prog/pvclust/>).

Meiotic toolbox genes

Protein sequences of each chytrid species were formatted into a BLAST database to identify genes involved in meiosis, following the methodology described by [35]. Meiotic genes in *Saccharomyces cerevisiae* and *Cryptococcus neoformans* were chosen as input queries (e-value < 1e-5) for the individual protein BLAST databases. When no blast hits were obtained with the *S. cerevisiae* and *C. neoformans* proteins, the analysis was repeated with *Allomyces macrogynus* orthologs. Absence of meiotic genes was verified with tblastn and manual curation (table S11).

Analysis of repeats and RIP activity

RepeatModeler v1.0.4 [36] was run on the individual genome sequences with default settings. Consensus repeat models were created with a 90% similarity for the combined chytrid species to allow comparison of repeat content shared between the species under investigation. Repeat sequences were masked with RepeatMasker v4.0.7 [37]. Repeats unique to *S. endobioticum* with repeat models of ≥ 1 kb were selected and aligned using MAFFT [38] incorporated in Geneious R10 [39]. Repeat sequences in the alignment with complete gene models, gene size ≥ 800 nt, and similarity $\geq 80\%$ were extracted from the alignment, realigned and tested for signatures of RIP activity using RIPcal [40] using the degenerate consensus sequence as reference. Putative RID proteins were identified using protein domains identified by InterProScan, and by blastp of known RID proteins of 18 fungal isolates (e-value < 1e-5) (table S13).

Analysis of candidate effectors

Secretomes were defined in a broad sense as proteins having a secretion signal peptide as determined by SignalP v4.1 [41], lacking transmembrane domains as determined by TMHMM2.0 [42]. From the secretome, motifs were identified with MEME [43] which used screen the entire protein set with MAST [44]. Subcellular location of chytrid proteins was determined using TargetP v1.1 [45]. Small secreted candidate effector proteins were identified using EffectorP [46], ApoplastP [47], and LOCALIZER [48].

Chapter 2

Candidate effector protein sequences were aligned per COG with MAFFT, and a consensus alignment was created for the complete protein set in Geneious. A Maximum Likelihood phylogeny was performed on the protein alignment in CLC genomics workbench using the WAG substitution model with 500 bootstrap replicates.

Identification of *S. endobioticum* sequences from a metagenome assembly using “ZOO selection”

ZOO selection is a comparative read mapping approach used to identify scaffolds and contigs for the species of interest from a metagenomic assembly. To allow comparison of the average read coverage per scaffold for the *S. endobioticum* isolates included in the analyses, normalization of the data was applied.

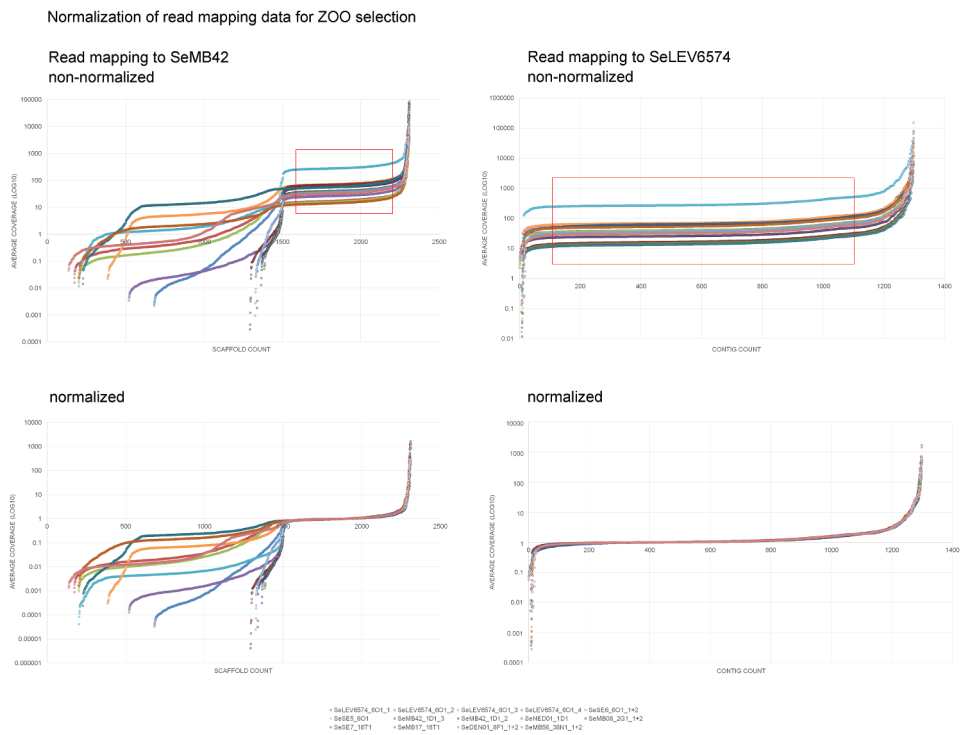


Figure S1. Selection of scaffolds and contigs for normalization (top) and the effect of normalization on distribution of average read coverage (bottom). Per isolate, average read mapping coverage values were sorted low to high for the 2310 SeMB42 pre-ZOO scaffolds and 1298 pre-ZOO SeLEV6574 contigs. Based on the similarities in average read coverage between isolates, areas boxed in red were believed to be true *S. endobioticum* sequences and were used for normalization.

For each of the three read mapping strategies in the ZOO selection (i.e. 1. ZOO_{mapping}: 14 *S. endobioticum* Illumina datasets mapped to metagenomic scaffolds; 2. ZOO_{RNAseq}: RNAseq data from *S. endobioticum* isolates MB42 and LEV6574 mapped to metagenomic scaffolds; and 3. ZOO_{heel-end}: HiSeq data from 24 potato heel end samples (healthy or infected with pathogens other than *S. endobioticum*) mapped to metagenomic scaffolds), scores and penalties were awarded resulting in a total ZOO score (Fig. S2). These scores were used to identify strong, potential and weak candidate *S. endobioticum* derived sequences. Potential and selected weak candidate scaffolds and contigs were manually checked to determine if they were *S. endobioticum* derived. The effect of the ZOO selection is visualized in Figure S3.

Distribution of total ZOO scores

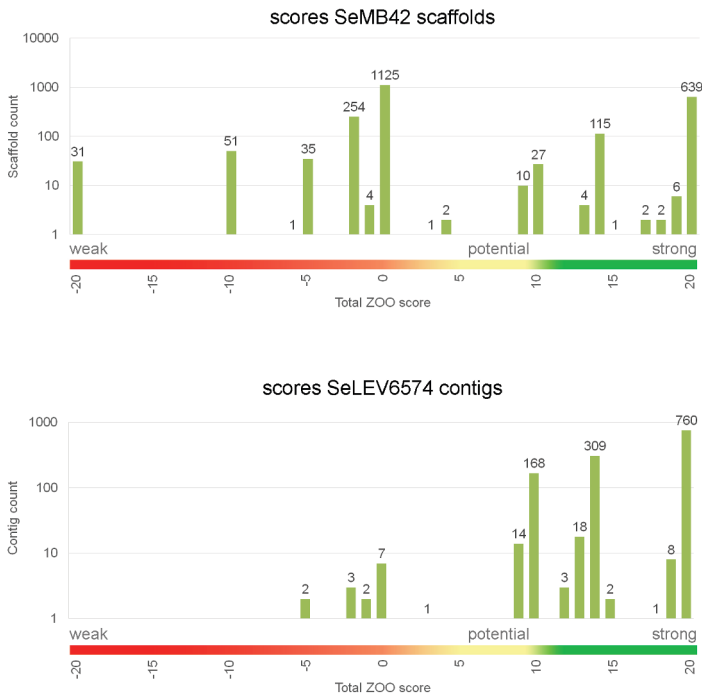


Figure S2. Total ZOO scores for the pre-ZOO selection sequences of SeMB42 and SeLEV6574. The number of scaffolds or contigs with a particular score is presented above the bars. The higher the total ZOO score, the more likely the scaffold truly belongs to *S. endobioticum*: total ZOO score of ≥ 12 were considered strong candidates, 1-11 were regarded as potential candidates, and scaffolds with total ZOO scores ≤ 0 were considered as weak candidates.

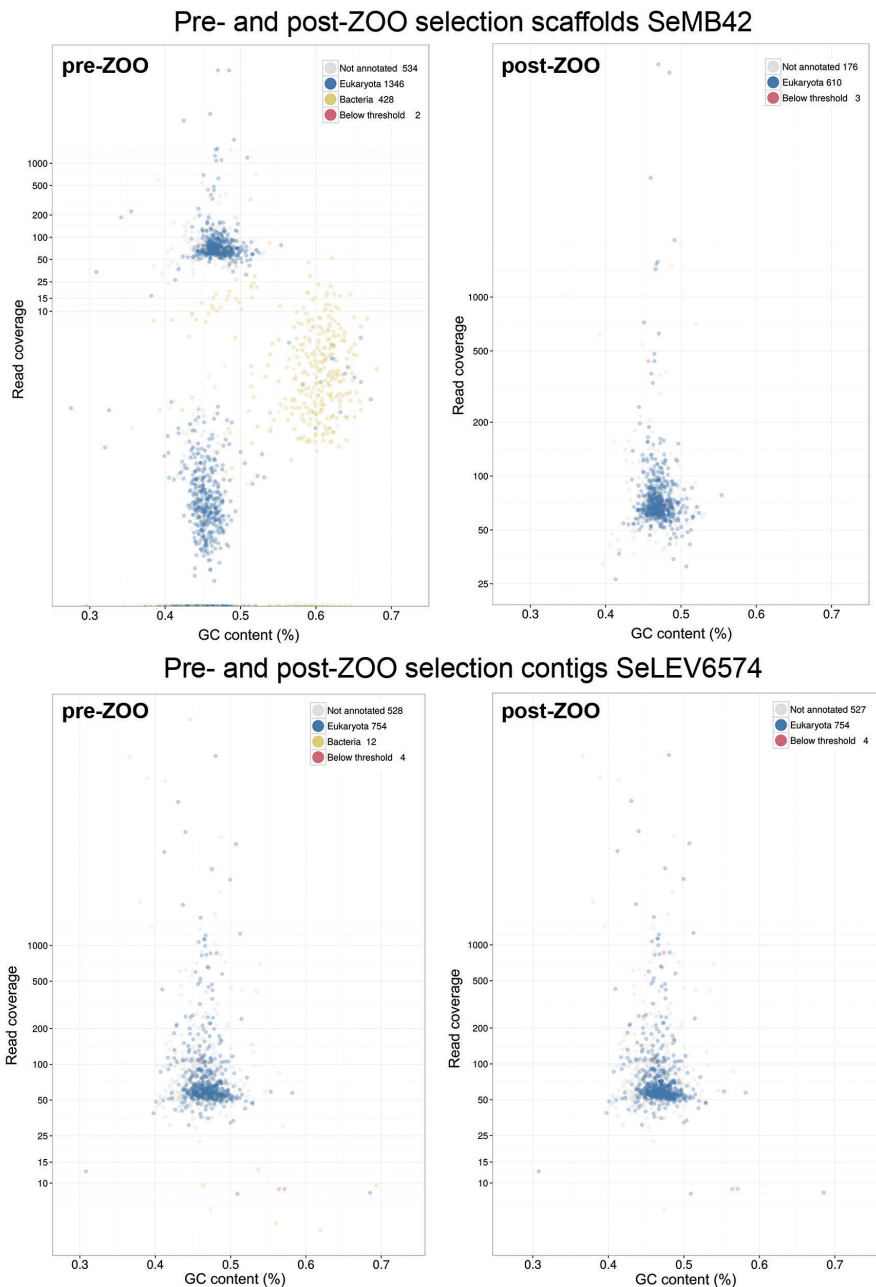


Figure S3. Effect of the ZOO selection illustrated with Blobology. Dots in the plots represent individual scaffolds or contigs which are plotted by read coverage (Y-axis) in function of GC content (X-axis). Kingdom level blastn based identification is used to color individual scaffolds. Scaffolds and contigs with the annotation “unassigned” did not produce a blastn hit, and sequences smaller than 200 were classified “below threshold”.

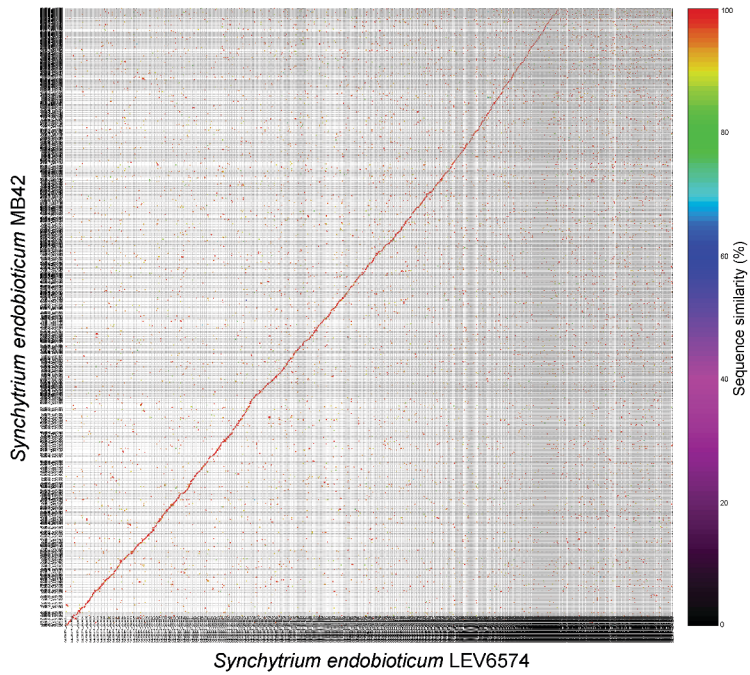


Figure S4. Mummerplot showing synteny between *S. endobioticum* pathotype 1(D1) isolate MB42 and pathotype 6(O1) isolate LEV6574. The 23.21 Mb genome of SeLEV6574 (x-axis, contigs sorted by sequence length in descending order) was used as reference for alignment of scaffolds from the 21.48 Mb SeMB42 genome (y-axis). The color scale indicates the sequence similarity between the two genomes, and the majority of sequences is (close to) 100% identical.

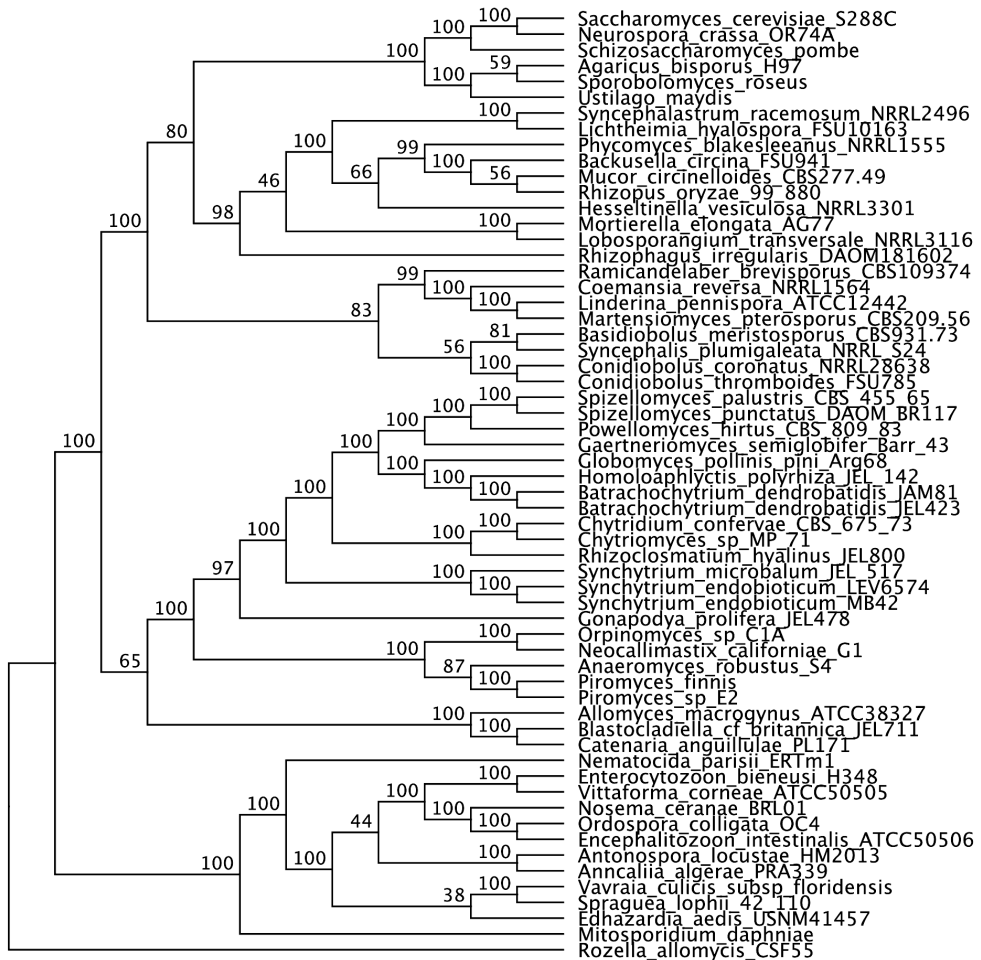


Figure S5. ASTRAL greedy consensus cladogram based on analyses of individual bootstrap trees for each of 192 conserved orthologous proteins. Bootstrap support values are shown on the individual branches.

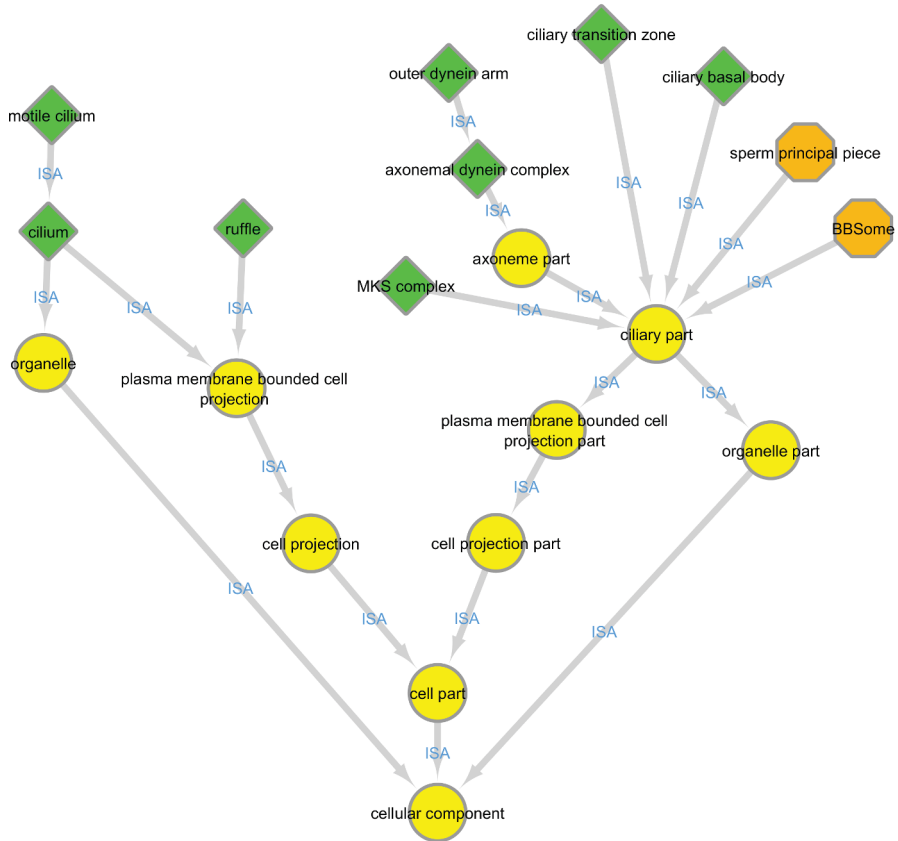
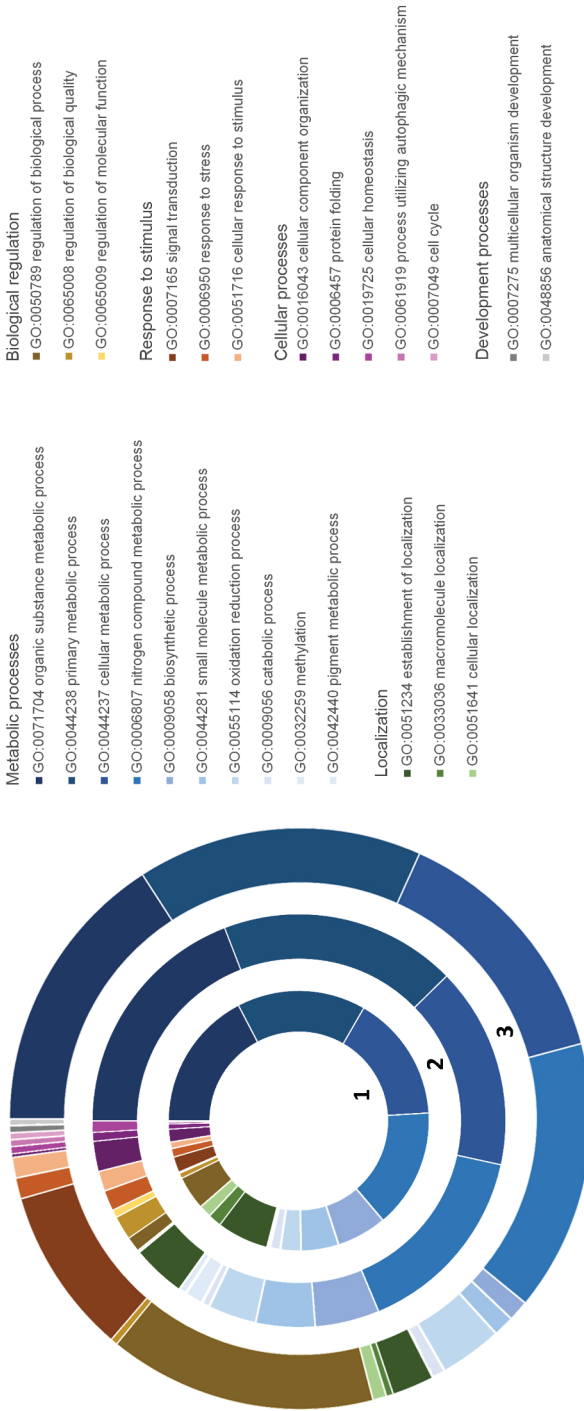


Figure S6. Cellular component GO-terms associated to flagella and movement, and their occurrence in chytrid species and higher fungi. Shapes in the network indicate the different GO-terms with “Cellular Compartment” being the highest level term. Grey arrows indicate an “is a” relation between the connected GO-terms. The different shapes indicate if a given GO-term is present in culturable and obligate biotrophic chytrid species but not in higher fungi (green diamonds), present in one or more culturable chytrid species (orange octagon), and unclassified (yellow circles).



1. chytrid core genes
2. genes absent in *S. endobioticum*
3. *S. endobioticum* specific genes

Figure S7. Predicted level 2 biological processes GO-term functions. With, **1** genes shared by all chytrid species (1848 COGs), **2** present in all chytrids except in *S. endobioticum* (76 COGs), and **3** unique to *S. endobioticum* (1413 COGs). Level 2 terms are grouped under respective level 1 biological processes terms (Metabolic processes: blue, Localization: green, Biological regulation: brown, Response to stimulus: red, cellular processes: pink, developmental processes: grey) and sorted on high to low incidence.

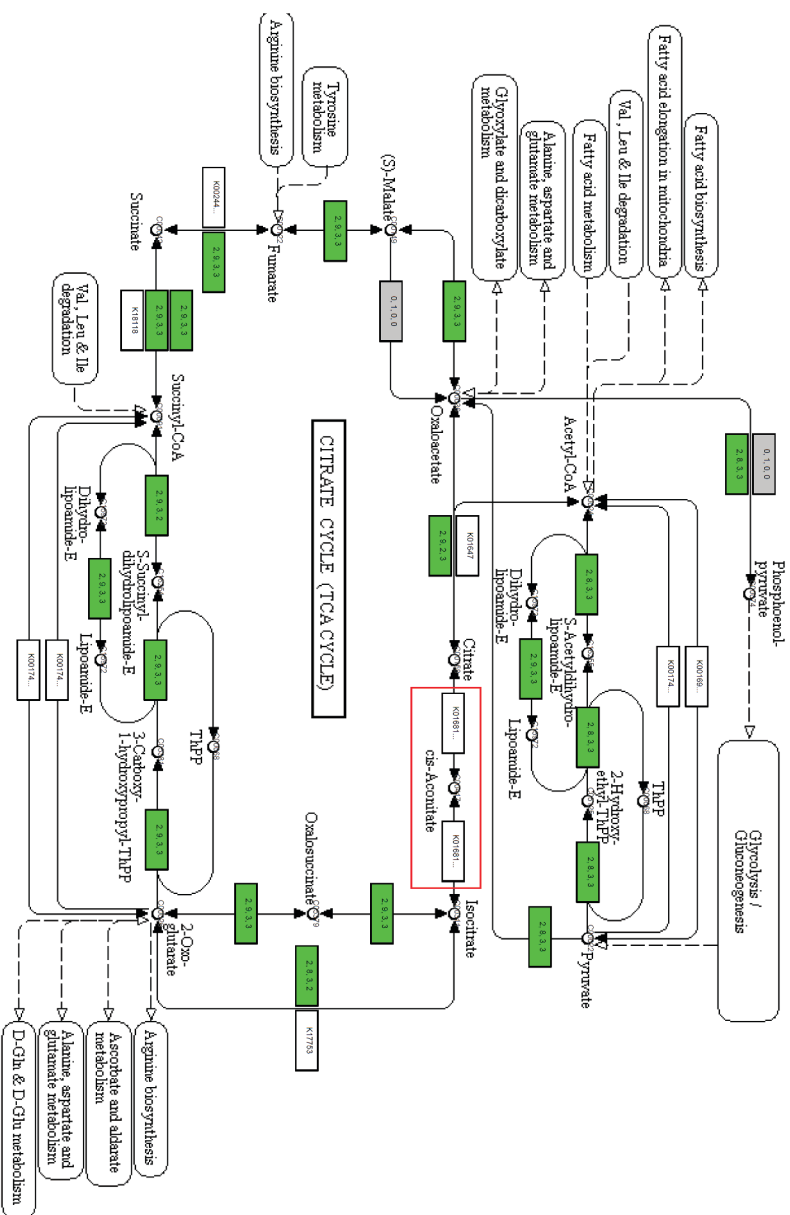
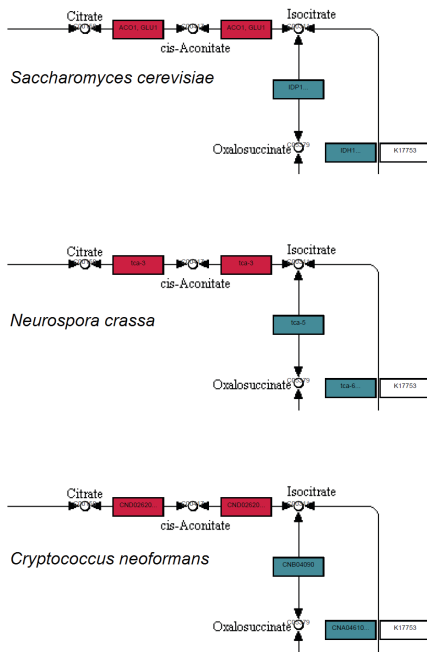


Figure S8. KEGG reference pathway 20: citrate cycle (TCA). Numbers (a, b, c, d) in the enzymatic steps indicate the number of isolates in a given group for which the corresponding gene was detected, in which a: obligate biotrophic chytrids (OCh, max = 2); b: culturable biotrophic chytrids (CCh, max = 9); c: culturable higher fungi (CHF, max = 3), and d: obligate biotrophic higher fungi (OHF, max = 3). The maximum score that can be obtained is "2,9,3,3". Colors indicate which groups (OCh, CCh, CHF, OHF) are represented for a given enzymatic step: present in ≥1 isolate in OCh, CCh, CHF and OHF (green); present in ≥1 isolate but not in a combination of interest linked to lifestyle or taxonomic placement (see supplementary file 3) (grey); unassigned (white). Boxed in red is the false-negative aconitate hydratase (EC:4.2.1.3).

Culturable controls



Obligate biotrophic controls

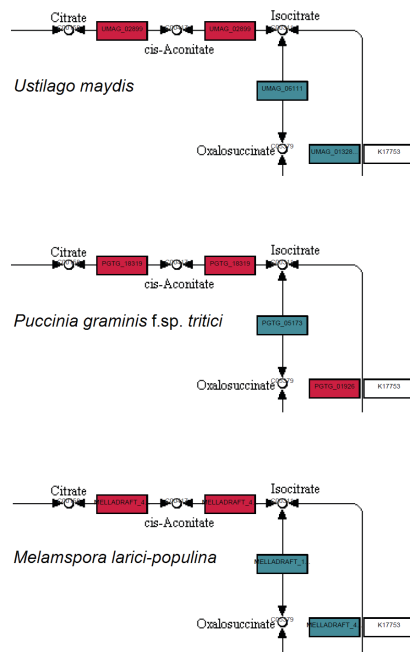


Figure S9. Detail KEGG reference pathway 20: citrate cycle (TCA) for six Ascomycota and Basidiomycota species with known KEGG pathways. Blue colors represent true positives, i.e. the element is described in the reference pathway and was correctly predicted in our analysis. Red colors represent false negatives (aconitate hydratase (EC:4.2.1.3) in this figure), i.e. the element is described in the reference pathway but was not predicted in our analysis. Blanc enzymatic steps are unassigned.

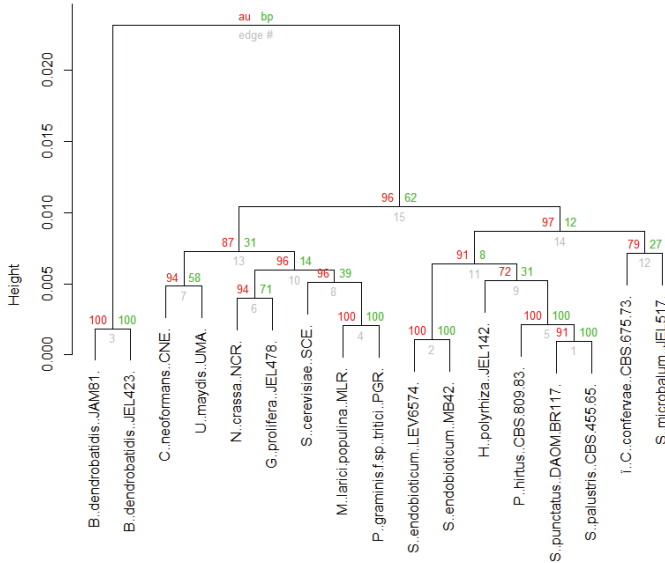


Figure S10. Cluster analysis based on all assigned CAZyme modules. Clustering was performed by pvclust (<http://stat.sys.i.kyoto-u.ac.jp/prog/pvclust/>) in R. AU = Approximately Unbiased p-value and BP = Bootstrap Probability value.

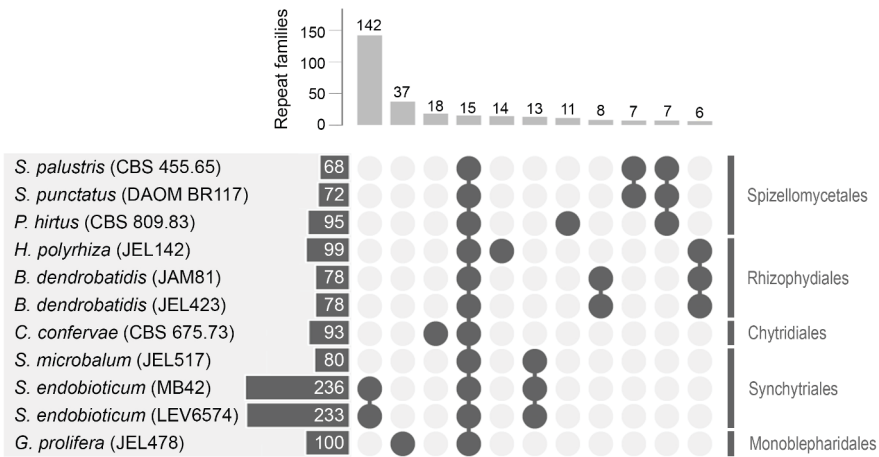


Figure S11. Repeat families identified in the chytrid species analyzed. Horizontal bars indicate the number of repeat families in the respective isolates. Vertical bars indicate the number of repeat families in intersecting selections. Intersections with more than five repeat families are shown sorted from the most abundant to the least abundant repeat families. With 236 and 233 different repeat families, *S. endobioticum* isolates MB42 and LEV6574 have the highest diversity of complex repeats on their respective genomes. In addition, the largest intersection of repeat families, with 142 different repeat families, is specific to *S. endobioticum*. This in contrast to 15 repeat families being shared by all chytrid genomes, and only eight being shared by both *B. dendrobatidis* genomes.

Chapter 2

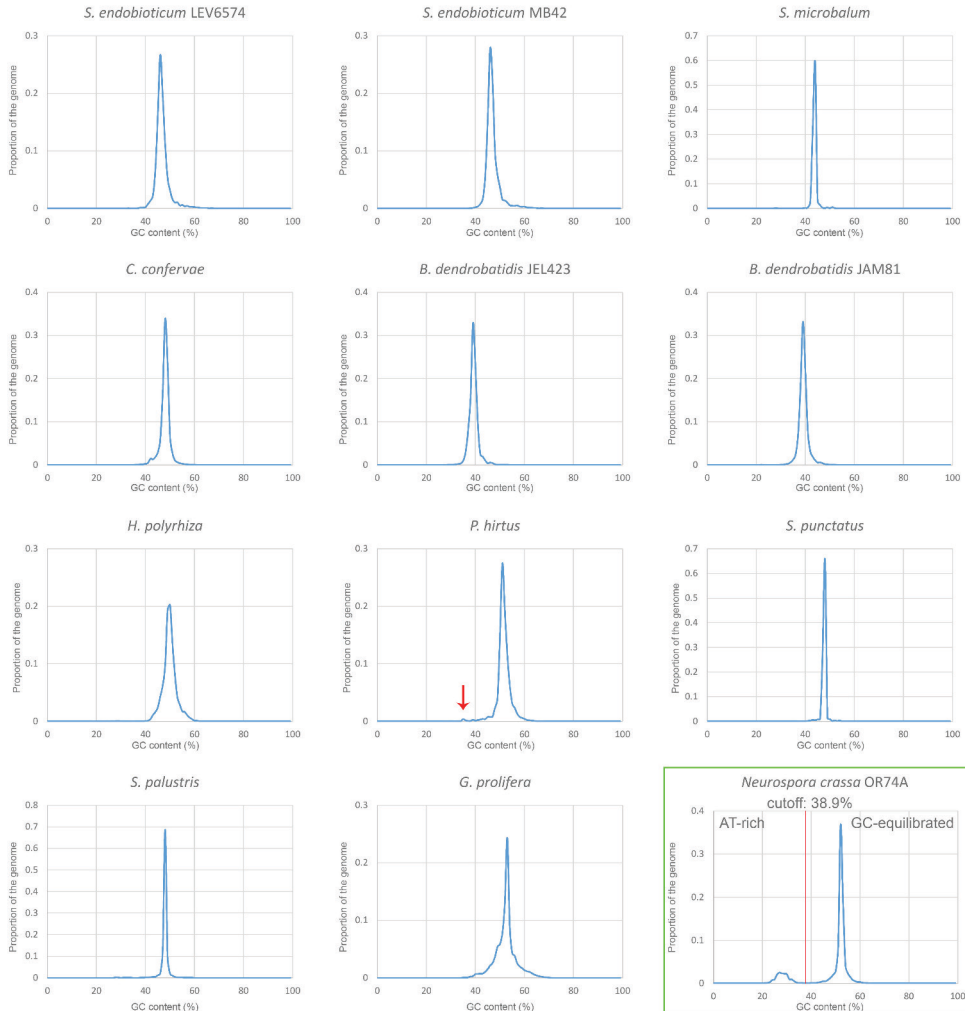


Figure S12. Detection of putative RIP-affected genomic regions. GC-content (x-axis) and their frequencies (y-axis) in eleven chytrid genomes determined with Occultercut. In the chytrid species analyzed, no AT-rich segments could be identified and GC-content plots show unimodal distribution. In the RIP proficient *N. crassa* genome (boxed in green), categorization of AT-rich segments and GC-equilibrated could be performed (cutoff: 38.9% GC-content). In *P. hirtus*, a second peak was identified (red arrow) but no categorization of AT-rich segments and GC-equilibrated could be made. This is likely the result of AT-rich segments overlapping too much with GC-equilibrated segments as described by [49]. The latter would suggest the AT-rich proportion of the *P. hirtus* genome is not the result of RIP activity.

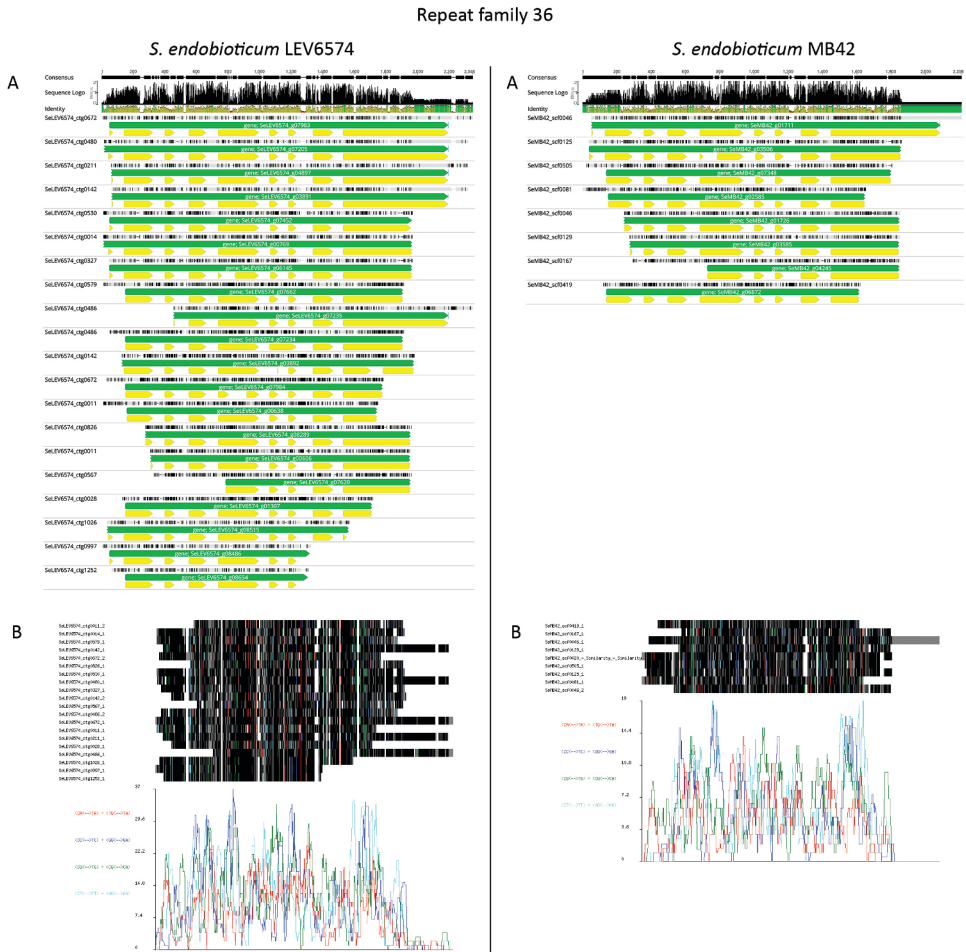


Figure S13. RIP analysis of *S. endobioticum* repeat family 36. **A.** MAFFT alignment of repeat sequences identified as “family 36” containing (near) complete gene models of *S. endobioticum* isolates LEV6574 (left) and MB42 (right), representing 20 and 8 sequences with alignment lengths of 2348 bp and 2208 bp respectively. Gene annotations are green whereas coding sequences are annotated in yellow. In the alignment, differences to the overall consensus sequence are highlighted in black. **B.** RIPcal analysis of the alignments presented under **A** using the degenerate consensus sequence as reference. Substitution frequencies between CpA - TpA + TpG - TpA (red), CpC - TpC + GpG - GpA (dark blue), CpG - TpG + CpG - CpA (green), and CpT - TpT + ApG - ApA (light blue) are similar and no dominance of CpA - TpA mutations (red) which are typical for RIP [40] was observed.

Repeat family 356

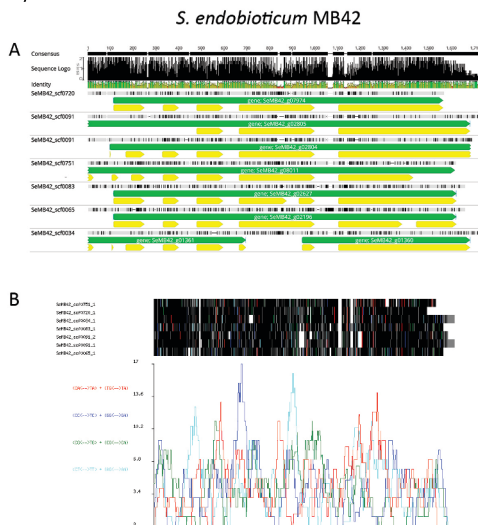
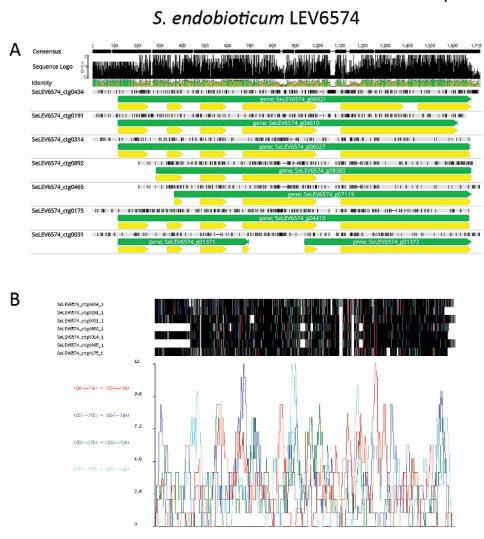


Figure S14. RIP analysis of *S. endobioticum* repeat family 356. **A.** Alignment of family 356 repeat sequences, representing 7 sequences with alignment lengths of 1719 bp for both *S. endobioticum* isolates. Gene annotations are green whereas coding sequences are annotated in yellow. In the alignment, differences to the overall consensus sequence are highlighted in black. **B.** RIPcal analysis of the alignments presented under **A** using the degenerate consensus sequence as reference.



Figure S15. Detail of motif VI as defined by [50] in chytrid proteins with predicted C-5 cytosine methyltransferase activity. Proteins with such activity were detected in all chytrid isolates analyzed. The conserved NV (asparagine-valine: boxed) is present in all chytrid proteins, except for two proteins from *S. punctatus* and *S. palustris*. In none of the chytrid proteins the consecutive QT or ET amino acid sequences specific to the RIP defective protein (RID) and methylation induced premeiotically protein (MIP) were observed. The QT di-amino acid sequence is present in the *N. crassa* RID protein (ESA43828.1).

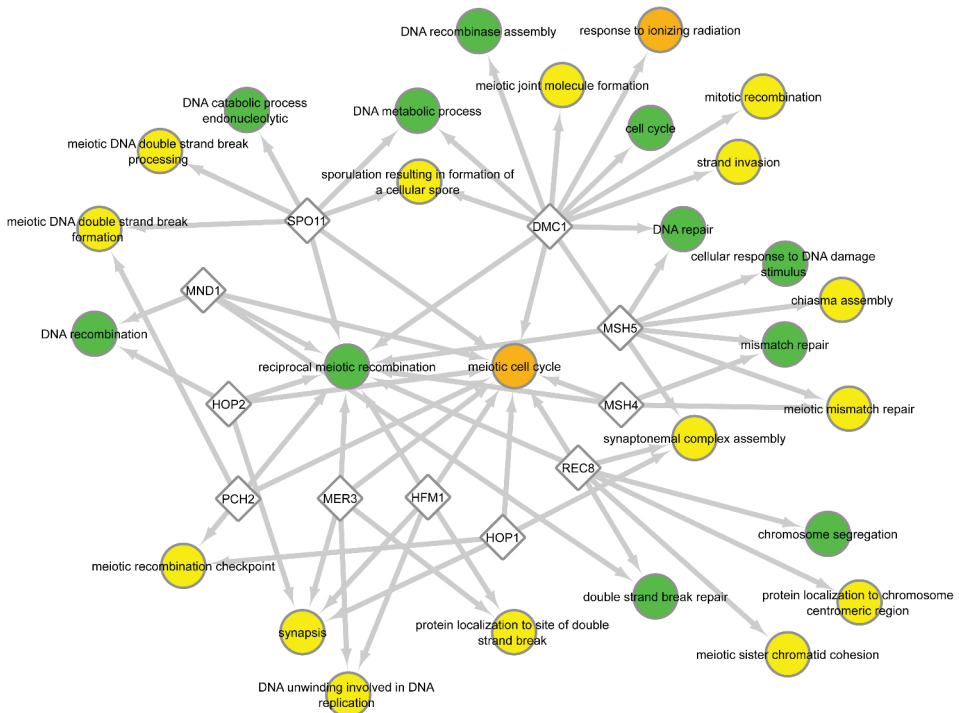


Figure S16. GO-terms associated with eleven meiosis specific proteins and their presence in the proteomes of chytrid species analyzed. Meiosis specific proteins are represented as white diamonds and associated GO-terms (circles) are indicated with grey links. GO-terms detected in proteomes of both obligate biotrophic chytrids and culturable chytrids are colored green, whereas GO-terms detected in only culturable chytrids are colored orange. Unassigned terms are colored yellow.

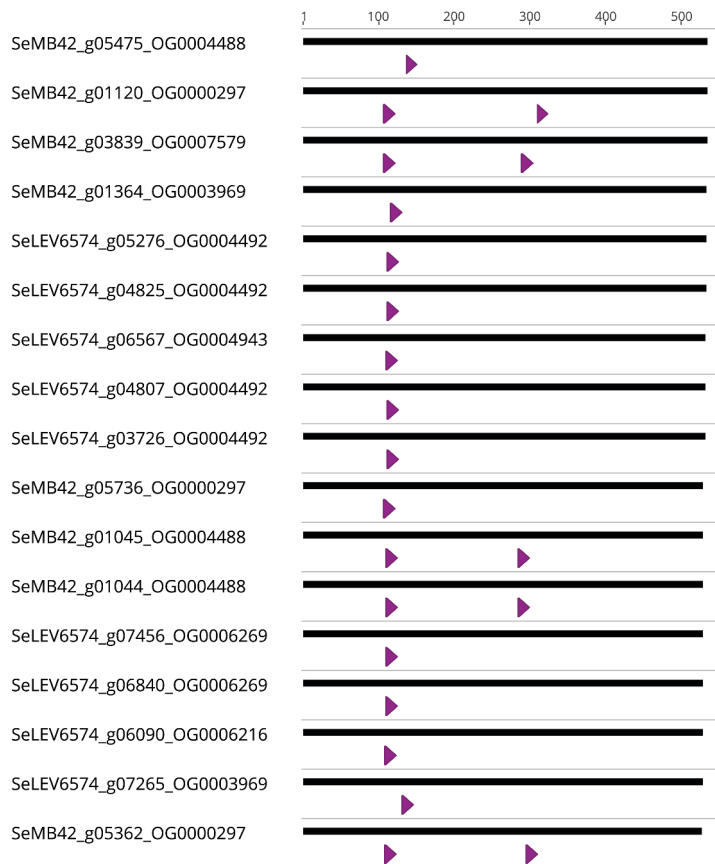


Figure S17. Annotation of the RAYH-motif (purple) on several LEV6574 and MB42 protein sequences (not aligned). Sequence names include COG membership. The scale bar indicates the protein length in number of amino acids. The first occurrence of the motif is typically found between amino acid positions 90 and 120, whereas the second motif (if present) is typically found between amino acid positions 280 and 310.

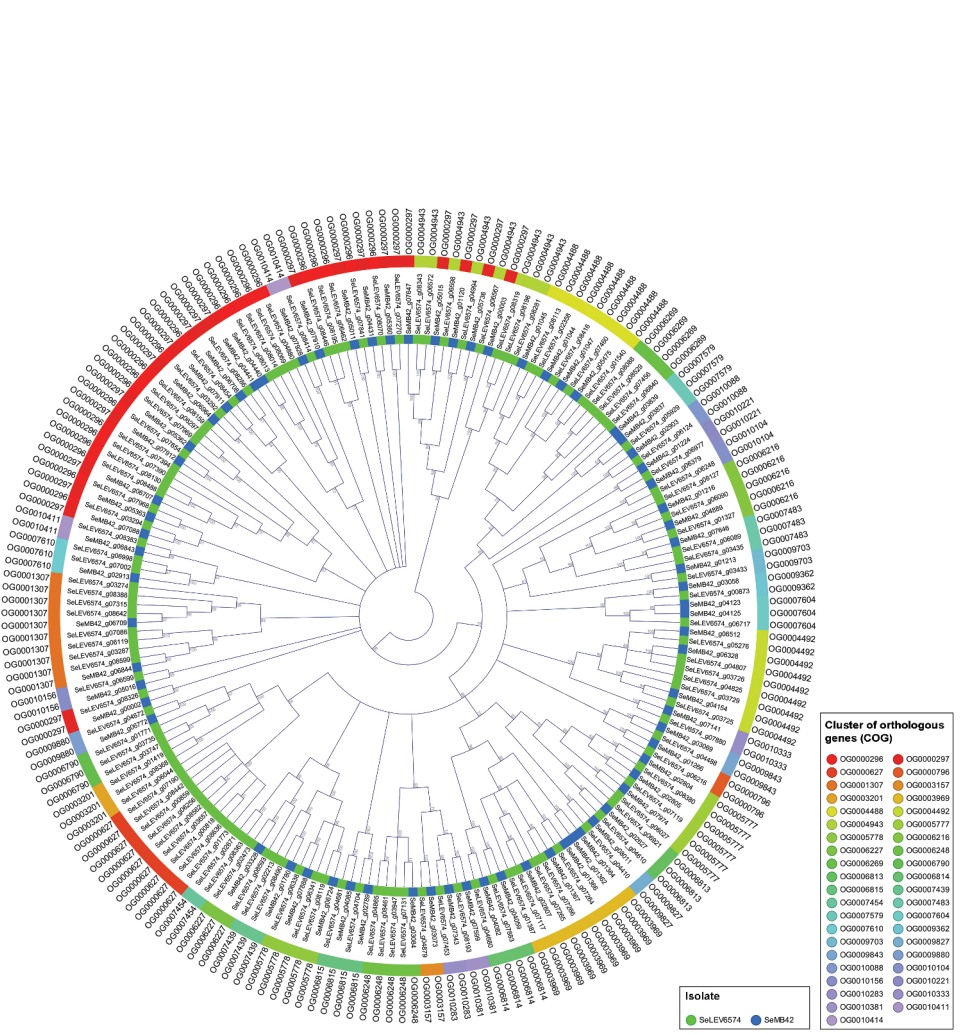


Figure S18. Maximum Likelihood cladogram of 193 *S. endobioticum* proteins carrying the RAYH-motif. The tree represents 122 LEV6574, and 71 MB42 sequences respectively. Bootstrap values $\geq 50\%$ are shown on tree nodes. The inner colored ring represents the isolate identity where LEV6574 is shown in green and MB42 is shown in blue. The outer colored ring represents the COG membership of the individual protein sequences. The gene names are displayed as labels in the inner node and COG identifiers are shown as labels on the outer node.

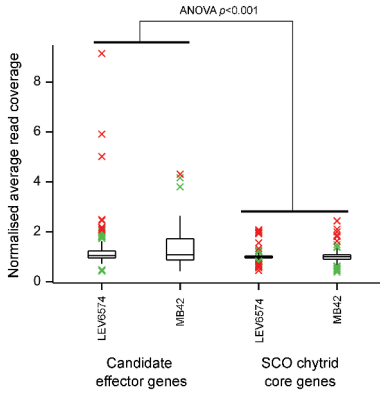


Figure S19. Normalized average read coverage for candidate effectors and chytrid core SCOs. Candidate effectors represent 148 and 75 proteins for LEV6574 and MB42 respectively, whereas the chytrid core SCOs constitute 694 proteins for both isolates. The MB42 candidate effector genes have a much wider range in the second and third quartiles compared to those of LEV6574.

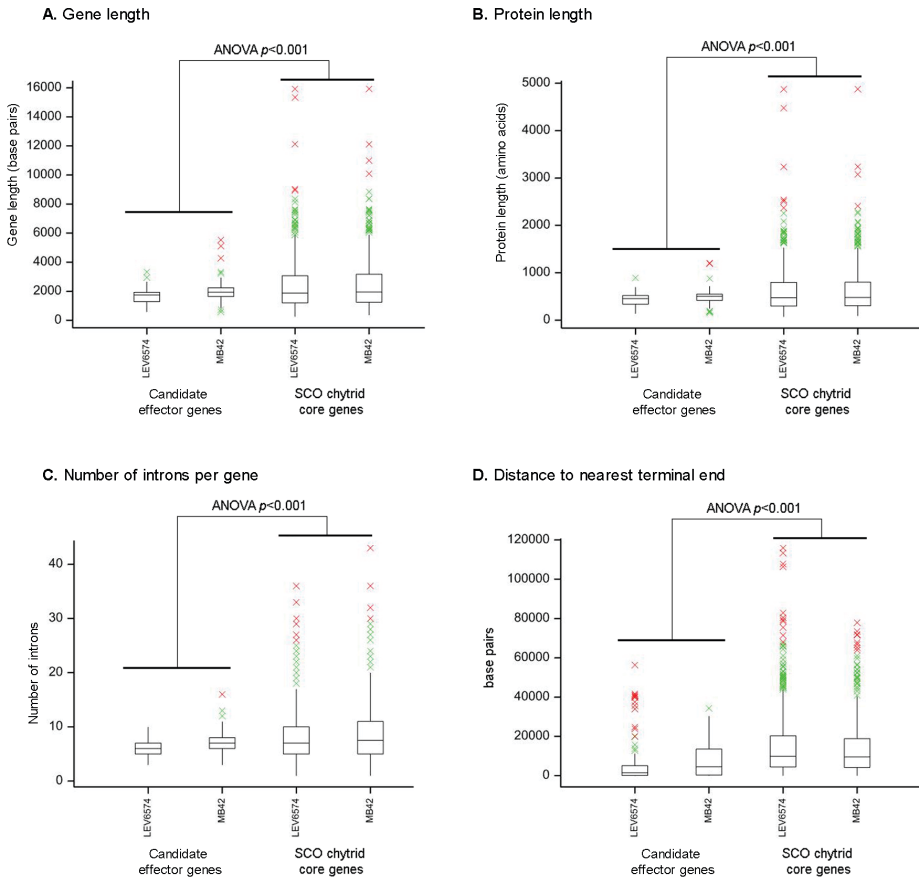


Figure S20. Comparison of gene and protein statistics for candidate effectors. With **A.** Gene length; **B.** protein length, **C.** the number of introns; and **D.** the distance to terminal ends of scaffolds and contigs.

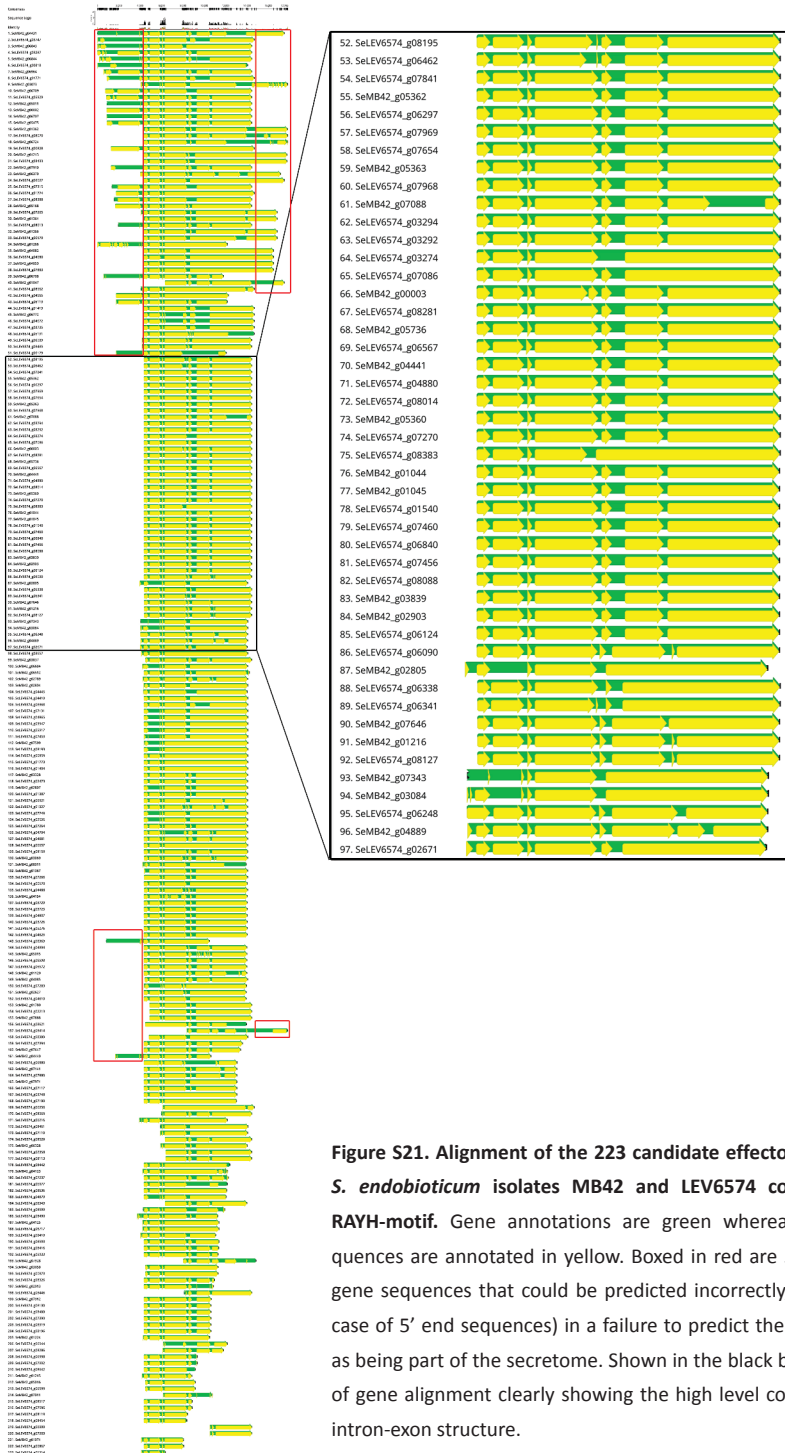


Figure S21. Alignment of the 223 candidate effector genes from *S. endobioticum* isolates MB42 and LEV6574 containing the RAYH-motif. Gene annotations are green whereas coding sequences are annotated in yellow. Boxed in red are 5' and 3' end gene sequences that could be predicted incorrectly, resulting (in case of 5' end sequences) in a failure to predict the gene models as being part of the secretome. Shown in the black box, is a detail of gene alignment clearly showing the high level conservation of intron-exon structure.

Table S1. Information on fungal isolates, NextGen sequence data, and genome sequences used in this study.

Species	Isolate	Origin	Sequence data used	SRA accession	Genome accession
<i>Chytridiomycota</i> isolates used in functional comparative genomics					
<i>Synchytrium endobioticum</i> pathotype 1(D1)	MB42	Langenboom, the Netherlands	454 Roche, GS-FLX	SRR8080735 - SRR8080742	QEAN000000000
			Titanium, 865.0 Mb		
			HiSeq PE100, 6.18 Gb	ERR2286952, ERR2286953	
			MiSeq PE250, 5.12 Gb	ERR2286954	
			PacBio SMRT, 65.3 Mb	ERR2698824	
			RNAseq, Pe125, 96.6 Gb	SRR8074753 - SRR8074755	
<i>Synchytrium endobioticum</i> pathotype 6(O1)	LEV6574	Prince Edward Island, Canada	MiSeq PE300, 38.9 Gb	ERR2286942, ERR2286943, ERR2286944, ERR2286945	QEAM000000000
				ERR2698823	
			PacBio SMRT, 35.5 Mb	ERR2698823	
			RNAseq, Pe150, 20.3 Gb	SRR8068361 - SRR8068376	
<i>Chytridium confervae</i>	CBS 675.73	Canada	HiSeq PE125, 3.50 Gb	ERR2451316	QEAP000000000
<i>Powellomyces hirtus</i>	CBS 809.83	the Netherlands	HiSeq PE125, 28.6 Gb	ERR2451317	QEAO000000000
<i>Spizellomyces palustris</i>	CBS 455.65	Germany	HiSeq PE125, 3.57 Gb	ERR2451319	QEAO000000000
<i>Synchytrium microbalum</i>	JEL17	USA	HiSeq PE125, 4.23 Gb	ERR2451318	QEAO000000000
<i>Batrachochytrium dendrobatidis</i>	JAM81	N.A.	N.A.	N.A.	GCF_000203795.1
<i>Batrachochytrium dendrobatidis</i>	JEL423	N.A.	N.A.	N.A.	GCA_000149865.1
<i>Gonapodya prolifera</i>	JEL478	N.A.	N.A.	N.A.	GCA_001574975.1
<i>Homolaphlyctis polyrhiza</i>	JEL142	N.A.	N.A.	N.A.	GCA_000235945.1
<i>Spizellomyces punctatus</i>	DAOM BR117	N.A.	N.A.	N.A.	GCF_000182565.1

Table S1 (continued). Information on fungal isolates, NextGen sequence data, and genome sequences used in this study.

Species	Isolate	Origin	Sequence data used	SRA accession	Genome accession
Ascomycota and Basidiomycota isolates used in functional comparative genomics					
<i>Cryptococcus neoformans</i> var. <i>neoformans</i>	JEC21	N.A.	N.A.	N.A.	GCA_000091045.1
<i>Melampsora larici-populina</i>	98AG31	N.A.	N.A.	N.A.	GCA_000204055.1
<i>Neurospora crassa</i>	OR74A	N.A.	N.A.	N.A.	GCA_000182925.2
<i>Puccinia graminis</i> f. sp. <i>tritici</i>	CR1 75-36-700-3	N.A.	N.A.	N.A.	GCA_000149925.1
<i>Saccharomyces cerevisiae</i>	S288C	N.A.	N.A.	N.A.	GCF_000146045.2
<i>Ustilago maydis</i>	521	N.A.	N.A.	N.A.	GCF_000328475.2
<i>Synchytrium endobioticum</i> isolates used in ZOO selection					
pathotype 1(D1)	NED01	the Netherlands	HiSeq PE100, 7.56 Gb	ERR2286957	N.A.
pathotype 2(G1)	MB08	the Netherlands	HiSeq PE125, 44.4 Gb	ERR2286949, ERR2286950	N.A.
pathotype 2(G1)	BBA 2(G1) 09-04 (SE4)	Germany	HiSeq PE125, 51.1 Gb	ERR2286960, ERR2286961	N.A.
pathotype 6(O1)	HLB 6(O1) 02-06 (SE5)	the Netherlands	HiSeq PE100, 6.05 Gb	ERR2286962	N.A.
pathotype 6(O1)	BBA 6(O1) 05_8.3 (SE6)	Germany	HiSeq PE100, 42.6 Gb	ERR2286963, ERR2286964	N.A.
pathotype 8(F1)	DEN01	Denmark	HiSeq PE125, 62.1 Gb	ERR2286937, ERR2286938	N.A.
pathotype 18(T1)	MB17	the Netherlands	HiSeq PE100, 6.94 Gb	ERR2286951	N.A.
pathotype 18(T1)	HLB P18(T1) -02-06 (SE7)	the Netherlands	HiSeq PE100, 7.91 Gb	ERR2286965	N.A.
pathotype 38(N1)	MB56	Turkey	HiSeq PE100, 34.5 Gb	ERR2286955, ERR2286956	N.A.

Table S2. Genome and annotation statistics of Chytridiomycota isolates included in this study.

Genome Assembly	<i>S. endobioticum</i> (MB42)	<i>S. endobioticum</i> (LEV6574)	<i>S. microbium</i> (JEL517)	<i>C. confervae</i> (CBS 675.73)	<i>B. dendrobatidis</i> (JAM81)	<i>B. dendrobatidis</i> (JEL423)	<i>H. polyrhiza</i> (JEL142)	<i>P. hitus</i> (CBS 809.83)	<i>S. punctatus</i> (DAOM BR117)	<i>S. palustris</i> (CBS 455.65)	<i>G. prolifera</i> (JEL478)
number of contigs	786	1285	169	2034	127	70	2582	482	38	213	352
genome size (Mb)	21.48	23.21	26.24	35.97	24.32	23.90	21.95	26.24	24.13	22.94	48.79
largest contig (Kb)	175.1	244.4	1,671.5	284.2	4,429.3	4,440.1	227.1	764.2	2,242.4	972.7	1,572.2
N50 (Kb)	44.1	40.5	518.6	44.3	1,484.5	1,707.3	54.5	157.5	1,465.7	219.3	347.3
GC content	47.0%	46.8%	43.7%	47.9%	39.3%	39.3%	49.8%	51.5%	47.6%	47.8%	52.4%
N's per 100 kbp	503	0	26	276	760	1339	1023	195	933	32	1244
median read coverage	229	215	146	69	n.a. ^b	n.a.	n.a.	959	n.a.	125	n.a.
Structural annotation											
number of genes	8031	8671	6308	10712	9879	8700	4849	6536	8950	8519	13827
mean gene length	1927	1819	2342	2045	1823	2045	2490	2063	2209	1946	1814
mean exons per gene	7	6	8	5	4	5	6	5	6	6	6
mean introns per gene	6	5	7	4	3	4	5	4	5	5	5
genome covered by genes	72.0%	68.0%	56.3%	60.9%	65.2%	73.9%	55.0%	51.4%	81.9%	72.3%	51.4%
gene models with support ^a	92.8%	89.7%	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Clusters of orthologous genes (COGs)											
Number of COGs	6208	6278	4982	4980	6473	6409	3897	5317	7469	7369	5243
Isolate specific genes	588 (7.3%)	569 (6.6%)	248 (3.9%)	2127 (19.9%)	604 (6.9%)	607 (6.1%)	370 (7.6%)	276 (4.2%)	397 (4.2%)	208 (2.4%)	5596 (40.5%)

a. RNAseq support at ≥ 3 reads per kilobase per million (RPKM); b. not analysed; no DNA and/or RNA NGS data was generated

The *S. endobioticum* nuclear genome**Table S3. Benchmarking Universal Single Copy Orthologs (BUSCO) statistics.** Statistics for reference set Fungi containing 290 conserved single copy genes are shown. The number and percentage of BUSCO genes found for each category (complete, complete (duplicated), fragmented and missing) is provided for the chytrid isolates analyzed.

Complete	260 (89.7%)	266 (91.7%)	273 (94.1%)	260 (89.7%)	282 (97.2%)	260 (89.7%)	244 (84.1%)	274 (94.5%)	281 (96.9%)	276 (95.2%)	268 (92.4%)
Complete (Duplicated)	7 (2.4%)	8 (2.8%)	4 (1.4%)	67 (23.1%)	7 (2.4%)	23 (7.9%)	2 (0.7%)	12 (4.1%)	18 (6.2%)	7 (2.4%)	10 (3.4%)
Fragmented	15 (5.2%)	12 (4.1%)	8 (2.8%)	13 (4.5%)	5 (1.7%)	23 (7.9%)	28 (9.7%)	9 (3.1%)	4 (1.4%)	10 (3.4%)	11 (3.8%)
Missing	15 (5.2%)	12 (4.1%)	9 (3.1%)	16 (5.5%)	3 (1.0%)	7 (2.4%)	18 (6.2%)	7 (2.4%)	5 (1.7%)	4 (1.4%)	11 (3.8%)

Chapter 2

Table S4. Genomes and download sources used for phylogenomics.

Phylum	Organism	Isolate	Version	Source ^a
Ascomycota	<i>Neurospora crassa</i>	OR74A	v2.0	1
	<i>Saccharomyces cerevisiae</i>	S288C		1
	<i>Schizosaccharomyces pombe</i>			1
Basidiomycota	<i>Agaricus bisporus</i> var. <i>bisporus</i>	H97	v2.0	1
	<i>Sporobolomyces roseus</i>		v1.0	1
	<i>Ustilago maydis</i>			1
Blastocladiomycota	<i>Blastocladiella</i> cf. <i>britannica</i>	JEL711		1
	<i>Allomyces macrogynus</i>	ATCC 38327	v1.0	1
	<i>Catenaria anguillulae</i>	PL171	v1.0	1
Chytridiomycota	<i>Chytrium</i> sp.	MP 71	v1.0	1
	<i>Gaertneriomyces semiglobifer</i>	Barr 43	v1.0	1
	<i>Globomyces pollinis-pini</i>	Arg68	v1.0	1
	<i>Batrachochytrium dendrobatidis</i>	JAM81		2
	<i>Batrachochytrium dendrobatidis</i>	JEL423		2
	<i>Gonapodya prolifera</i>	JEL478		2
	<i>Homolaphlyctis polyrhiza</i>	JEL142		2
	<i>Spizellomyces punctatus</i>	DAOM BR117	v1.0	2
	<i>Anaeromyces robustus</i>	S4	v1.0	1
	<i>Neocallimastix californiae</i>	G1	v1.0	1
	<i>Orpinomyces</i> sp.	C1A	v1.0	1
	<i>Piromyces finnis</i>		v3.0	1
	<i>Piromyces</i> sp.	E2	v1.0	1
	<i>Rhizoclostridium globosum</i>	JEL800		1
	<i>Chytridium confervae</i>	CBS 675.73		this study
	<i>Spizellomyces palustris</i>	CBS 455.65		this study
	<i>Powellomyces hirtus</i>	CBS 809.83		this study
	<i>Synchytrium endobioticum</i>	LEV6574		this study
	<i>Synchytrium endobioticum</i>	MB42		this study
	<i>Synchytrium microbalum</i>	JEL517		this study
Cryptomycota	<i>Rozella allomycis</i>	CSF55	v1.0	1
Glomeromycotina	<i>Rhizophagus irregularis</i>	DAOM 181602	v1.0	1
Microsporidia	<i>Anncaliia algerae</i>	PRA339		2
	<i>Edhazardia aedis</i>	USNM 41457		2
	<i>Mitosporidium daphniae</i>	UGP3		2
	<i>Ordospora colligata</i>	OC4		2
	<i>Spraguea lophii</i>	42_110		2

a. Genomic sequences downloaded from: JGI MycoCosm = 1; NCBI GenBank = 2

Table S4 (continued). Genomes and download sources used for phylogenomics.

Phylum	Organism	Isolate	Version	Source ^a
Microsporidia	<i>Vavraia culicis</i> subsp. <i>floridensis</i>			2
(continued)	<i>Vittaforma corneae</i>	ATCC 50505		2
	<i>Antonospora locustae</i>	HM-2013	v1.0	1
	<i>Encephalitozoon intestinalis</i>	ATCC 50506	v1.0	1
	<i>Enterocytozoon bieneusi</i>	H348	v1.0	1
	<i>Nematocida parisii</i>	ERTm1	v1.0	1
	<i>Nosema ceranae</i>	BRL01		1
Mortierellomycotina	<i>Lobosporangium transversale</i>	NRRL 3116	v1.0	1
	<i>Mortierella elongata</i>	AG-77	v2.0	1
Mucoromycota	<i>Backusella circina</i>	FSU 941	v1.0	1
	<i>Hesseltinella vesiculosa</i>	NRRL3301	v2.0	1
	<i>Lichtheimia hyalospora</i>		v1.0	1
	<i>Mucor circinelloides</i>	CBS 277.49	v2.0	1
	<i>Phycomyces blakesleeianus</i>	NRRL1555	v2.0	1
	<i>Rhizopus oryzae</i>		v1.0	1
	<i>Syncephalastrum racemosum</i>	NRRL 2496	v1.0	1
Zoopagomycota	<i>Basidiobolus meristosporus</i>	CBS 931.73	v1.0	1
	<i>Conidiobolus coronatus</i>	NRRL 28638	v1.0	1
	<i>Conidiobolus thromboides</i>	FSU 785	v1.0	1
	<i>Coemansia reversa</i>	NRRL 1564	v1.0	1
	<i>Linderina pennispora</i>	ATCC 12442	v1.0	1
	<i>Martensiomycetes pterosporus</i>	CBS 209.56	v1.0	1
	<i>Ramicandelaber brevisporus</i>	CBS 109374	v1.0	1
	<i>Syncephalis plumigaleata</i>	NRRL S24	v1.0	1

a. Genomic sequences downloaded from: JGI MycoCosm = 1; NCBI GenBank = 2

Table S7. Validation of KEGG pathway analysis using six Ascomycota and Basidiomycota organisms represented in the KEGG database. Accuracy (%) of the KEGG pathway prediction relative to the known pathway is presented.

Map	Pathway	Culturable species			Obligate biotrophic species		
		<i>Saccharomyces cerevisiae</i> (SCE)	<i>Neurospora crassa</i> (NCR)	<i>Cryptococcus neoformans</i> (CNE)	<i>Ustilago maydis</i> (UMA)	<i>Puccinia graminis f.sp. tritici</i> (PGR)	<i>Melampsora larici-populina</i> (MLR)
Carbohydrate metabolism							
10	Glycolysis / Gluconeogenesis	85%	87%	91%	88%	85%	79%
20	Citrate cycle (TCA cycle)	86%	87%	86%	87%	82%	87%
30	Pentose phosphate pathway	92%	88%	88%	84%	90%	90%
40	Pentose and glucuronate interconversions	89%	80%	87%	84%	92%	90%
51	Fructose and mannose metabolism	91%	86%	88%	91%	88%	89%
52	Galactose metabolism	93%	89%	89%	85%	89%	89%
53	Ascorbate and aldarate metabolism		95%	92%	89%	92%	92%
500	Starch and sucrose metabolism	91%	86%	86%	88%	84%	88%
520	Amino sugar and nucleotide sugar metabolism	93%	91%	93%	92%	90%	91%
562	Inositol phosphate metabolism	83%	79%	77%	75%	79%	77%
620	Pyruvate metabolism	84%	84%	86%	82%	86%	89%
630	Glyoxylate and dicarboxylate metabolism	83%	80%	76%	78%	83%	85%
640	Propanoate metabolism	93%	89%	91%	89%	89%	89%
650	Butanoate metabolism	87%	83%	84%	84%	86%	86%
660	C5-Branched dibasic acid metabolism	94%	89%	89%	76%	89%	89%
Energy metabolism							
190	Oxidative phosphorylation	92%	67%	62%	67%	65%	66%
680	Methane metabolism	95%	93%	94%	93%	96%	91%
910	Nitrogen metabolism	97%	81%	94%	90%	97%	97%
920	Sulfur metabolism	91%	87%	89%	89%	94%	91%
Lipid metabolism							
61	Fatty acid biosynthesis	52%	52%	48%	52%	43%	52%
62	Fatty acid elongation	70%	64%	64%	64%	64%	55%
71	Fatty acid degradation	71%	64%	70%	67%	65%	65%

Table S7 (continued). Validation of KEGG pathway analysis using six Ascomycota and Basidiomycota organisms represented in the KEGG database.

Map	Pathway	Culturable species			Obligate biotrophic species		
		<i>Saccharomyces cerevisiae</i> (SCE)	<i>Neurospora crassa</i> (NCR)	<i>Cryptococcus neoformans</i> (CNE)	<i>Ustilago maydis</i> (UMA)	<i>Puccinia graminis f.sp. tritici</i> (PGR)	<i>Melampsora larici-populina</i> (MLR)
Lipid metabolism (continued)							
72	Synthesis and degradation of ketone bodies	83%	67%	67%	50%	50%	50%
100	Steroid biosynthesis	54%	50%	57%	54%	64%	64%
561	Glycerolipid metabolism	74%	76%	77%	80%	74%	80%
564	Glycerophospholipid metabolism	70%	69%	75%	77%	74%	73%
565	Ether lipid metabolism	75%	81%	81%	82%	88%	81%
590	Arachidonic acid metabolism	92%	72%	88%	80%	88%	81%
592	alpha-Linolenic acid metabolism	88%	75%	73%	73%	73%	81%
600	Sphingolipid metabolism	61%	59%	59%	56%	59%	59%
Nucleotide metabolism							
230	Purine metabolism	83%	84%	85%	84%	87%	89%
240	Pyrimidine metabolism	83%	85%	87%	87%	87%	89%
Amino acid metabolism							
220	Arginine biosynthesis	87%	83%	87%	87%	90%	90%
250	Alanine, aspartate and glutamate metabolism	83%	80%	85%	85%	85%	85%
260	Glycine, serine and threonine metabolism	87%	82%	77%	80%	77%	79%
270	Cysteine and methionine metabolism	82%	87%	82%	84%	84%	84%
280	Valine, leucine and isoleucine degradation	85%	69%	69%	64%	64%	64%
290	Valine, leucine and isoleucine biosynthesis	92%	92%	92%	92%	92%	92%
300	Lysine biosynthesis	82%	79%	79%	85%	88%	85%
310	Lysine degradation	90%	73%	82%	85%	85%	80%
330	Arginine and proline metabolism	89%	86%	89%	91%	94%	94%

Table S7 (continued). Validation of KEGG pathway analysis using six Ascomycota and Basidiomycota organisms represented in the KEGG database.

Map	Pathway	Culturable species			Obligate biotrophic species		
		<i>Saccharomyces cerevisiae</i> (SCE)	<i>Neurospora crassa</i> (NCR)	<i>Cryptococcus neoformans</i> (CNE)	<i>Ustilago maydis</i> (UMA)	<i>Puccinia graminis f.sp. tritici</i> (PGR)	<i>Melampsora larici-populina</i> (MLR)
Amino acid metabolism (continued)							
340	Histidine metabolism	90%	80%	83%	90%	87%	93%
350	Tyrosine metabolism	91%	85%	87%	94%	89%	87%
360	Phenylalanine metabolism	88%	87%	88%	90%	92%	92%
380	Tryptophan metabolism	85%	75%	83%	81%	85%	88%
400	Phenylalanine, tyrosine and tryptophan biosynthesis	81%	74%	76%	79%	74%	76%
Metabolism of other amino acids							
410	beta-Alanine metabolism	82%	86%	86%	86%	86%	86%
430	Taurine and hypotaurine metabolism	100%	94%	100%	100%	100%	100%
440	Phosphonate and phosphinate metabolism	86%		92%	92%	89%	89%
450	Selenocompound metabolism	88%	88%	76%	82%	76%	82%
460	Cyanoamino acid metabolism	96%	82%	96%	96%	96%	96%
480	Glutathione metabolism	79%	79%	79%	76%	79%	76%
Glycan biosynthesis and metabolism							
510	N-Glycan biosynthesis	55%	61%	66%	58%	67%	56%
513	Various types of N-glycan biosynthesis	59%	68%	81%	81%	81%	78%
531	Glycosaminoglycan degradation		87%	87%	80%	87%	73%
563	Glycosylphosphatidylinositol (GPI)-anchor biosynthesis	17%	53%	56%	70%	56%	50%
Metabolism of cofactors and vitamins							
130	Ubiquinone and other terpenoid-quinone biosynthesis	78%	78%	80%	90%	88%	85%
670	One carbon pool by folate	88%	88%	88%	75%	88%	88%
730	Thiamine metabolism	91%	96%	92%	83%	91%	91%
740	Riboflavin metabolism	84%	92%	92%	88%	88%	88%
750	Vitamin B6 metabolism	90%	90%	81%	86%	90%	81%

Table S7 (continued). Validation of KEGG pathway analysis using six Ascomycota and Basidiomycota organisms represented in the KEGG database.

Map	Pathway	Culturable species			Obligate biotrophic species		
		<i>Saccharomyces cerevisiae</i> (SCE)	<i>Neurospora crassa</i> (NCR)	<i>Cryptococcus neoformans</i> (CNE)	<i>Ustilago maydis</i> (UMA)	<i>Puccinia graminis f.sp. tritici</i> (PGR)	<i>Melampsora larici-populina</i> (MLR)
Metabolism of cofactors and vitamins (continued)							
760	Nicotinate and nicotinamide metabolism	88%	91%	91%	93%	91%	91%
770	Pantothenate and CoA biosynthesis	76%	76%	76%	80%	80%	80%
780	Biotin metabolism	82%	75%	56%	69%	75%	69%
785	Lipoic acid metabolism	75%	75%	75%		75%	75%
790	Folate biosynthesis	89%	82%	89%	84%	91%	86%
860	Porphyrin and chlorophyll metabolism	99%	97%	96%	97%	98%	98%
Metabolism of terpenoids and polyketides							
900	Terpenoid backbone biosynthesis	86%	81%	79%	79%	81%	81%
909	Sesquiterpenoid and triterpenoid biosynthesis	98%	98%	98%	98%	98%	98%
Biosynthesis of other secondary metabolites							
261	Monobactam biosynthesis	100%	95%	95%	100%	100%	100%
332	Carbapenem biosynthesis	100%	100%	100%	100%	100%	100%
Xenobiotics biodegradation and metabolism							
791	Atrazine degradation	100%	100%				

Chapter 2

Table S8. Protein sequences assigned to enzymatic steps in KEGG pathways for the eleven Chytridiomycota species analyzed. Per reference pathway, the number of predicted enzymatic steps is presented.

Map	Pathway	<i>S. endobioticum</i> (MB42)	<i>S. endobioticum</i> (LEV6574)	<i>S. microbalum</i> (JEL517)	<i>C. confervae</i> (CBS 675.73)	<i>B. dendrobatidis</i> (JAM81)	<i>B. dendrobatidis</i> (JEL423)	<i>H. polyrhiza</i> (JEL142)	<i>P. hirtus</i> (CBS 809.83)	<i>S. punctatus</i> (DAOM BR117)	<i>S. palustris</i> (CBS 455.65)	<i>G. prolifera</i> (JEL478)
Carbohydrate metabolism												
10	Glycolysis / Gluconeogenesis	17	16	18	16	17	18	17	20	19	19	19
20	Citrate cycle (TCA cycle)	15	15	15	12	16	15	15	14	15	15	15
30	Pentose phosphate pathway	12	13	15	15	14	14	12	14	14	14	13
40	Pentose and glucuronate interconversions	5	5	6	5	3	3	3	4	3	3	5
51	Fructose and mannose metabolism	14	14	14	14	12	12	11	12	15	15	13
52	Galactose metabolism	8	8	9	10	7	8	8	9	8	9	8
53	Ascorbate and aldarate metabolism	3	3	3	3	1	1	2	3	2	2	2
500	Starch and sucrose metabolism	10	11	15	11	10	10	10	12	13	13	12
520	Amino sugar and nucleotide sugar metabolism	18	18	19	20	17	16	16	16	20	20	16
562	Inositol phosphate metabolism	8	8	6	6	7	6	5	9	9	9	5
620	Pyruvate metabolism	18	18	19	15	16	14	14	18	20	20	22
630	Glyoxylate and dicarboxylate metabolism	5	5	7	6	7	6	6	7	8	8	9
640	Propanoate metabolism	10	12	13	10	9	9	10	11	12	12	12
650	Butanoate metabolism	7	11	11	8	7	6	8	9	9	8	11
660	C5-Branched dibasic acid metabolism	3	3	3	3	1	1	2	3	3	3	3
Energy metabolism												
190	Oxidative phosphorylation	59	12	65	59	60	12	12	59	59	65	59
680	Methane metabolism	15	15	15	13	14	14	13	17	18	17	18

Table S8 (continued). Protein sequences assigned to enzymatic steps in KEGG pathways for the eleven Chytridiomycota species analyzed. Per reference pathway, the number of predicted enzymatic steps is presented.

Map	Pathway	<i>S. endobioticum</i> (MB42)	<i>S. endobioticum</i> (LEV6574)	<i>S. microbalum</i> (JEL517)	<i>C. confervae</i> (CBS 675.73)	<i>B. dendrobatidis</i> (JAM81)	<i>B. dendrobatidis</i> (JEL423)	<i>H. polyrhiza</i> (JEL142)	<i>P. hirtus</i> (CBS 809.83)	<i>S. punctatus</i> (DAOM BR117)	<i>S. palustris</i> (CBS 455.65)	<i>G. prolifera</i> (JEL478)
Energy metabolism (continued)												
910	Nitrogen metabolism	5	5	5	6	6	6	4	5	5	5	8
920	Sulfur metabolism	6	5	8	7	3	3	5	8	8	7	9
Lipid metabolism												
61	Fatty acid biosynthesis	9	9	9	9	8	9	9	9	9	9	5
62	Fatty acid elongation	3	6	6	6	7	7	6	6	7	7	6
71	Fatty acid degradation	0	5	7	7	7	7	6	7	8	8	6
72	Synthesis and degradation of ketone bodies	3	3	3	2	2	2	2	2	2	1	2
100	Steroid biosynthesis	1	1	1	1	1	1	1	1	1	1	1
561	Glycerolipid metabolism	5	6	8	7	6	6	6	8	7	7	9
564	Glycerophospholipid metabolism	12	13	16	16	17	17	17	20	22	22	18
565	Ether lipid metabolism	1	2	2	2	2	2	2	3	4	4	5
590	Arachidonic acid metabolism	2	2	2	2	2	2	1	2	6	6	3
592	alpha-Linolenic acid metabolism	0	1	3	3	3	3	2	3	5	5	4
600	Sphingolipid metabolism	5	6	6	7	1	3	2	5	5	5	4
Nucleotide metabolism												
230	Purine metabolism	35	36	37	35	38	33	26	34	32	37	38
240	Pyrimidine metabolism	16	16	19	19	18	17	16	17	17	17	19
Amino acid metabolism												
220	Arginine biosynthesis	11	13	14	14	10	9	10	14	14	14	14
250	Alanine, aspartate and glutamate metabolism	16	16	17	15	13	13	10	15	15	15	16
260	Glycine, serine and threonine metabolism	18	18	21	16	17	17	16	20	21	21	19
270	Cysteine and methionine metabolism	22	24	28	27	26	25	24	28	28	28	27

Table S8 (continued). Protein sequences assigned to enzymatic steps in KEGG pathways for the eleven Chytridiomycota species analyzed. Per reference pathway, the number of predicted enzymatic steps is presented.

Map	Pathway	<i>S. endobioticum</i> (MB42)	<i>S. endobioticum</i> (LEV6574)	<i>S. microbalum</i> (JEL517)	<i>C. confervae</i> (CBS 675.73)	<i>B. dendrobatidis</i> (JAM81)	<i>B. dendrobatidis</i> (JEL423)	<i>H. polyrhiza</i> (JEL142)	<i>P. hirtus</i> (CBS 809.83)	<i>S. punctatus</i> (DAOM BR117)	<i>S. palustris</i> (CBS 455.65)	<i>G. prolifera</i> (JEL478)
Amino acid metabolism (continued)												
280	Valine, leucine and isoleucine degradation	8	12	13	12	11	10	12	11	12	11	11
290	Valine, leucine and isoleucine biosynthesis	9	9	9	9	5	5	7	9	9	9	9
300	Lysine biosynthesis	5	6	6	6	7	7	6	7	7	7	7
310	Lysine degradation	6	8	8	7	7	7	7	7	7	7	7
330	Arginine and proline metabolism	9	10	11	11	11	11	9	11	11	11	13
340	Histidine metabolism	9	9	9	12	9	9	8	9	9	9	7
350	Tyrosine metabolism	5	5	6	5	6	6	4	6	7	7	4
360	Phenylalanine metabolism	3	4	4	4	4	4	3	5	5	5	3
380	Tryptophan metabolism	4	6	8	7	9	8	7	8	8	8	10
400	Phenylalanine, tyrosine and tryptophan synthesis	16	18	23	22	22	23	20	23	23	23	21
Metabolism of other amino acids												
410	beta-Alanine metabolism	6	8	8	5	5	5	4	6	7	7	6
430	Taurine and hypotaurine metabolism	5	5	6	4	3	3	3	5	5	5	6
440	Phosphonate and phosphinate metabolism	0	1	1	0	1	1	1	1	1	1	1
450	Selenocompound metabolism	4	4	6	5	3	3	4	5	5	5	6
460	Cyanoamino acid metabolism	2	2	2	3	2	2	2	2	2	2	3
480	Glutathione metabolism	14	14	14	14	12	12	13	14	14	14	14
Glycan biosynthesis and metabolism												
510	N-Glycan biosynthesis	6	6	6	6	6	6	6	7	6	6	7
513	Various types of N-glycan biosynthesis	7	7	7	7	7	7	7	7	7	7	7

Table S8 (continued). Protein sequences assigned to enzymatic steps in KEGG pathways for the eleven Chytridiomycota species analyzed. Per reference pathway, the number of predicted enzymatic steps is presented.

Map	Pathway	<i>S. endobioticum</i> (MB42)	<i>S. endobioticum</i> (LEV6574)	<i>S. microbalum</i> (JEL517)	<i>C. confervae</i> (CBS 675.73)	<i>B. dendrobatidis</i> (JAM81)	<i>B. dendrobatidis</i> (JEL423)	<i>H. polyrhiza</i> (JEL142)	<i>P. hirtus</i> (CBS 809.83)	<i>S. punctatus</i> (DAOM BR117)	<i>S. palustris</i> (CBS 455.65)	<i>G. prolifera</i> (JEL478)
Glycan biosynthesis and metabolism (continued)												
531	Glycosaminoglycan degradation	4	4	4	4	2	3	3	2	2	2	5
563	GPI-anchor biosynthesis	2	2	2	1	1	1	1	1	1	1	1
Metabolism of cofactors and vitamins												
130	Ubiquinone and other terpenoid-quinone synthesis	4	4	4	4	5	5	4	5	5	5	3
670	One carbon pool by folate	9	10	11	11	10	10	10	12	12	12	11
730	Thiamine metabolism	4	4	8	8	6	3	3	8	8	8	6
740	Riboflavin metabolism	6	6	5	5	6	6	6	6	6	6	6
750	Vitamin B6 metabolism	4	4	4	4	3	3	4	4	4	4	3
760	Nicotinate and nicotinamide metabolism	9	9	11	12	10	9	11	10	9	7	12
770	Pantothenate and CoA biosynthesis	11	11	11	11	8	8	8	11	11	11	10
780	Biotin metabolism	2	2	2	5	4	3	4	3	3	3	0
785	Lipoic acid metabolism	2	2	2	2	2	2	2	2	2	2	2
790	Folate biosynthesis	9	10	12	10	11	11	10	13	13	13	11
860	Porphyrin and chlorophyll metabolism	13	12	15	17	11	11	12	13	14	14	15
Metabolism of terpenoids and polyketides												
900	Terpenoid backbone biosynthesis	8	10	9	10	10	10	10	10	11	9	9
909	Sesquiterpenoid and triterpenoid biosynthesis	1	1	1	1	1	1	1	1	1	1	1
Biosynthesis of other secondary metabolites												
261	Monobactam synthesis	3	3	3	3	2	2	3	3	3	3	4
332	Carbapenem biosynthesis	2	2	2	2	2	2	0	2	2	2	2
Xenobiotics biodegradation and metabolism												
791	Atrazine degradation	1	1	1	1	1	0	0	1	1	1	1

Table S11. Presence of core meiotic genes as defined by Halary et al. (2011) in Chytridiomycota isolates analyzed. Meiosis specific genes are indicated with “*”, and genes that were not identified in all species or a given taxon by Halary et al. (2011) are indicated with “\$”. Core meiosis genes that were identified are indicated with “+”, whereas genes that were not detected are indicated with “-”.

Meiotic toolbox genes	Spol1*	Hfm1*/Mer3*	Mre11	Rad50	Rad51	Dmc1*	Rad52	Pch2*	Mih1	Mih3	Msh4*	Msh5*	Rad1	Hop1*	Hop2*	Mnd1*
Halary et al. 2011																
Ascomycota (4)	+	+	+	+	+	\$	+	\$	+	+	+	+	+	\$	\$	\$
Basidiomycota (3)	+	+	+	+	+	\$	+	+	+	+	+	+	+	\$	\$	\$
Zygomycota (2)	+	\$	+	+	+	+	+	-	+	+	+	+	+	-	+	+
Glomeromycota (1)	+	-	+	+	+	+	+	-	+	+	+	+	+	-	+	+
Chytridiomycota (2)	+	+	+	+	+	\$	+	+	+	\$	+	+	+	\$	+	+
This study																
<i>S. endobioticum</i> (MB42)	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	-
<i>S. endobioticum</i> (LEV6574)	+	+	+	+	-	+	+	+	+	-	+	+	+	+	+	-
<i>S. microbalum</i> (JEL517)	+	+	+	+	-	+	+	+	+	-	+	+	+	+	+	+
<i>C. confervae</i> (CBS 675.73)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>B. dendrobatidis</i> (JAM81)	+	+	+	+	-	+	+	+	+	-	+	+	+	+	+	+
<i>B. dendrobatidis</i> (JEL423)	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+
<i>H. polyrhiza</i> (JEL142)	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	+
<i>P. hirtus</i> (CBS 809.83)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>S. punctatus</i> (DAOM BR117)	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+
<i>S. palustris</i> (CBS 455.65)	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+
<i>G. prolifera</i> (JEL478)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Table S11 (continued). Presence of core meiotic genes as defined by Halary *et al.* (2011) in Chytridiomycota isolates analyzed. Meiosis specific genes are indicated with “*”, and genes that were not identified in all species or a given taxon by Halary *et al.* (2011) are indicated with “\$”. Core meiosis genes that were identified are indicated with “+”, whereas genes that were not detected are indicated with “-”.

Meiotic toolbox genes	Halary <i>et al.</i> 2011														
	<i>Smc5</i>	<i>Smc6</i>	<i>Msh2</i>	<i>Msh6</i>	<i>Mlh2</i>	<i>Pms1</i>	<i>Mus81</i>	<i>Rec8*</i>	<i>Rad21</i>	<i>Smc1</i>	<i>Smc2</i>	<i>Smc3</i>	<i>Smc4</i>	<i>Scs3</i>	<i>Pds5</i>
Ascomycota (4)	+	+	+	+	\$	+	+	+	+	+	+	+	+	+	+
Basidiomycota (3)	+	+	+	+	-	+	\$	+	+	+	+	+	+	+	+
Zygomycota (2)	+	+	+	+	-	+	+	+	+	\$	\$	+	+	+	+
Glomeromycota (1)	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+
Chytridiomycota (2)	+	+	+	+	-	+	\$	+	+	+	+	+	+	-	+
This study															
<i>S. endobioticum</i> (MB42)	+	+	+	+	+	+	-	+	+	+	+	+	+	-	+
<i>S. endobioticum</i> (LEV6574)	+	+	+	+	+	+	-	+	+	+	+	+	+	-	+
<i>S. microbalum</i> (JEL517)	+	+	+	+	+	+	-	+	+	+	+	+	+	-	+
<i>C. confervae</i> (CBS 675.73)	+	+	+	+	+	+	-	+	+	+	+	+	+	-	+
<i>B. dendrobatidis</i> (JAM81)	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
<i>B. dendrobatidis</i> (JEL423)	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
<i>H. polyrhiza</i> (JEL142)	+	+	+	+	+	+	+	-	+	+	+	+	+	-	+
<i>P. hirtus</i> (CBS 809.83)	+	+	+	+	+	+	-	+	+	+	+	+	+	-	+
<i>S. punctatus</i> (DAOM BR117)	+	+	+	+	+	+	-	+	+	+	+	+	+	-	+
<i>S. palustris</i> (CBS 455.65)	+	+	+	+	+	-	+	+	+	+	+	+	+	-	+
<i>G. proliferans</i> (JEL478)	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+

Chapter 2

Table S12. Repeat content (total, small simple repeats, and complex repeats) in the Chytridiomycota species analyzed.

Genome	Genome size (Mb)	Total repeat sequences		Simple repeats		Complex repeats	
		(kb)	(%)	(kb)	(%)	(kb)	(%)
<i>S. endobioticum</i> (MB42)	21.5	2,705.5	13%	128.3	5%	2,577.3	95%
<i>S. endobioticum</i> (LEV6574)	23.2	3,754.7	16%	131.8	4%	3,623.0	96%
<i>B. dendrobatidis</i> (JAM81)	24.3	4,897.8	20%	63.1	1%	4,834.7	99%
<i>B. dendrobatidis</i> (JEL423)	23.9	4,821.7	20%	62.0	1%	4,759.6	99%
<i>C. confervae</i> (CBS 675.73)	36.0	1,558.5	4%	456.5	29%	1,102.0	71%
<i>G. prolifera</i> (JEL478)	48.8	2,528.8	5%	485.8	19%	2,043.1	81%
<i>H. polyrhiza</i> (JEL142)	21.9	1,625.1	7%	815.8	50%	809.3	50%
<i>P. hirtus</i> (CBS 809.83)	26.2	1,243.7	5%	325.2	26%	918.4	74%
<i>S. palustris</i> (CBS 455.65)	22.9	2,84.6	1%	127.9	45%	156.7	55%
<i>S. microbalum</i> (JEL517)	26.2	242.9	1%	133.6	55%	109.4	45%
<i>S. punctatus</i> (DOAM BR117)	24.1	1,183.3	5%	126.1	11%	1,057.2	89%

Table S13. RID protein sequences used for the detection of RID homologs in Chytridiomycota genomes

Accession	Organism	Length (AA)
ELA28362	<i>Colletotrichum gloeosporioides</i>	722
KZL67337	<i>Colletotrichum tofieldiae</i>	749
KKY30018	<i>Diaporthe ampelina</i>	759
KKY33519	<i>Diaporthe ampelina</i>	766
KKY38719	<i>Diaporthe ampelina</i>	1158
OCW32809	<i>Diaporthe helianthi</i>	1271
XP_007790886	<i>Eutypa lata</i> UCREL1	437
EJT70672	<i>Gaeumannomyces tritici</i> R3-111a-1	954
EHA46203	<i>Magnaporthe oryzae</i> 70-15	882
KLU83146	<i>Magnaporthe oryzae</i> ATCC 64411	895
AAM27408	<i>Neurospora crassa</i>	845
ESA43828	<i>Neurospora crassa</i> OR74A	837
AAM27409	<i>Neurospora intermedia</i>	839
AAM27410	<i>Neurospora tetrasperma</i>	820
EGO51717	<i>Neurospora tetrasperma</i> FGSC 2508	817
EGZ78284	<i>Neurospora tetrasperma</i> FGSC 2509	817
XP_007916180	<i>Phaeoacremonium minimum</i> UCRPA7	317
XP_003007578	<i>Verticillium alfalfae</i> VaMs.102	909

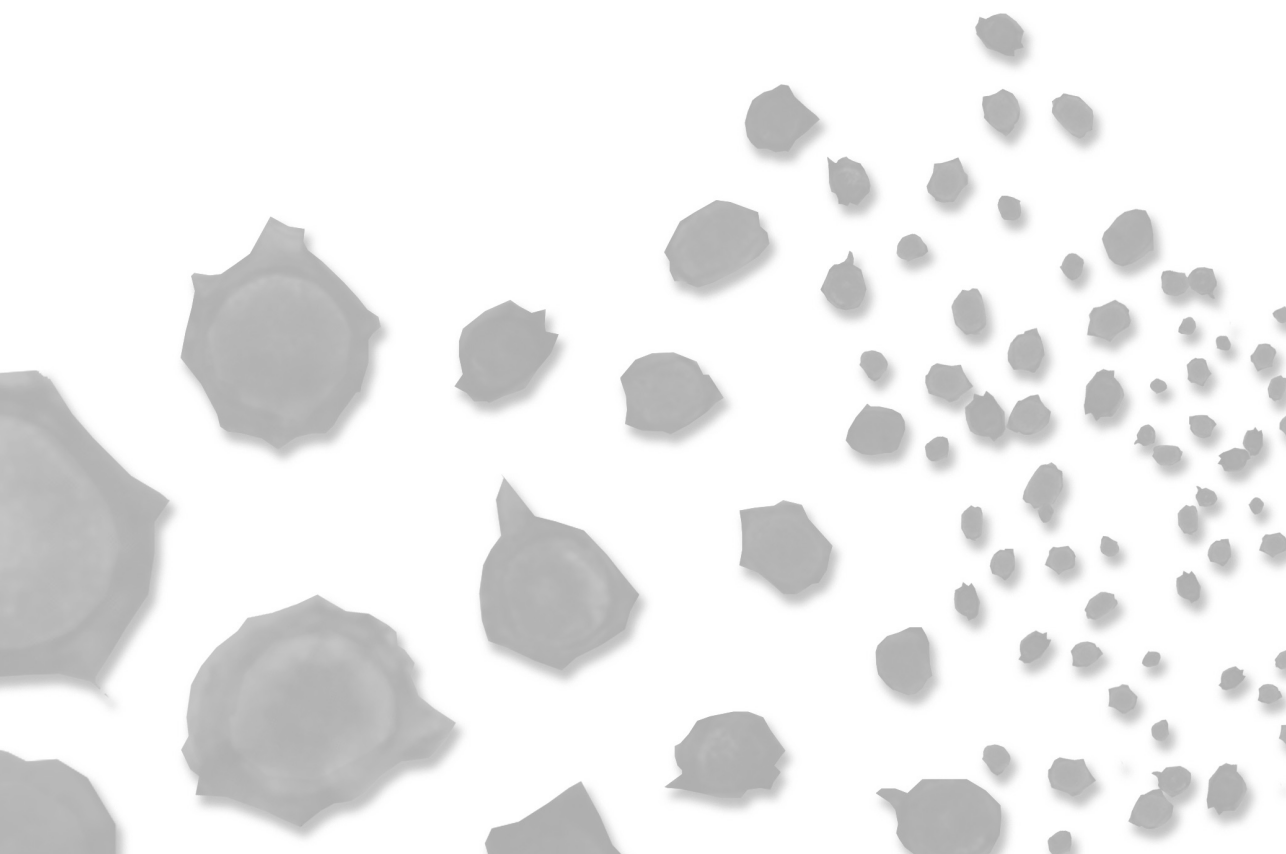
References supplementary material

1. van de Vossenbergh, B.T.L.H., *et al.*, The linear mitochondrial genome of the quarantine chytrid *Synchytrium endobioticum*; insights into the evolution and recent history of an obligate biotrophic plant pathogen. *BMC Evolutionary Biology*, 2018. 18(1): p. 136.
2. Bonants, P.J.M., *et al.*, A real-time TaqMan PCR assay to discriminate between pathotype 1 (D1) and non-pathotype 1 (D1) isolates of *Synchytrium endobioticum*. *European Journal of Plant Pathology*, 2015. 143(3): p. 495-506.
3. Bolger, A.M., M. Lohse, and B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 2014. 30(15): p. 2114-2120.
4. The Potato Genome Sequencing, C., *et al.*, Genome sequence and analysis of the tuber crop potato. *Nature*, 2011. 475: p. 189.
5. Bankevich, A., *et al.*, SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, 2012. 19(5): p. 455-477.
6. Li, H. and R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 2009. 25(14): p. 1754-60.
7. Myers, E.W., *et al.*, A whole-genome assembly of *Drosophila*. *Science*, 2000. 287(5461): p. 2196-204.
8. Walker, B.J., *et al.*, Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One*, 2014. 9(11): p. e112963.
9. Kumar, S., *et al.*, Blobology: exploring raw genome data for contaminants, symbionts and parasites using taxon-annotated GC-coverage plots. *Front Genet*, 2013. 4: p. 237.
10. Trapnell, C., L. Pachter, and S.L. Salzberg, TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*, 2009. 25(9): p. 1105-1111.
11. Hoff, K.J., *et al.*, BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics*, 2016. 32(5): p. 767-9.
12. Holt, C. and M. Yandell, MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics*, 2011. 12: p. 491-491.
13. Lomsadze, A., P.D. Burns, and M. Borodovsky, Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm. *Nucleic Acids Res*, 2014. 42(15): p. e119.
14. Ter-Hovhannisyan, V., *et al.*, Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Res*, 2008. 18(12): p. 1979-90.
15. Chang, Y., *et al.*, Phylogenomic Analyses Indicate that Early Fungi Evolved Digesting Cell Walls of Algal Ancestors of Land Plants. *Genome Biol Evol*, 2015. 7(6): p. 1590-601.
16. Mondo, S.J., *et al.*, Widespread adenine N6-methylation of active genes in fungi. *Nat Genet*, 2017. 49(6): p. 964-968.
17. Russ, C., *et al.*, Genome Sequence of *Spizellomyces punctatus*. *Genome Announc*, 2016. 4(4).
18. Haitjema, C.H., *et al.*, A parts list for fungal cellulosomes revealed by comparative genomics. *Nat Microbiol*, 2017. 2: p. 17087.
19. Youssef, N.H., *et al.*, The genome of the anaerobic fungus *Orpinomyces* sp. strain C1A reveals the unique evolutionary history of a remarkable plant biomass degrader. *Appl Environ Microbiol*, 2013. 79(15): p. 4620-34.

Chapter 2

20. Simao, F.A., *et al.*, BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 2015. 31(19): p. 3210-3212.
21. Jones, P., *et al.*, InterProScan 5: genome-scale protein function classification. *Bioinformatics*, 2014. 30(9): p. 1236-40.
22. Gurevich, A., *et al.*, QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, 2013. 29(8): p. 1072-5.
23. Mohanta, T.K. and H. Bae, The diversity of fungal genome. *Biol Proced Online*, 2015. 17: p. 8.
24. Smith, T.F. and M.S. Waterman, Identification of common molecular subsequences. *J Mol Biol*, 1981. 147(1): p. 195-7.
25. Emms, D.M. and S. Kelly, OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, 2015. 16(1): p. 1-14.
26. Spatafora, J.W., *et al.*, A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia*, 2016. 108(5): p. 1028-1046.
27. James, Timothy Y., *et al.*, Shared Signatures of Parasitism and Phylogenomics Unite Cryptomycota and Microsporidia. *Current Biology*, 2013. 23(16): p. 1548-1553.
28. Eddy, S.R., Accelerated Profile HMM Searches. *PLOS Computational Biology*, 2011. 7(10): p. e1002195.
29. Capella-Gutiérrez, S., J.M. Silla-Martínez, and T. Gabaldón, trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 2009. 25(15): p. 1972-1973.
30. Stamatakis, A., RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 2014. 30(9): p. 1312-1313.
31. Mirarab, S. and T. Warnow, ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics*, 2015. 31(12): p. i44-52.
32. Zerillo, M.M., *et al.*, Carbohydrate-Active Enzymes in Pythium and Their Role in Plant Cell Wall and Storage Polysaccharide Degradation. *PLOS ONE*, 2013. 8(9): p. e72572.
33. Yin, Y., *et al.*, dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Research*, 2012. 40(Web Server issue): p. W445-W451.
34. Cantarel, B.L., *et al.*, The Carbohydrate-Active EnZymes database (CAZY): an expert resource for Glycogenomics. *Nucleic Acids Res*, 2009. 37(Database issue): p. D233-8.
35. Halary, S., *et al.*, Conserved meiotic machinery in *Glomus* spp., a putatively ancient asexual fungal lineage. *Genome Biol Evol*, 2011. 3: p. 950-8.
36. Smit, A.F.A. and R. Hubley. RepeatModeler Open-1.0. 2008 [cited 2018 21 September]; Available from: <http://www.repeatmasker.org>.
37. Smit, A.F.A. and R. Hubley. RepeatMasker Open-4.0. 2013-2015 [cited 2018 21 September]; Available from: www.repeatmasker.org.
38. Katoh, K. and D.M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*, 2013. 30(4): p. 772-80.
39. Kearse, M., *et al.*, Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 2012. 28(12): p. 1647-9.
40. Hane, J.K. and R.P. Oliver, RIPCAL: a tool for alignment-based analysis of repeat-induced point mutations in fungal genomic sequences. *BMC Bioinformatics*, 2008. 9(1): p. 478.
41. Petersen, T.N., *et al.*, SignalP 4.0: discriminating signal peptides from transmembrane regions. 2011. 8: p. 785.

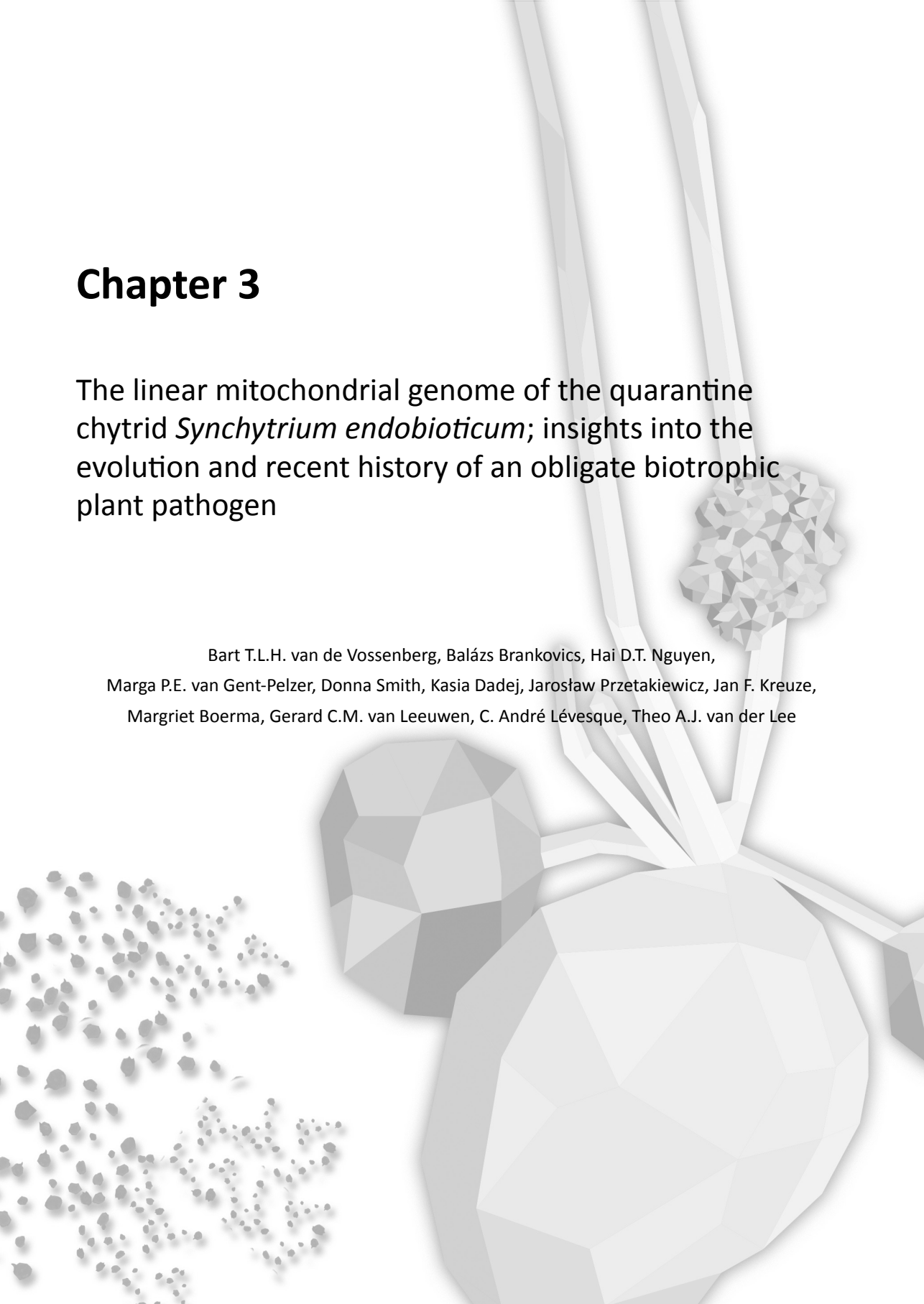
42. Krogh, A., *et al.*, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*, 2001. 305(3): p. 567-80.
43. Bailey, T.L. and C. Elkan, Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol*, 1994. 2: p. 28-36.
44. Bailey, T.L. and M. Gribskov, Combining evidence using p-values: application to sequence homology searches. *Bioinformatics*, 1998. 14(1): p. 48-54.
45. Emanuelsson, O., *et al.*, Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol*, 2000. 300(4): p. 1005-16.
46. Sperschneider, J., *et al.*, EffectorP: predicting fungal effector proteins from secretomes using machine learning. *New Phytol*, 2016. 210(2): p. 743-61.
47. Sperschneider, J., *et al.*, ApoplastP: prediction of effectors and plant proteins in the apoplast using machine learning. *bioRxiv*, 2017.
48. Sperschneider, J., *et al.*, LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. 2017. 7: p. 44598.
49. Testa, A.C., R.P. Oliver, and J.K. Hane, OcculterCut: A Comprehensive Survey of AT-Rich Regions in Fungal Genomes. *Genome Biol Evol*, 2016. 8(6): p. 2044-64.
50. Gladyshev, E., Repeat-Induced Point Mutation (RIP) and Other Genome Defense Mechanisms in Fungi. *Microbiology spectrum*, 2017. 5(4): p. 10.1128/microbiolspec.FUNK-0042-2017.



Chapter 3

The linear mitochondrial genome of the quarantine chytrid *Synchytrium endobioticum*; insights into the evolution and recent history of an obligate biotrophic plant pathogen

Bart T.L.H. van de Vossenberg, Balázs Brankovics, Hai D.T. Nguyen,
Marga P.E. van Gent-Pelzer, Donna Smith, Kasia Dadej, Jarosław Przetakiewicz, Jan F. Kreuze,
Margriet Boerma, Gerard C.M. van Leeuwen, C. André Lévesque, Theo A.J. van der Lee



Chapter 3

This chapter was published as “van de Vossenbergh B.T.L.H., Brankovics B., Nguyen H.D.T., van Gent-Pelzer M.P.E., Smith D., Dadej K., Przetakiewicz J., Kreuze J.F., Boerma M., van Leeuwen G.C.M., Lévesque C.A., and van der Lee T.A.J. (2018) The linear mitochondrial genome of the quarantine chytrid *Synchytrium endobioticum*; insights into the evolution and recent history of an obligate biotrophic plant pathogen, BMC Evol Biol. 18: 136” with additional supplementary materials. Slight formatting changes were made to facilitate thesis consistency.

Abstract

Chytridiomycota species (chytrids) belong to a basal lineage in the fungal kingdom. Inhabiting terrestrial and aquatic environments, most are free-living saprophytes but several species cause important diseases: e.g. *Batrachochytrium dendrobatidis*, responsible for worldwide amphibian decline; and *Synchytrium endobioticum*, causing potato wart disease. *S. endobioticum* has an obligate biotrophic lifestyle and isolates can be further characterized as pathotypes based on their virulence on a differential set of potato cultivars. Quarantine measures have been implemented globally to control the disease and prevent its spread. We used a comparative approach using chytrid mitogenomes to determine taxonomical relationships and to gain insights into the evolution and recent history of introductions of this plant pathogen.

We assembled and annotated the complete mitochondrial genome of 30 *S. endobioticum* isolates and generated mitochondrial genomes for five additional chytrid species. The mitochondrial genome of *S. endobioticum* is linear with terminal inverted repeats which was validated by tailing and PCR amplifying the telomeric ends. Surprisingly, no conservation in organization and orientation of mitochondrial genes was observed among the Chytridiomycota except for *S. endobioticum* and its sister species *Synchytrium microbalum*. However, the mitochondrial genome of *S. microbalum* is circular and comprises only a third of the 72.9 Kbp found for *S. endobioticum* suggesting recent linearization and expansion. Four mitochondrial lineages were identified in the *S. endobioticum* mitochondrial genomes. Several pathotypes occur in different lineages, suggesting that these have emerged independently. In addition, variations for polymorphic sites in the mitochondrial genome of individual isolates were observed demonstrating that *S. endobioticum* isolates represent a community of different genotypes. Such communities were shown to be complex and stable over time, but we also demonstrate that the use of semi-resistant potato cultivars triggers a rapid shift in the mitochondrial haplotype associated with increased virulence.

Mitochondrial genomic variation shows that *S. endobioticum* has been introduced into Europe multiple times, that several pathotypes emerged multiple times, and that isolates represent communities of different genotypes. Our study represents the most comprehensive dataset of chytrid mitogenomes, which provides new insights into the extraordinary dynamics and evolution of mitochondrial genomes involving linearization, expansion and reshuffling.

Background

Synchytrium endobioticum (Schilb.) Perc. is a chytrid fungus (Chytridiomycota) causing potato wart disease, which is one of the most important quarantine diseases on cultivated potato [1, 2]. Host tissues infected with this fungus develop tumor-like galls (i.e. warts) rendering the crop unmarketable. The malformed tissue produces resting spores which are released into the surrounding soil when the host tissue decays, and once released, spores can remain viable and infectious for decades [3, 4]. Individual isolates can be further characterized as pathotypes based on their virulence on a differential set of potato cultivars. Reported from many countries worldwide, *S. endobioticum* is believed to originate from potato-growing areas in the Andean region from where it was taken into the United Kingdom in the aftermath of the Irish potato famine (~1880s) [5]. Initially, the disease spread quickly and widely in Europe, but phytosanitary measures restricted its movement to other parts of the world [2, 5]. Today, quarantine measures have been implemented through phytosanitary legislation world-wide. In addition, *S. endobioticum* is on the federal select agent program of the United States of America [6]. Restricting potato cultivation in infested fields and the use of resistant potato cultivars are the only known efficient control strategies. Characterization of potato wart resistance in potato cultivars is essential for efficient control.

Forty-five years after the first description of *S. endobioticum* [7], only a single pathotype was known, which is nowadays known as pathotype 1(D1). In 1941, wart development was discovered on formerly resistant potato cultivars (Edda, Edelragis, Parnassia, Primula, Sabina, and Sickingen) in Gießübel, Germany [8] and it was recognized that a new pathotype, pathotype 2(G1), of the fungus had emerged. Subsequently, new pathotypes were found in Germany, the Czech Republic, and Ukraine [9], and at present, 39 pathotypes are described with the most recent one being discovered in Piekielnik, Poland [10]. Although pathotype identity is essential for authorities to enforce suitable phytosanitary measures, these tests do not provide information on the migration and spread of the fungus, or the emergence of new pathotypes.

To infer intraspecific genetic variation, Gagnon *et al.* [11] developed polymorphic microsatellite loci for *S. endobioticum*. However, isolates of the same pathotype were found to display highly variable genotypes and none of the markers correlated with a specific pathotype. Apart from several Canadian isolates, no clustering based on geographical origin was observed. We have pursued a different approach, exploiting the mitochondrial genome

to determine taxonomical relationships, and to gain insights into the introductions and population dynamics of this plant pathogen.

The mitochondrion is a cell organelle involved in ATP production, and is descended from an α -proteobacterium [12]. They have maintained the double membrane characteristic of their ancestors and retained a highly reduced bacterial-like genome [13, 14]. The mitochondrial genome is typically assumed to be a single maternally inherited circular molecule that is present in high copy number, ranging from a few hundred to a few thousand copies per cell. Loci from the mitochondria are most frequently used in phylogenetic and population-level studies [15-17], because of their copy number, commonly available generic primers for amplification and sequencing, and phylogenetic signal on higher taxonomical levels [18, 19]. With the introduction of next generation sequencing technologies, the number of complete mitochondrial genomes has rapidly increased [14]. However, complete mitochondrial genomes of the early diverging fungal lineages of the Chytridiomycota and Blastocladiomycota are scarce with only eight publically available complete and annotated mitochondrial genomes. This limits the possibilities for species level analyses using the mitochondrial genome in these divisions.

The aims for this study were to: 1 sequence, assemble and annotate the complete mitochondrial genome of *S. endobioticum*; 2 perform phylogenetic analyses using the mitochondrial genes to determine its taxonomical relationships to 14 other species from the Chytridiomycota including *Synchytrium taraxaci*, which is the type species for the genus [20], and *Synchytrium microbalum*, the first culturable *Synchytrium* species to be described [21]; and 3 perform network analyses using a panel of *S. endobioticum* isolates from different countries covering several pathotypes and identify potential linkages between genotypes, geographical origins and pathotypes.

Results

Assembly of the *S. endobioticum* mitochondrial genome

We assembled and annotated the mitochondrial genome of *S. endobioticum* pathotype 1(D1) isolate MB42. Three independent assemblies using different datasets were performed and the results were combined to obtain a consensus mitochondrial genome; 1 from the MB42 reference genome, seven putative mitochondrial scaffolds were identified representing a total size of 68,068 base pairs; 2 using GRAB, a single 63,627 bp putative mitochondrial

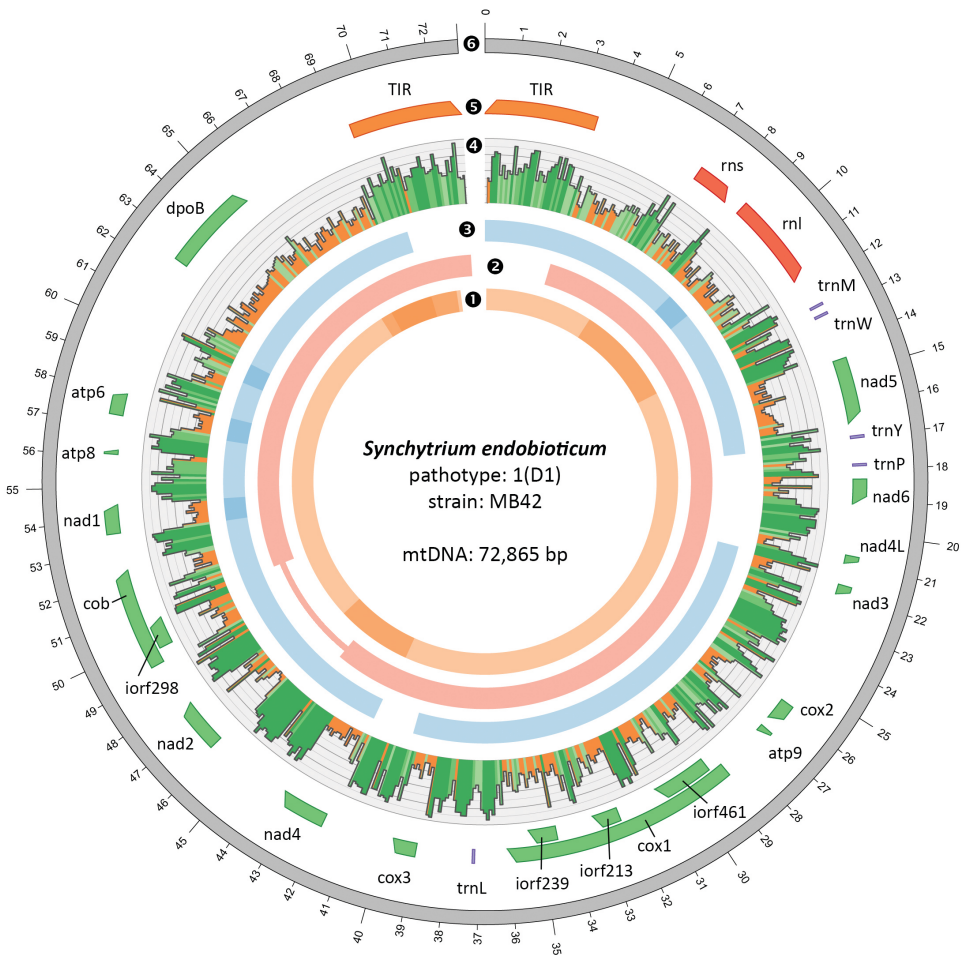


Figure 1. Assembly and annotation of the linear *S. endobioticum* mtDNA genome. Tracks 1 to 3 represent putative mtDNA scaffolds from three different assemblies (respectively putative mtDNA scaffolds from the *S. endobioticum* MB42 genome assembly, reference guided assembly with GRABb using *S. endobioticum* MB42 HiSeq data, and a reference guided assembly using *S. endobioticum* MB42 PacBio CCS) mapped to the linear mtDNA genome. Darker shades are used to indicate regions where scaffolds from the same assembly overlap. The narrow line in track 2 indicates a gap relative to the linear mtDNA genome. **4** G+C content determined with a 100 bp sliding window in which orange ≤ 40% G+C, light green 40 - 50 % G+C, green 50 - 60% G+C, and dark green >60% G+C. **5** mtDNA annotation track showing genes (green), rRNA (red), tRNA (purple) and the terminal inverted repeats (orange). All genes, rRNAs and tRNAs are orientated in the 5' to 3'direction. **6** the 72,865 bp linear mitochondrial genome of *S. endobioticum*.

scaffold was obtained; 3 an assembly of PacBio data resulted in five putative mitochondrial scaffolds with a total size of 90,248 bp. Putative mitochondrial scaffolds were combined to create a single 72,865 bp consensus assembly. The assembly was verified by mapping the long-read PacBio data to the consensus sequence which resulted in contiguous sequence with no reads that exceed the 5' or 3' terminal ends (Fig. S1). No circular mapping was obtained for the linear mtDNA, i.e. no reads were found that spanned the terminal ends of the consensus sequence, and terminal deoxynucleotidyl transferase (TdT) tailing followed by PCR amplification further verified the linear ends of the molecule. The ends of the linear molecule contain two identical 3,529 bp repeats in inverted orientation, which we refer to as Terminal Inverted Repeats (TIR) (Fig. 1, Fig. S2).

Annotation of the *S. endobioticum* mitochondrial genome recovered the mitochondrial genes typically found in fungal species: *atp6*, *atp8*, *atp9*, *cob*, *cox1*, *cox2*, *cox3*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *rns* and *rnl*. In addition, a reduced set of five tRNA recognizing methionine, tryptophan, tyrosine, proline and leucine was found. Four homing endonuclease genes (HEGs) were identified encoded by introns of protein coding genes: three HEGs residing in *cox1* introns carry the LAGLIDADG_1 Pfam domain (PF00961), whereas the HEG residing in the *cob* intron carries the LAGLIDADG_2 Pfam domain (PF03161). No presence/absence variations were found for the intron and HEGs content in other *S. endobioticum* isolates. Interestingly a gene coding for DNA polymerase type B (*dpoB*) containing the DNA_pol_B_2 Pfam domain (PF03175), which are also found in linear mitochondrial plasmids, was identified on the *S. endobioticum* mitogenome. The *dpoB* gene is partially repeated in the A+T rich region (Fig. S3).

Intergenic regions are found to have an unusual high G+C content (57.3%), with several intergenic spacers having G+C contents close to 70% (e.g. *nad4-nad2*: 68.5%, *nad2-cob*: 67.6%, and *cox3-nad4*: 65.5%). In contrast, the intergenic regions in the mitochondrial genome of a closely related species, *S. microbalum*, consist of 28.6% G+C, with the highest G+C percentage found in the *nad6-nad4L* intergenic spacer (34.2%). *S. endobioticum* has the highest intergenic G+C content in chytrid mitogenomes reported to date, which is only almost matched by the obligate biotrophic *S. taraxaci* (51.4%). For Chytridiomycota other than *S. endobioticum* and *S. taraxaci*, a mean intergenic G+C content of 34.3% (standard deviation 9.7%) was observed.

Chapter 3

Of the five additional chytrid species sequenced in this study, three resulted in a single circular mtDNA scaffold containing all 14 mtDNA genes, i.e. *C. confervae* (225.6 kb), *S. palustris* (126.5 kb), and *S. microbalum* (23.8 kb). The assembly of *P. hirtus* resulted in three mtDNA scaffolds with a total size of 298.6 kb, in which all mtDNA genes were detected except for *nad6*. With three scaffolds we could not determine if the mitochondrial genome is circular. The mtDNA assembly of the obligate biotrophic species *S. taraxaci* extremely

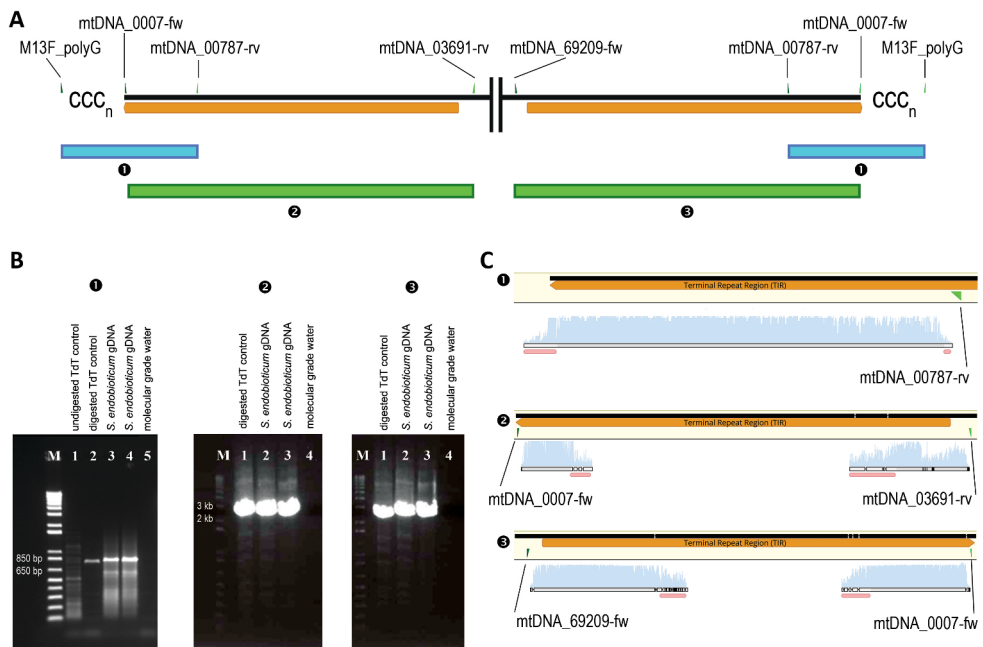


Figure 2. Verification of linearity of the *S. endobioticum* mtDNA. **A** Design of the verification experiments using a graphical representation of the 5' and 3' TIRs (orange) and forward and reverse primer sites (green and light green). Three PCR reactions, performed after TdT tailing, are displayed: **1** M13F_polyG/mtDNA_00787-rv to verify linear conformation of the mtDNA. This reaction takes place at both telomeric ends since the TIRs are inverse orientated. **2** mtDNA_0007-fw/mtDNA_03691-rv to verify presence of the TIR with its specific flanking sequence on the 5' end of the mtDNA. **3** mtDNA_0007-fw/mtDNA_69209-fw to verify presence of the TIR with its specific flanking sequence on the 3' end of the mtDNA. **B** Gel images corresponding with reactions described under **A**. The 1kb-plus size marker (M) was used for amplicon size estimation. **C** Sanger sequence data mapped to the *S. endobioticum* mtDNA corresponding with the reactions described under **A**. The mtDNA is used as reference, and bases similar to the reference are shown in grey, and differences to the reference are highlighted (black). Phred scores for the individual peaks are shown as blue bars, and low quality data (Phred <30) is annotated in pink. For reaction 1, the first 800 bases of the 5' end TIR are shown, whereas for reactions 2 and 3 the full TIR including the specific flanking sequences are shown (~4 kb).

fragmented with 14 mtDNA scaffolds with a total size of 39.2 kb coding for all mtDNA genes. Annotation of the mtDNA scaffold of *B. dendrobatidis* JEL423 (DS022322; single scaffold, 175.3 Kb) resulted in the detection of all mtDNA genes except for *nad2* (table S1).

Verification of the *S. endobioticum* linear mtDNA

PCR amplification of the poly-C tailed terminal end to a primer site in the TIR (mtDNA_00787-rv), performed to verify the linearity of the mitogenome, resulted in an amplicon length of expected size (~850 bp). Sanger sequencing of the amplicon and alignment to the mitochondrial genome showed a perfect fit to the TIR region.

Long-range PCR amplification from the telomeric ends, over the TIR sequence, to a 5' or 3' specific primer site (mtDNA_03691-rv and mtDNA_69209-fw), performed to verify the presence of identical TIRs and 5' and 3' specific sequences, resulted in amplicons of the expected length (~3,680 bp, and ~3,710 bp respectively), and Sanger sequence data of the amplicons verified the presence of the repeat on both telomeric ends of the linear molecule (Fig. 2). A linear concatemeric structure of the *S. endobioticum* mitochondrion can be ruled out as no sequence data could be mapped beyond the 5' and 3' ends, and no amplicons were obtained using a primer pointing outwards in the TIR (mtDNA_00787-rv). This primer anneals to the TIR and would act as both forward and reverse primer in a PCR reaction bridging the ends of both TIRs should a concatemeric structure exist.

Bayesian inference of phylogeny

A 16,329 nt alignment (including gaps) was constructed from concatenating individual mitochondrial genes from 44 isolates. The Bayesian phylogeny showed all nodes with a posterior probability of 1.00 (Fig. 3).

Mitochondrial haplotypes and intraspecies variation

Reference assembly of 29 additional *S. endobioticum* isolates to the MB42 mtDNA resulted in 72,941 bp to 72,775 bp mitogenomes with 50 to 163,200x mean coverage. Alignment of all 30 *S. endobioticum* mitogenomes resulted in a 73,030 bp alignment (including gaps). Sequence homology between isolates was high at $\geq 99.21\%$ and sequence variation was found mostly in the intergenic regions (96.8% of the variation). Two Polish isolates, 03WS and 04WS, had identical mitochondrial genome sequences. Extraction of variable sites resulted in a 347 bp alignment (including gaps) that contained 122 parsimony-informative

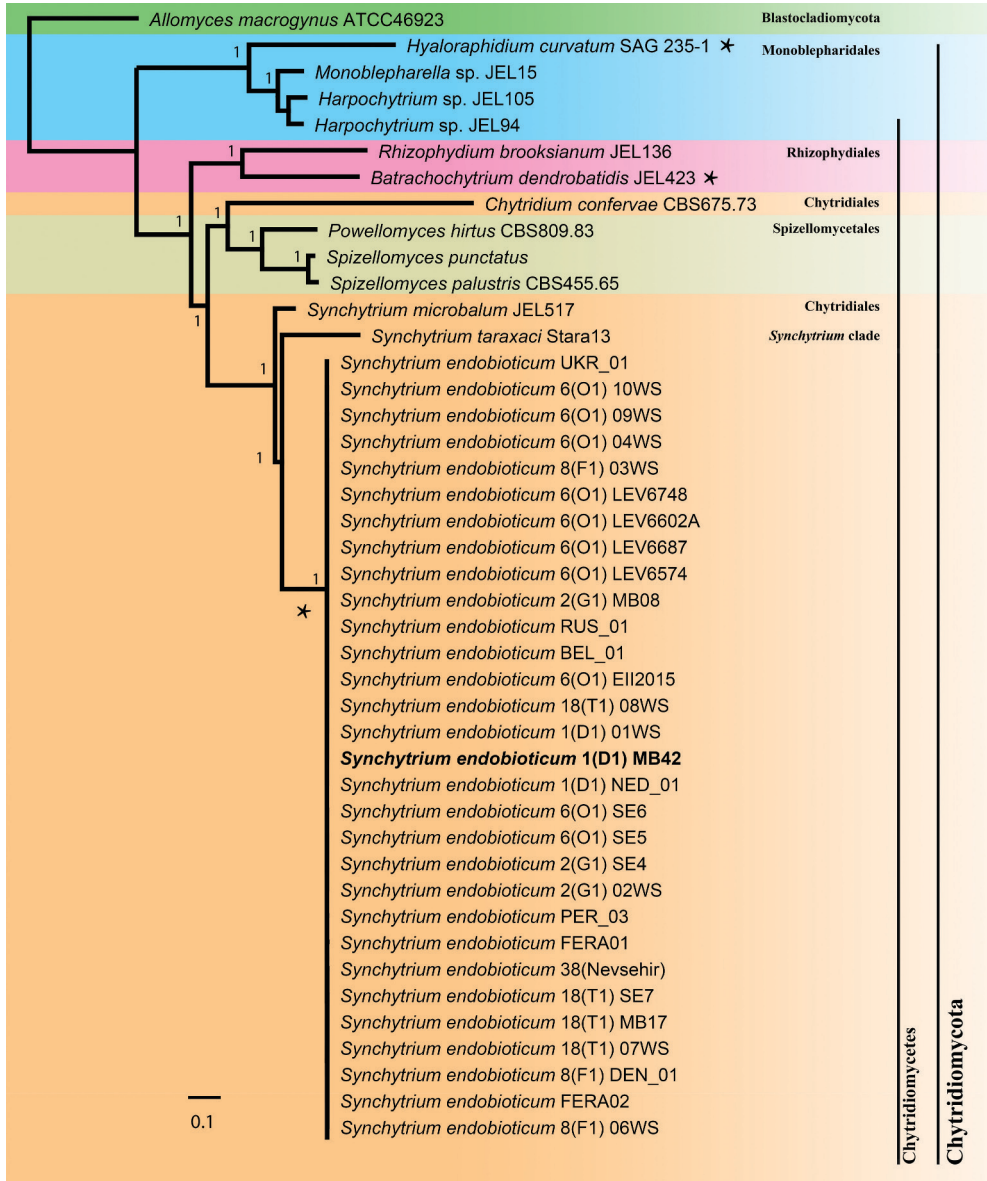


Figure 3. Inference of phylogeny. Bayesian tree (GTR model, G+I distributed sites) based on a concatenated alignment of *atp6*, *atp8*, *atp9*, *cob*, *cox1*, *cox2*, *cox3*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, and *nad6* of 43 chytrid isolates with *A. macrogynus* (Blastocladiomycota) serving as outgroup. Bayesian posterior probabilities are displayed at branch nodes. Highlighted in bold is the *S. endobioticum* mtDNA reference isolate MB42. Species with linear mtDNA genomes are indicated with an asterisk “*”.

sites, which were used to determine mitochondrial haplotypes. Four main groups (i.e. mitochondrial lineages) are distinguished by >20 mutations (Fig. 4). Clustering based on polymorphic sites in coding sequences of mtDNA genes showed a similar clustering with lower resolution (Fig. S5).

Composition of *S. endobioticum* populations

Read mapping to the mitochondrial genome showed unexpected variations of SNP frequency for parsimony-informative sites for the majority of the sequenced *S. endobioticum* isolates. Next to the dominant type, low levels of the polymorphic base were observed. For instance, the polymorphism at position 5,803 (reference: A, alternative: C) is fixed (>98 %) in nine isolates from group 1, but this SNP is also present in variable frequencies (12 - 60%) in seven isolates from groups 1, 2 and 3. Similar differences were found for all 122 parsimony-informative sites (Table S2). In ten isolates from group 1 originating from Canada, Poland and the Netherlands, no signatures for multiple haplotypes were found (Fig. 4). An exception is found in a 40 bp hypervariable intergenic (*atp9-cox1*) region containing 13 polymorphic sites which was found to be present in all *S. endobioticum* isolates.

Table 2. Virulence profiles of *S. endobioticum* isolates 1(D1)_01WS and E/II/2015 using differential cultivars described in EPPO PM7/28(1). Susceptible interactions (wart formation) are indicated with “S”, intermediated reactions (sprouts partially proliferated) are indicated with “±”, and resistant interactions are indicated with “-”.

Differential cultivar	1(D1)	E/II/2015
Deodara	S	S
Tomensa	S	S
Eersteling	S	S
Producent	-	S
Combi	-	S
Saphir	-	-
Delcora	-	±
Miriam	-	-
Karolin	-	-
Ulme	-	-

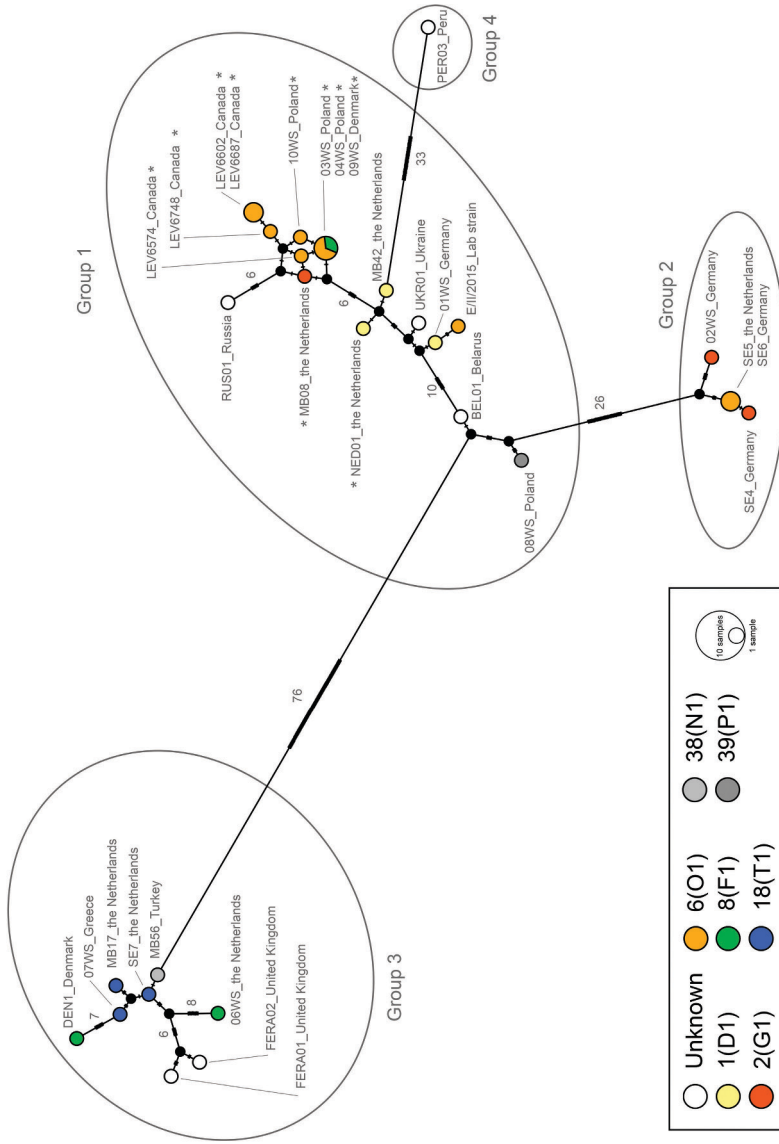


Figure 4. Median joining haplotype network. The network is based on 141 polymorphic sites on the mitochondrial genome of *S. endobioticum*, of which 122 were parsimony-informative. Nodes in the network are coloured based on pathotype identity. Colours are used for pathotypes of major importance in Europe and Canada, whereas a greyscale is used for pathotypes of lesser importance [9]. Black nodes represent hypothetical ancestors. Marks on the branches indicate the number of mutations, and the numbers are shown on branches with > 5 mutations. No signatures for multiple mitochondrial haplotypes were detected for isolates with an asterisk “*”.

Pathotype formation

The dynamic process of pathotype formation is clearly illustrated by pathotype 6(O1) isolate E/II/2015. This isolate was obtained after two multiplications of pathotype 1(D1) isolate O1WS on the partial resistant potato cultivar Erika in 2015. Before multiplication on this partial resistant cultivar a clear pathotype 1(D1) phenotype was obtained using the EPPO differential set, whereas after multiplication on the partial resistant potato cultivar, a clear pathotype 6(O1) phenotype was obtained (table 2). The resulting isolate E/II/2015 is virulent on cultivars Producent and Combi. Repeating the experiment, i.e. multiplying isolate O1WS on cultivar Erika, resulted in the same increased virulence and the same pathotype 6(O1) pathotype profile was obtained. After multiplication, the within isolate variation of E/II/2015 is higher compared to the O1WS isolate.

Discussion

We assembled and annotated the mitogenomes of 30 *S. endobioticum* isolates and five additional chytrid species, and compared these to publically available Chytridiomycota mitogenomes. Using the mitogenome sequences, we determined taxonomical relationships of *S. endobioticum* to 14 other chytrid species, and to performed network analyses to identify potential linkages between *S. endobioticum* genotypes, geographical origins and pathotypes.

The linear mitogenome of *S. endobioticum*

Mitochondrial genomes are typically assumed to have a circular monomeric conformation. However, after the description of a linear mitochondrial genome in *Tetrahymena pyriformis* [22], other linear mitochondrial genomes were reported in fungi, oomycetes, ciliates, Chlorophycean green algae and cnidarians, [23-25]. From our analyses, the *S. endobioticum* mitogenome was found to be linear, encoding a *dpoB* homolog and having two identical TIRs.

In contrast, the mitochondrial genome of the sister species *S. microbalum* is circular, lacks the *dpoB* homolog and comprises only a third of the size of the *S. endobioticum* mitogenome, suggesting recent linearization and expansion in the latter species. Integration of linear mitochondrial plasmids have been hypothesized to be cause of linearization of mitogenomes in medusozoan cnidarians (24) and yeasts (25), and the *dpoB* gene identified in the *S. endobioticum* genome is believed to be reminiscent of such event. Among chytrid

Chapter 3

fungi, a linear mitochondrial genome is described for *H. curvatum* [26] and *B. dendrobatidis* [27], of which the latter contains a *dpoB* ortholog, and consists of three linear segments, each having inverted repeats at the termini.

As was found in other chytrid species [28, 29], *S. endobioticum* contains a reduced set of tRNAs. The tRNA content of the mitochondrial genome determines the genetic code that applies for a given mitogenome. In particular, the presence of *trnL(CTA)*, translating TAG to Leucine instead of introducing a stop codon, indicates that the Chlorophycean Mitochondrial Code (tbl16) applies to the mitogenomes of the species assembled and annotated in this study. The organisation and orientation of the mitochondrial genes, tRNAs and rRNAs are conserved between the linear mtDNA of *S. endobioticum* and the circular mitochondrial genome of *S. microbalum* (Fig. S4). Otherwise, conservation was observed neither in the organization nor orientation of mitochondrial genes in the Chytridiomycota, underlining the long evolutionary distance among these species.

Bayesian inference of phylogeny

Clustering of species in their respective orders is identical compared to the nuclear rDNA based phylogeny (18S+5.8S+28S subunits) as described by James *et al.* [30], supporting the polyphyly in the order Chytridiales and monophyly in the genus *Synchytrium* as described by Smith *et al.* [31]. This further underlines the use and successful application of whole mitochondrial genomes for phylogeny providing the required resolution and robustness in such a wide and scarcely populated evolutionary branch.

Mitochondrial haplotypes and intraspecies variation

Mitochondrial genotypes did not show direct association with pathotypes or origin (table S3). Similar observations were reported by Gagnon *et al.* [11] using nuclear based micro-satellite markers. Isolates of pathotypes 2(G1) and 6(O1) are found both in groups 1 and 2. Also, isolates originating from the Netherlands (e.g. 1(D1)_MB42, 2(G1)_MB08, 6(O1)_SE5, and 18(T1)_MB17), are represented in groups 1, 2 and 3. European pathotype 6(O1) isolates are found in both groups 1 and 2, and pathotype 8(F1) isolates are present in clusters 1 and 3. Nevertheless, pathotype 18(T1) isolates are found exclusively in group 3 even though they originate from different countries. Similarly, pathotype 6(O1) isolates originating from Prince Edward Island, Canada are found only in group 1.

The capacity of *S. endobioticum* for natural dispersion is limited, and the pest is mainly spread as a result of human crop exchange and agricultural activities [2, 32]. Our observation that *S. endobioticum* isolates with identical multi-gene mitochondrial haplotypes can have a global distribution corroborates with distribution as a result of human activities. If regional or worldwide distribution would pre-date agricultural development, a much larger variation in mitochondrial haplotypes between isolates from distant locations would be expected. Based on the mitochondrial lineages identified, we conclude that *S. endobioticum* has been introduced at least three times in Europe, and that from those introductions the disease has spread in Europe and to other continents. Isolates from pathotype 18(T1), identified first in Germany in 1978 [33], cluster in mtDNA group 3, suggesting that the emergence of this pathotype was the result of a new introduction opposed to the gain of increased virulence of isolates already present in Europe. The fourth group is represented by a single isolate originating from Peru and fits well with the hypothesis that more genetic diversity can be expected from the Andean region, the presumed center of origin of this pathogen.

Pathotype formation in *S. endobioticum* is a dynamic and ongoing process. This is supported by the occurrence of the same pathotypes in different mitochondrial lineages, e.g. pathotypes 2(G1) and 6(O1) being found in both mitochondrial groups 1 and 2. Our data suggests that the emergence of those pathotypes have occurred independently in a different genetic background. Different SNP frequencies in read mappings demonstrate the presence of more than one mitochondrial haplotype in the majority of isolates. This was not observed in the other chytrid species sequenced in this study, with the exception of *S. taraxaci*; a closely related species with a similar obligate biotrophic pathogenic life style on its host dandelion. The number of SNPs was not high enough for the size of the fragments to assemble the individual haplotypes for isolates MB42 and LEV6574 using long read PacBio sequences and the recently published Canu assembler [34]. The composition of the *S. endobioticum* communities of different genotypes seem to be conserved. Isolates from group 1 typically portray low levels of diversity, whereas isolates from group 3 have high levels of diversity (Fig. S6). Ten isolates clustering together in group 1 are strikingly scarce in diversity, apart from the 40 bp *atp9-cox1* intergenic region which found to be variable in all isolates. We suspect that this region could represent the origin of replication which may involve an RNA intermediate as was described for linear chromosomal fragments [35].

Chapter 3

Inoculation experiments show that the use of semi-resistant potato cultivars triggers a rapid shift (increase) in virulence, which was associated with a shift in the mitochondrial haplotypes found in the isolate. In fact an increase in diversity was observed. Most likely this variation was already present, but below the detection limit of the sequencing depth. We postulate that this increased virulence was the result of a shift in the heterogeneous 1(D1) population of *S. endobioticum* due to the selection of virulent genotypes already present in the isolate. Functional variation could be present on the nuclear genome as we found evidence for genetic heterogeneity not only on the mitochondrial genome but also on some nuclear scaffolds in all sequenced isolates (data not shown). The alternative hypothesis that many new mutations could be induced during the two multiplications seems highly unlikely. The possible rapid increase in virulence due to selection of the population underlines that caution is needed when adopting the deployment of resistance (*R*) genes in a control strategy, in particular when they provide partial or QTL-based resistance.

Conclusions

The mitochondrial genome of *S. endobioticum* and other Chytridiomycota provides insight into the evolution of mitochondria. Of special interest are the linearization of what is often assumed to be a circular genome which happened multiple times during the evolution of Chytridiomycota, the extremely variable size range (19,5 to 298,5 Kb), and the remarkable high GC content for the intergenic regions (close to 70%). The mitochondrial genome also provides insight in the more recent history of this plant pathogen. From this study we conclude there were at least three independent introductions in Europe, most likely as the result of human activities. With one of those introductions, pathotype 18(T1) was most likely introduced into Europe and so far this pathotype seems to be confined to Europe. A fourth mtDNA lineage obtained from a Peruvian isolate represents a genotype not found in other *S. endobioticum* isolates sequenced so far. Mitochondrial data shows that *S. endobioticum* isolates should not be considered as a single genotype, but rather populations of different genotypes. The observed population diversity seems to be consistent in four different mitochondrial lineages identified in this study. Also, we conclude that pathotypes 2(G1) and 6(O1) have emerged at least twice independently, most likely by selection of virulent genotypes from diverse population as was observed for a pathotype 6(O1) derivative from a pathotype 1(D1) isolate.

Our data indicate that characterization by pathotyping alone is insufficient for characterizing isolates, as underlying diversity may remain unnoticed. To keep track of genetic drift and selection of individuals from such populations we recommend sequencing of isolates before and after each multiplication, particularly when (partial) resistant potato cultivars are used. Obtaining genetically pure cultures may not be possible since *S. endobioticum* is an obligate biotroph. This situation may be similar for other plant pathogens in this genus as illustrated by the fact that a heterogeneous genotypic profile in the mtDNA of *S. taraxaci* was observed. To our knowledge *S. endobioticum* represents a unique case where although the number of (new) outbreaks of the disease is low, the genetic diversity is high.

Methods

Fungal material

A total of 30 *S. endobioticum* isolates were included in this study covering eleven different pathotypes and six isolates with unknown pathotype identity. Material was collected from countries in Europe, North America, South America and Asia (table 1). Pathotype identification for the Canadian isolates was performed according to the Glynne-Lemmerzahl method [36]. The samples were provided either as wart material, resting spores or DNA extracts. When provided with wart material, resting spores were extracted following the procedure described by Bonants *et al.* [37]. For the Canadian isolates, resting spores collected in a 37 µm sieve were collected and treated with a 1/10 dilution of viscozyme (Sigma-Aldrich) in 100 mM phosphate buffered saline at pH 5.0 at 30 °C for 16 hours. Spores were layered onto a 50% solution of sucrose and centrifuging at 2,000 x g for 20 minutes. Spores were collected in a sieve, washed, and further purified by pelleting through a 90% solution of Percoll (Sigma-Aldrich) at 1,000 x g for 5 min. Cultures of *Chytridium confervae* (CBS 675.73), *Powellomyces hirtus* (CBS 809.83), *Spizellomyces palustris* (= *Phlyctochytrium palustre*) (CBS 455.65) and *Synchytrium microbalum* (JEL517) were maintained on ARCH medium and ARCH broth (Broth: peptone 2 g (Oxoid), malt extract 3 g (Oxoid), glucose 5 g/L, pH 6, Medium contained 8 g/L agar (Oxoid) in addition) at 21 °C. Dandelion leaves containing *Synchytrium taraxaci* spores were collected from a meadow in Bennekom the Netherlands (GPS coordinates: 51.997314, 5.662961) in April 2015. *S. taraxaci* spores were extracted by gentle disruption of epidermal leaf tissue using a pipette tip. Extracted spores were washed with tap water and stored at -20 °C until DNA extraction.

Chapter 3

Table 1. *S. endobioticum* isolates and other fungal materials used in this study.

Pathotype, based on EPPO PM7/28(1)	Additional cultivars ^a	Isolate	Origin	mtDNA accession
1(D1)	n.a.	MB42	Langenboom, the Netherlands	ERZ668671
1(D1)	n.a.	1/2007/D1 (01WS)	Germany	ERZ668651
1(D1)	n.a.	NED01	Brabant, the Netherlands	ERZ668673
2(G1)	n.a.	4/2005/G1 (02WS)	Germany	ERZ668652
2(G1)	n.a.	MB08	Mussel, the Netherlands	ERZ668669
2(G1)	n.a.	BBA 2(G1) 09-04 (SE4)	Germany	ERZ668676
6(O1)	3(M1)	PL28/2007/2 (04WS)	Poland	ERZ668654
6(O1)	P40	DK17/2015 (09WS)	Denmark	ERZ668658
6(O1)	P41	PL2/2015 (10WS)	Poland	ERZ668659
6(O1)	n.a.	E/II/2015	Isolate obtained after two multiplications of isolate 01WS on cultivar Erika	ERZ668662
6(O1)	n.a.	LEV6574	Prince Edward Island, Canada	ERZ668665
6(O1)	n.a.	LEV6602	Prince Edward Island, Canada	ERZ668666
6(O1)	n.a.	LEV6687	Prince Edward Island, Canada	ERZ668667
6(O1)	n.a.	LEV6748	Prince Edward Island, Canada	ERZ668668
6(O1)	n.a.	HLB 6(O1) 02-06 (SE5)	the Netherlands	ERZ668677
6(O1)	n.a.	BBA 6(O1) 05_8.3 (SE6)	Germany	ERZ668678
8(F1)	n.a.	DEN01	Jylland, Denmark	ERZ668661
8(F1)	n.a.	3/2005/F1 (06WS)	The Netherlands	ERZ668655
8(F1)	2(Ch1)	2/2005/Ch1 (03WS)	Poland	ERZ668653
18(T1)	n.a.	GR2/2015 (07WS)	Greece	ERZ668656
18(T1)	n.a.	MB17	Borgercompagnie, the Netherlands	ERZ668670
18(T1)	n.a.	HLB P18(T1) -02-06 (SE7)	Borgercompagnie, the Netherlands	ERZ668679
38(N1)	n.a.	MB56	Nevsehir, Turkey	ERZ668672
39(P1)	n.a.	PL69/2009 (08WS)	Piekielnik, Poland	ERZ668657
unknown	n.a.	BEL01	Belarus	ERZ668660
unknown	n.a.	FERA01	United Kingdom	ERZ668663
unknown	n.a.	FERA02	United Kingdom	ERZ668664
unknown	n.a.	PER03	Aco Paucartambo, Pasco, Peru	ERZ668674
unknown	n.a.	RUS01	St. Petersburg, Russian Federation	ERZ668675
unknown	n.a.	UKR01	Ukraine	ERZ668680

a. In this paper, pathotype identities are used based on the potato cultivar differential set as presented in EPPO standard PM7/28(1). In some cases, additional cultivars were used resulting in a further differentiation of pathotype identity (n.a. when not applicable).

Table 1 (continued). Fungal material used in this study.

Species	Strain	Origin	ITS accession	mtDNA accession(s)
<i>Chytridium confervae</i>	CBS 675.73	Canada	MH660417	ERZ681044
<i>Powellomyces hirtus</i>	CBS 809.83	the Netherlands	MH660418	ERZ681050
<i>Spizellomyces palustris</i>	CBS 455.65	Germany	MH660420	ERZ681046
<i>Synchytrium microbalum</i>	JEL517	Hancock Co., Maine, USA	MH660419	ERZ681045
<i>Synchytrium taraxaci</i>	Stara13	Bennekom, the Netherlands	MH660421	ERZ681051

Multiplication of pathotype 1(D1) isolate 01WS on potato cultivar Erika

Eye pieces of cultivar Erika were inoculated with winter sporangia of pathotype 1(D1) isolate 01WS. After two months of incubation, matured winter sporangia from small warts were removed and used for a second round of inoculation on cultivar Erika. After each multiplication the virulence was checked on a differential set of potato cultivars. When wart development on cultivars known to be resistant to pathotype 1(D1) was observed, a full virulence profile was determined on an extended set of potato cultivars. After two multiplications, isolate 01WS produced a pathotype 6(O1) virulence spectrum and the isolate was renamed to E/II/2015. The experiment was repeated at a different moment with similar results.

DNA extraction

A total of 200 μ l suspensions containing $\geq 5,000$ *S. endobioticum* resting spores were homogenized in a Hybaid Ribolyser multiple bead beater (Thermo Electron Corporation, the Netherlands) at 5,000 bpm for 100 s using three stainless steel beads (3.2 mm). For the WS and E/II/2015 isolates, resting spores were homogenized using a combination of vortexing and ultrasound sonication. Subsequently the Ultra Clean Soil DNA kit (MoBio) was used according to manufacturer's instructions to extract genomic DNA. For the Canadian isolates, DNA was extracted from resting spores as previously described (14). Genomic DNA of *S. microbalum* was extracted from fungal cells grown in 8 mL ARCH broth and 50 μ l spore suspensions containing *S. taraxaci* spores using the Wizard Magnetic DNA Purification System for Food (Promega) following manufacturer's instructions. When sufficient starting material was available multiple DNA extracts were generated per isolate. The nuclear ITS region was amplified and sequenced to verify the species identity using primers ITS4 [38] and Chy18S-269A [31] as a check before Next Generation Sequencing (NGS).

Quantification of obligate biotrophs

S. endobioticum DNA was quantified using a *S. endobioticum* specific real-time PCR [39]. For the quantification of *S. taraxaci*, a species specific real-time PCR was designed based on publicly available ITS1-5.8S-ITS2 sequences covering 20 *Synchytrium* species (Fig. S7). DNA extracts with exponential amplification curves and Cq values below 30 were selected for NGS.

Next Generation Sequencing

NGS data was generated using one or more of three sequencing platforms: HiSeq 2500 (Illumina), MiSeq (Illumina), PacBio RSII (Pacific Biosciences) (table S4).

Assembly and annotation of the *S. endobioticum* mitochondria

A consensus mitochondrial sequence of *S. endobioticum* isolate MB42 pathotype 1(D1) was created using the results of three independent assemblies with two different datasets. In the first assembly, Illumina HiSeq reads were mapped to the *S. endobioticum* MB42 reference genome (Van de Vossenbergh & Warris *et al.*, unpublished) in CLC genomics workbench v8.0.2. Scaffolds coverage >50 times above the expected coverage for single copy nuclear regions were selected and annotated using Mfannot with the Chlorophycean Mitochondrial Code (tbl 16) [40]. Contigs containing an Mfannot annotation and/or those with an e-value of < 1e-20 when compared by blastn to other publically available chytrid mitochondrial genome sequences, were selected. In the second assembly, the GRAbB assembly tool [41] was used to reconstruct the *S. endobioticum* mitochondrial genome using the putative *S. endobioticum* scaffolds identified in the first assembly, as bait for the raw genomic library. The reads collected by the final iteration of the GRAbB run were used to perform *de novo* assemblies using three different assemblers using SPAdes v3.6 [42], Velvet v1.2.10 [43] and Edena v3.131028 [44].

Table 3. Amplification and sequencing primers used for verification of linearity mtDNA genome.

Primer name	Primer sequence (5' - 3' direction)
M13F_polyG	TGTAAAACGACGGCCAGTGGGGGGGGGGGGGGGGGG
M13F	TGTAAAACGACGGCCAGT
mtDNA_0007-fw	GTTTTCTTAGGCCATCCT
mtDNA_00787-rv	CAGTTTTCTTAGGGTCCA
mtDNA_03691-rv	AATAGACTATCAGAACCACCAAG
mtDNA_69209-fw	CTCAAACACAGAAATTAACAACG

Overlaps between contigs produced by the assemblers were manually identified and merged using the helper tools distributed with the GRAB assembler until all regions were included in a single contig. Raw reads were mapped back to the final sequence and errors were corrected with Pilon v1.19 [45] iteratively until no further errors were identified. In the third assembly, the mitochondrial genome of *S. endobioticum* MB42 was assembled with Celera Assembler version 8.2 [46] using the PacBio corrected reads [47]. Using the largest scaffold as a backbone sequence, all putative mitochondrial scaffolds from the three independent assemblies were assembled in Geneious R10 [48] using default parameters. Conflicts were resolved using a majority rule. In addition, the Illumina HiSeq and PacBio reads from *S. endobioticum* MB42 were mapped to this consensus sequence to further identify and resolve conflicts using the quality based majority rule resulting in a final consensus *S. endobioticum* mitochondrial genome. This final consensus was verified using the PacBio reads mapping to the consensus which were assembled using the Canu assembler [34]. The final consensus sequence was then annotated using Mfannot. Putative mitochondrial genes, tRNAs and rRNAs were manually curated. Amino acid sequences from predicted gene models were compared to those in InterProScan [49] and the NCBI nr database by blastp and those with an e-value of $<1e-20$ with a known predicted function were kept as valid and those that did not meet these criteria were filtered out.

Molecular verification of mitochondrial genome linearity

A TdT tailing method based on [50, 51] was used as a method to verify that the linear conformation of the *S. endobioticum* mitochondrial genome. Positive and negative controls (i.e. synthetic oligo resembling the mitochondrial genome sequence in a linear and circular form respectively, see Fig. S8) were included to monitor the effectiveness of the TdT tailing. Sample DNA (30 to 45 ng/ μ L) was denatured at 95 °C for 5 min, placed on ice and then immediately used in a 25 μ L of 3'-TdT reaction consisting of 1x TdT buffer (TaKaRa), 0.01% BSA, 300 μ mol of dCTP (TaKaRa), 7 U of TdT enzyme (TaKaRa), and 1 μ L of sample DNA. TdT tailing was performed at 37 °C for 30 min, followed by the inactivation of the TdT enzyme at 65 °C for 10 min. Linearity of the mtDNA was tested by PCR amplification using the primers listed in table 3. Amplification from the poly-C 3' ends of the mtDNA into the terminal inverted repeat (TIR) was performed with primers M13F_polyG and mtDNA_00787-rv using a reaction mix containing 1x GoTaq flexi buffer (Promega), 1.5 mM MgCl₂, 200 μ M dNTP mix, 300 nM of each primer, 1U Taq polymerase (Promega), and 1 μ L poly-C tailed genomic DNA. Cycling conditions were as follows: 2 min at 95 °C, 35x (30 sec at 95 °C, 30 sec 60 °C, 1

Chapter 3

min 72 °C), 5 min 72 °C. The presence of TIRs on both the 5' and 3' end of the linear mtDNA was verified by Long-Range PCR from the telomeric ends of the mtDNA over the TIR to both site specific ends with primer pairs mtDNA_0007-fw/ mtDNA_03691-rv for the 5' end, and mtDNA_0007-fw /mtDNA_69209-fw for the 3' end. Reaction mixes for the latter two primer combinations consisted of 1x GoTaq Long Master Mix (Promega), 300 nM of each primer, and 1 µL poly-C tailed genomic DNA. Cycling conditions were as follows: 2 min at 95 °C, 35x (30 sec at 95 °C, 30 sec 60 °C, 3.5 min 72 °C), 5 min 72 °C. After electrophoresis, PCR products were excised from the gel, purified with the Wizard SV gel and PCR clean-up system (Promega) and submitted for sequencing using the generic M13F sequencing primer, and specific mtDNA amplification primers in separate reactions. The resulting sequence data were mapped to the assembled and annotated *S. endobioticum* consensus mitochondrial genome in Geneious.

Assembly and annotation mitochondrial genome of other species

Illumina NGS data of other chytrid species: *C. confervae*, *P. hirtus*, *S. palustris*, *S. microbalum* and *S. taraxaci* were used for *de novo* assembly using CLC genomics workbench (word size: automatic, bubble size: automatic, scaffolding: on) and scaffolds were updated using a read mapping approach (length fraction: 0.5, similarity fraction: 0.8) (Van de Vossenbergh & Warris *et al.*, unpublished). Scaffolds with relative high coverage were selected and annotated using Mfannot (tbl 16) as described above for *S. endobioticum* MB42.

Bayesian inference of phylogeny

NGS datasets from other isolates of *S. endobioticum* were mapped to the annotated *S. endobioticum* MB42 mitochondrial genome in CLC genomics workbench (length fraction: 0.8, similarity fraction: 0.9, non-specific match: map random, resolve conflicts: quality based) and regions with $\geq 5x$ coverage were extracted. Annotations from the *S. endobioticum* MB42 reference mitochondrial genome were transferred to the consensus sequence. Coding sequences of 14 mitochondrial genes were extracted from the consensus sequences and these were aligned with MAFFT v7.308 [52] to coding sequences of the same genes from publically available complete chytrid mitochondrial genomes (table S1): *Allomyces macrogynus* ATCC46923 (NC_001715), *Harpochytrium* sp. JEL94 (AY182005), *Harpochytrium* sp. JEL105 (AY182006), *Hyaloraphidium curvatum* SAG 235-1 (AF402142), *Monoblepharella* sp. JEL15 (NC_004624), *Rhizophyidium brooksianumb* JEL136 (NC_003053), and *Spizellomyces punctatus* (NC_003052). In addition, a mitochondrial scaffold was identified

from the *Batrachochytrium dendrobatidis* JEL423 genome (GCA_000149865.1) by blastn (e-value < 1e-20) to other publically available chytrid mitochondrial genome sequences. *B. dendrobatidis* mitochondrial scaffold DS022322 was annotated using Mfannot (tbl 16). Alignments were concatenated and phylogenies were constructed using Bayesian inference (BI) [53]. Models for nucleotide substitution were obtained using JModelTest 2.1.10 [54] was run on default settings and the best nucleotide substitution model, according to AIC, AICc, BIC and DT, was the GTR model with gamma distribution and estimation of invariable sites (GTR+G+I). BI was performed with the MrBayes v3.2.6 plugin incorporated in Geneious with a random starting tree and four Monte Carlo Markov Chains for 10⁶ generations. Trees were sampled every 200 generations, and the first 10⁵ generations were discarded as burn-in. Remaining trees were combined to generate a 50% majority rule consensus tree with posterior probabilities.

Mitochondrial haplotypes and intraspecies variation

The mitochondrial genomes of 30 *S. endobioticum* isolates were aligned with MAFFT v7.308 to estimate the relationship between *S. endobioticum* isolates. Low coverage terminal sequences (<5x coverage) were trimmed from the alignment and variable sites were extracted with “Show variable sites only” from the online fasta sequence toolbox FaBox [55]. A Median Joining (MJ) network ($\epsilon = 0$, uninformative sites = masked) was calculated with Popart v1.7 [56]. Variation within a isolate for a given polymorphic site was determined for informative sites in the MJ network with the basic variant calling tool, including only paired reads but allowing non-specific matches to include TIRs, in CLC genomics workbench (mincoverage = 5x, mincount = 2, minfrequency = 10%, minbase quality = 20, direction frequency filter = 5%).

Acknowledgements

We thank Ednar G. Wulff (Danish Veterinary and Food Administration, Denmark), Ian Adams (FERA, United Kingdom), Olga Afanasenko (All Russian Research Institute for Plant Protection, Russian Federation) for kindly providing *S. endobioticum* infected potatoes or DNA extracted from warts for Illumina sequencing; Joyce E. Longcore (University of Maine, United States of America) for kindly providing cultures of *S. microbalum* (JEL517) for Illumina sequencing; Henri C. van der Geest and Sven Warris (Wageningen University and Research, the Netherlands) for help with the PacBio assemblies using Celera and Canu, and the Circos

Chapter 3

visualization tool; Wilmer Perez and Soledad Gamboa for collection and DNA extraction of Peruvian samples respectively; Peter J.M. Bonants (Wageningen University and Research, the Netherlands) for acquisition of funds; Richard G.F. Visser and J.H. Vossen (Wageningen University and Research, the Netherlands) for their support and advice on the project.

Funding

The PhD project of B.T.L.H. van de Vossen was funded by Topsector Tuinbouw & Uitgangsmaterialen grant 1406-056 entitled “An integrated genomics and effectomics impulse for potato wart resistance management and breeding” Genome sequencing of the Canadian *Synchytrium endobioticum* isolates was funded by Agriculture and Agri-Food Canada grant J-000985 entitled “Next generation sequencing - genomics and metagenomics of quarantine fungal and bacterial crop pathogens” (www.agr.gc.ca). Research at CIP was undertaken as part of, and funded by, the CGIAR Research Program on Roots, Tubers and Bananas (RTB) and supported by CGIAR Fund Donors.

Author contributions

BV, TL designed the study. CL, DS, GL, JK, JP, MB, MG collected, prepared and pathotyped (where relevant) biological material. MG, KD organized sequence data collection. BB, BV, HN performed bioinformatic analyses. BV and MG verified the linearity of the mitogenome. CL, JK, TL acquired funding for the study. BB, BV, HN, TL drafted the manuscript. All authors contributed to the final version of the manuscript and gave final approval for publication.

Author affiliations

BV, MG, TL: Wageningen University and Research, Droevendaalsesteeg 1, 6708 PB, Wageningen, the Netherlands; BV, GL: Dutch National Plant Protection Organization, Geertjesweg 15, 6706 EA, Wageningen, the Netherlands; BB: Westerdijk Fungal Biodiversity Institute, Uppsalalaan 8, 3584 CT, Utrecht, the Netherlands; HN, KD, AL: Agriculture and Agri-Food Canada, 960 Carling Avenue, Ottawa, Canada; DS: Canadian Food Inspection Agency, 93 Mount Edward Road, Charlottetown, Canada; JP: Plant Breeding and Acclimatization Institute, Radzikow, 05-870 Blonie, Poland; JK: International Potato Centre, Avenida La Molina 1895, Lima, Peru; and MB: Hilbrands Laboratorium BV, Kampsweg 27, 9418 PD Wijster, the Netherlands.

References

1. Smith, I.M., *et al.*, Quarantine Pests for Europe - Data Sheets on quarantine pests for the European Union and for the European and Mediterranean Plant Protection Organization. 2 ed. 1997: CAB International, Wallingford, UK.
2. Obidiegwu, J.E., K. Flath, and C. Gebhardt, Managing potato wart: a review of present research status and future perspective. *Theor Appl Genet*, 2014. 127(4): p. 763-80.
3. Laidlaw, W.M.R., A method for the detection of the resting sporangia of potato wart disease (*Synchytrium endobioticum*) in the soil of old outbreak sites. *Potato Research*, 1985. 28(2): p. 223-232.
4. Przetakiewicz, J., The Viability of Winter Sporangia of *Synchytrium endobioticum* (Schilb.) Perc. from Poland. *American Journal of Potato Research*, 2015. 92(6): p. 704-708.
5. Hampson, M.C., History, biology and control of potato wart disease in Canada. *Canadian Journal of Plant Pathology*, 1993. 15(4): p. 223-244.
6. Federal Select Agent Program, Select Agents and Toxins List. 2018 [cited 2018 25 April]; Available from: www.selectagents.gov/selectagentsandtoxinslist.html.
7. Schilberszky, K., Ein neuer Schorfparasit der Kartoffelknollen. *Ber. Deut. Botan. Ges.*, 1896. 14: p. 36-37.
8. Braun, H.C., Biologische Spezialisierung bei *Synchytrium endobioticum* (Schilb.). *Zeitschrift für Pflanzenkrankheiten (Pflanzenpathologie) und Pflanzenschutz*, 1942. 52(11): p. 481-486.
9. Baayen, R.P., *et al.*, History of potato wart disease in Europe - a proposal for harmonisation in defining pathotypes. *European Journal of Plant Pathology*, 2006. 116(1): p. 21-31.
10. Przetakiewicz, J., First Report of New Pathotype 39(P1) of *Synchytrium endobioticum* Causing Potato Wart Disease in Poland. *Plant Disease*, 2015. 99(2): p. 285-285.
11. Gagnon, M.C., *et al.*, Development of Polymorphic Microsatellite Loci for Potato Wart from Next-Generation Sequence Data. *Phytopathology*, 2016. 106(6): p. 636-44.
12. Friedman, J.R. and J. Nunnari, Mitochondrial form and function. *Nature*, 2014. 505: p. 335.
13. Taanman, J.-W., The mitochondrial genome: structure, transcription, translation and replication. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1999. 1410(2): p. 103-123.
14. Smith, D.R., The past, present and future of mitochondrial genomics: have we sequenced enough mtDNAs? *Brief Funct Genomics*, 2016. 15(1): p. 47-54.
15. Gilbert, M.T.P., *et al.*, Intraspecific phylogenetic analysis of Siberian woolly mammoths using complete mitochondrial genomes. *Proceedings of the National Academy of Sciences*, 2008. 105(24): p. 8327-8332.
16. Ma, C., *et al.*, Mitochondrial genomes reveal the global phylogeography and dispersal routes of the migratory locust. *Mol Ecol*, 2012. 21(17): p. 4344-58.
17. Janssen, T., *et al.*, Mitochondrial coding genome analysis of tropical root-knot nematodes (*Meloidogyne*) supports haplotype based diagnostics and reveals evidence of recent reticulate evolution. *Scientific Reports*, 2016. 6: p. 22591.
18. Simon, C., *et al.*, Evolution, Weighting, and Phylogenetic Utility of Mitochondrial Gene Sequences and a Compilation of Conserved Polymerase Chain Reaction Primers. *Annals of the Entomological Society of America*, 1994. 87(6): p. 651-701.
19. Hebert, P.D.N., *et al.*, Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 2003. 270(1512): p. 313.
20. Karling, J.S., *Synchytrium*. 1 ed. 1964: Academic Press, New York.

Chapter 3

21. Longcore, J.E., D.R. Simmons, and P.M. Letcher, *Synchytrium microbalum* sp. nov. is a saprobic species in a lineage of parasites. *Fungal Biology*, 2016. 120(9): p. 1156-1164.
22. Suyama, Y. and K. Miura, Size and structural variations of mitochondrial DNA. *Proc Natl Acad Sci U S A*, 1968. 60(1): p. 235-42.
23. Nosek, J., et al., Linear mitochondrial genomes: 30 years down the line. *Trends Genet*, 1998. 14(5): p. 184-8.
24. Burger, G., M.W. Gray, and B.F. Lang, Mitochondrial genomes: Anything goes. *Trends in Genetics*, 2003. 19(12): p. 709-716.
25. Lavrov, D.V. and W. Pett, Animal Mitochondrial DNA as We Do Not Know It: mt-Genome Organization and Evolution in Nonbilaterian Lineages. *Genome Biology and Evolution*, 2016. 8(9): p. 2896-2913.
26. Forget, L., et al., *Hyaloraphidium curvatum*: A linear mitochondrial genome, tRNA editing, and an evolutionary link to lower fungi. *Molecular Biology & Evolution*, 2002. 19(3): p. 310-319.
27. O'Hanlon, S.J., et al., Recent Asian origin of chytrid fungi causing global amphibian declines. *Science*, 2018. 360(6389): p. 621.
28. Bullerwell, C.E., L. Forget, and B.F. Lang, Evolution of monoblepharidalean fungi based on complete mitochondrial genome sequences. *Nucleic Acids Research*, 2003. 31(6): p. 1614-1623.
29. Laforest, M., et al., Origin, evolution, and mechanism of 5' tRNA editing in chytridiomycete fungi. *Rna*, 2004. 10(8): p. 1191-9.
30. James, T.Y., et al., A molecular phylogeny of the flagellated fungi (Chytridiomycota) and description of a new phylum (Blastocladiomycota). *Mycologia*, 2006. 98(6): p. 860-71.
31. Smith, D.S., et al., Phylogeny of the genus *Synchytrium* and the development of TaqMan PCR assay for sensitive detection of *Synchytrium endobioticum* in soil. *Phytopathology*, 2014. 104(4): p. 422-32.
32. Hampson, M.C., A qualitative assessment of wind dispersal of resting spores of *Synchytrium endobioticum*, the causal agent of wart disease of potato. *Plant Disease*, 1996. 80(7): p. 779-782.
33. Stachewicz, H., Nachweis eines neuen Biotypen des Kartoffelkrebserreger *Synchytrium endobioticum* (Schilb.) Perc. in der DDR. *Nachrichtenblatt für den Pflanzenschutz in der DDR* 1978. 32: p. 215.
34. Koren, S., et al., Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research*, 2017.
35. Picardeau, M., J.R. Lobry, and B.J. Hinnebusch, Physical mapping of an origin of bidirectional replication at the centre of the *Borrelia burgdorferi* linear chromosome. *Molecular Microbiology*, 1999. 32(2): p. 437-445.
36. OEPP/EPPO, PM 7/28 (1) *Synchytrium endobioticum*. *EPPO Bulletin*, 2004. 34(2): p. 213-218.
37. Bonants, P.J.M., et al., A real-time TaqMan PCR assay to discriminate between pathotype 1 (D1) and non-pathotype 1 (D1) isolates of *Synchytrium endobioticum*. *European Journal of Plant Pathology*, 2015. 143(3): p. 495-506.
38. White, T.J.B., T.; Lee, S.; Taylor, J., Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics, in *PCR Protocols: a guide to methods and applications*, I.M.A.G.D.H.S.J.J.W. T.J., Editor. 1990, Academic Press: New York. p. 315-322.
39. van Gent-Pelzer, M.P.E., M. Krijger, and P.J.M. Bonants, Improved real-time PCR assay for detection of the quarantine potato pathogen, *Synchytrium endobioticum*, in zonal centrifuge extracts from soil and in plants. *European Journal of Plant Pathology*, 2010. 126(1): p. 129-133.
40. Mfannot. [cited 2018 25 April]; Available from: <http://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl>.

41. Brankovics, B., *et al.*, GRABb: Selective Assembly of Genomic Regions, a New Niche for Genomic Research. *PLoS Comput Biol*, 2016. 12(6): p. e1004753.
42. Bankevich, A., *et al.*, SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, 2012. 19(5): p. 455-477.
43. Zerbino, D.R. and E. Birney, Velvet: Algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Research*, 2008. 18(5): p. 821-829.
44. Hernandez, D., *et al.*, *De novo* bacterial genome sequencing: millions of very short reads assembled on a desktop computer. *Genome Res*, 2008. 18(5): p. 802-9.
45. Walker, B.J., *et al.*, Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One*, 2014. 9(11): p. e112963.
46. Koren, S., *et al.*, Reducing assembly complexity of microbial genomes with single-molecule sequencing. *Genome Biol*, 2013. 14(9): p. R101.
47. Berlin, K., *et al.*, Assembling large genomes with single-molecule sequencing and locality-sensitive hashing. *Nat Biotechnol*, 2015. 33(6): p. 623-30.
48. Kearse, M., *et al.*, Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 2012. 28(12): p. 1647-9.
49. Jones, P., *et al.*, InterProScan 5: genome-scale protein function classification. *Bioinformatics*, 2014. 30(9): p. 1236-40.
50. Hikosaka, K., *et al.*, Novel type of linear mitochondrial genomes with dual flip-flop inversion system in apicomplexan parasites, *Babesia microti* and *Babesia rodhaini*. *BMC Genomics*, 2012. 13(622).
51. Ogedengbe, M.E., *et al.*, A linear mitochondrial genome of *Cyclospora cayatanensis* (Eimeriidae, Eucoccidiorida, Coccidiasina, Apicomplexa) suggests the ancestral start position within mitochondrial genomes of eimeriid coccidia. *International Journal for Parasitology*, 2015. 45(6): p. 361-365.
52. Katoh, K. and D.M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*, 2013. 30(4): p. 772-80.
53. Huelsenbeck, J.P. and F. Ronquist, MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, 2001. 17(8): p. 754-755.
54. Darriba, D., *et al.*, jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods*, 2012. 9(8): p. 772.
55. FaBox (1.41) - an online fasta sequence toolbox. 2013 [cited 2018 25 April]; Available from: <http://users-birc.au.dk/biopv/php/fabox/>.
56. Bandelt, H.J., P. Forster, and A. Rohl, Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol*, 1999. 16(1): p. 37-48.

Supplementary material

Supplementary figures

- Fig. S1 PacBio read mapping to the *S. endobioticum* mtDNA
- Fig. S2 Dotplot internal repeat structure TIRs
- Fig. S3 Dotplot repeat in AT-rich region containing the *dpoB* gene
- Fig. S4 Comparison of the mitochondrial genomes of *S. endobioticum* MB42 and *S. microbalum* JEL517
- Fig. S5 Median Joining haplotype network based on polymorphic sites in seven mitochondrial genes
- Fig. S6 Within-isolate diversity in haplotype network based on polymorphic sites of the *S. endobioticum* mtDNA
- Fig. S7 *Synchytrium taraxaci* Taqman assay
- Fig. S8 TdT-tailing control

Selected supplementary tables

- Table S1 Mitogenomes and mitochondrial genes included in the Bayesian inference (BI) of phylogeny
- Table S3 Mitochondrial haplotype classification of *S. endobioticum* isolates
- Table S4 Fungal materials and NexGen sequence information

Supplementary table S2 was not included in this thesis on account of its size. This file is available online.

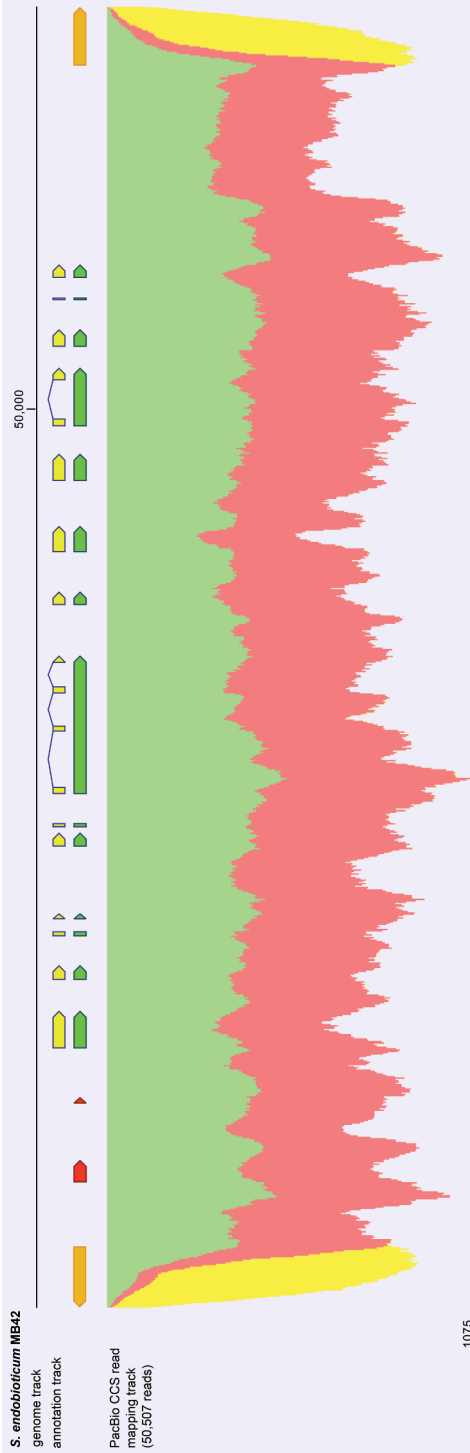


Figure S1. PacBio read mapping to the *S. endobioticum* mtDNA. Mapping of 50,507 PacBio circular consensus reads to the 72,865 bp linear *S. endobioticum* mtDNA genome. Annotations for mitochondrial genes (green), CDS (yellow), and the Terminal Inverted Repeats (orange) are provided in the annotation track. Mapped reads (forward reads: green; reverse reads: red; unspecific match: yellow) are shown in the mapping track. Reads mapping to the terminal inverted repeats are shown in yellow since they can map equally well at both terminal ends of the assembly. PacBio reads map contiguously to the mtDNA template, and no reads exceed the 5' or 3' terminal end of the reference sequence.

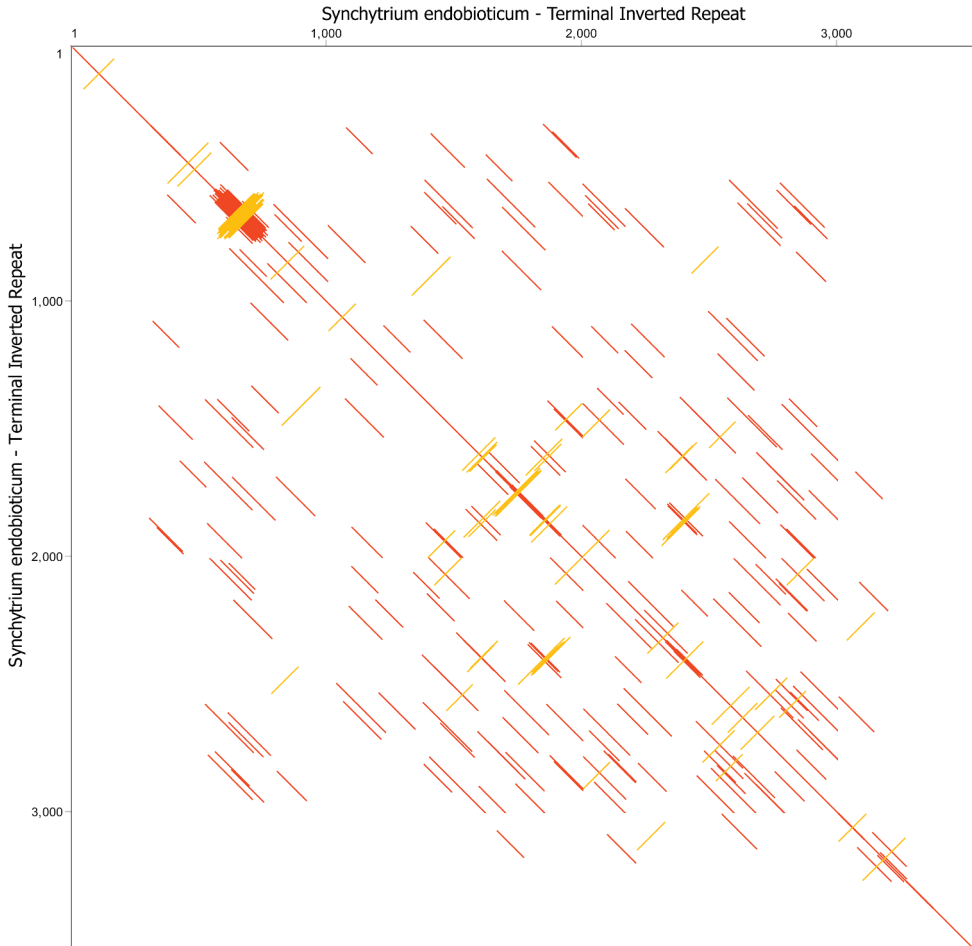


Figure S2. Dotplot internal repeat structure TIRs. Dotplot (self) of the 3,529 bp Terminal Inverted Repeat (TIR) of the 3' end of the linear *S. endobioticum* mtDNA. All positions from the TIR are compared with itself using the EDNAFULL (no ambiguous matches) substitution matrix using a window size of 100 bases and a similarity threshold of 50%. Similarities between a position from each sequence are plotted (in frame = red, reverse complement = orange).

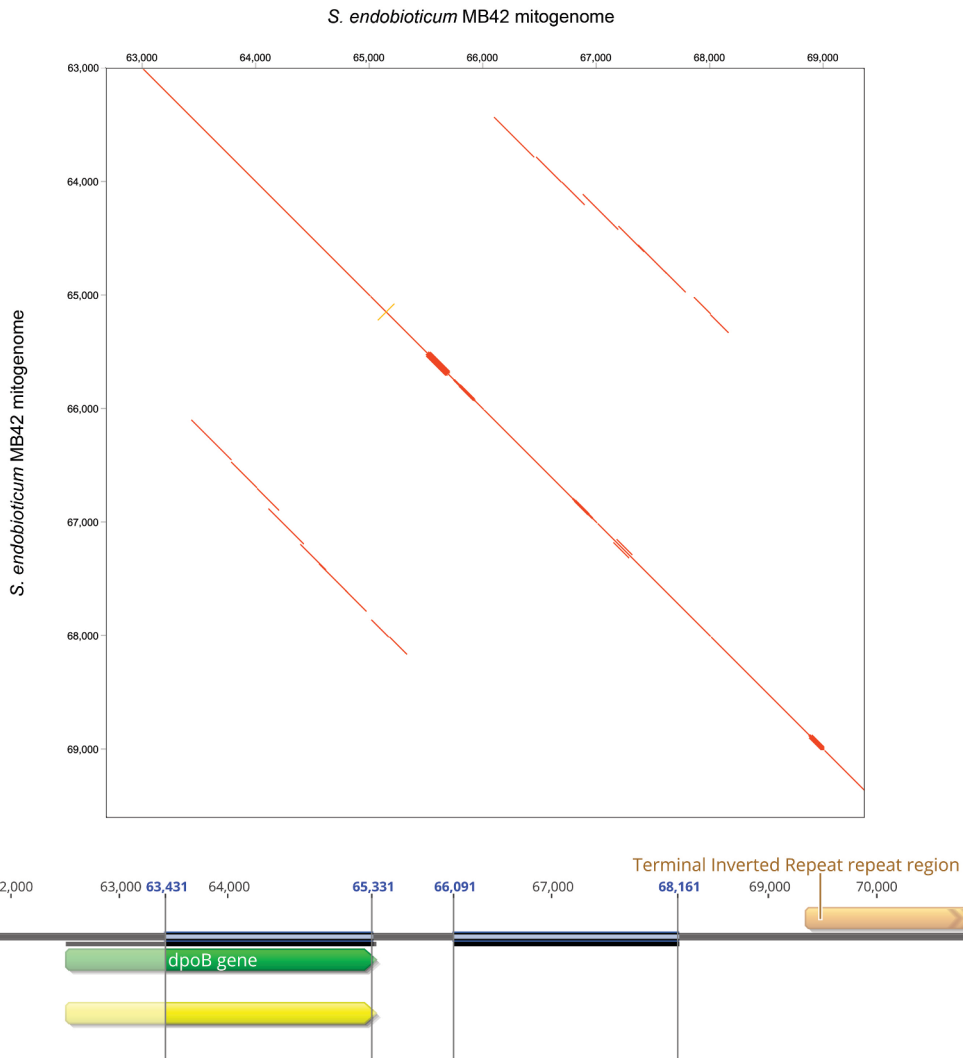


Figure S3 Dotplot repeat in AT-rich region containing the *dpoB* gene. Dotplot (self) of the 7,524 bp A-T rich region of the linear *S. endobioticum* mtDNA containing the *dpoB* gene. All positions from the A-T rich region are compared with itself using the EDNAFULL (no ambiguous matches) substitution matrix using a window size of 100 bases and a similarity threshold of 50%. Similarities between a position from each sequence are plotted (in frame = red, reverse complement = orange). The *dpoB* gene and its repeated sequence (similarity 54.7%) are highlighted in an annotated sequence track.

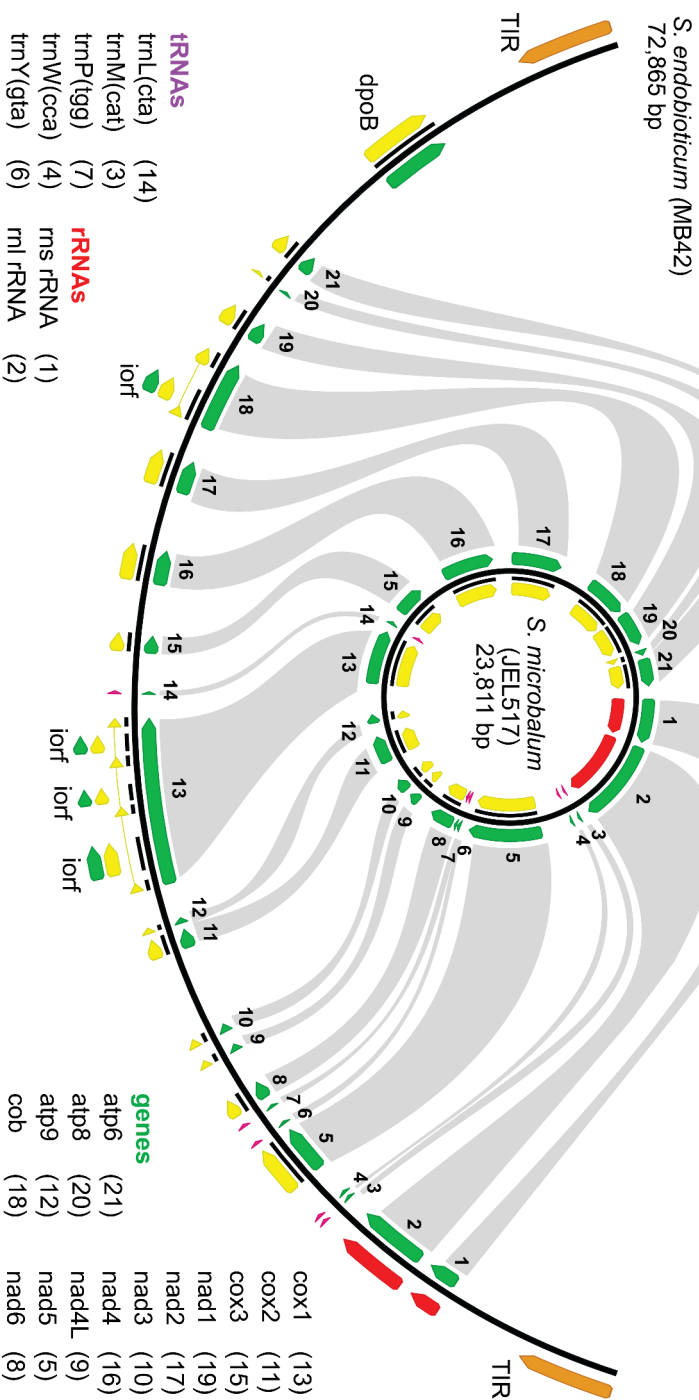


Figure S4. Comparison of the mitochondrial genomes of *S. endobioticum* MIB42 and *S. microbalum* JEL517. Organisation and orientation of the mitochondrial genes, tRNAs and rRNAs is conserved between the 72,865 bp linear mtDNA of *S. endobioticum* and the circular mapping 23,811 bp mtDNA genome of *S. microbalum*. Genes are annotated in green, CDS in yellow, tRNAs in purple, rRNAs in red, Terminal Inverted Repeats (TIR) in orange. Grey links between the two mtDNA genomes indicate the orthologous genes, tRNAs and rRNAs. Intronic open reading frames on the *S. endobioticum* genome are indicated "iof", and the DNA polymerase B gene is indicated "dpoB".

Haplotype network based on polymorphisms in CDS of mtDNA genes

Comparing 30 *S. endobioticum* isolates, 19 polymorphic sites were found in seven mitochondrial genes (*atp6*, *cob*, *cox1*, *nad1*, *nad4*, *nad5* and *nad6*). Three of the observed SNPs resulted in non-synonymous (dN) substitutions relative to the 1(D1) MB42 mtDNA: two in *atp6*, and one in *nad5*. In addition, the insertion of 6 bases resulted in the gain of two amino acids in *nad5* relative to the reference. Clustering of the informative sites using a Median Joining network resulted 10 different mitochondrial haplotypes (Fig. S5, table S3).

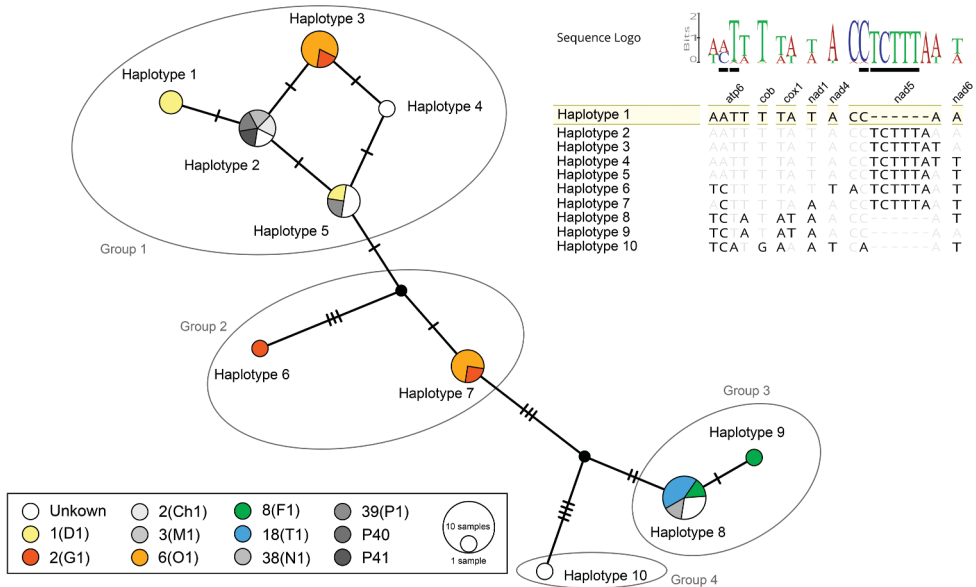


Figure 5. Median Joining haplotype network based on polymorphic sites in seven mitochondrial genes. Nodes in the network are coloured based on pathotype identity. Colours are used for pathotypes of major importance in Europe and Canada. Black nodes represent hypothetical ancestors. Marks on the branches indicate the number of mutations. Sequences of the informative sites and their corresponding genes are provided as an alignment using the genotype of isolate MB42 (haplotype 1) as reference. Differences to the reference are highlighted. The sequence logo shows the incidence of the different bases in the set of 30 *S. endobioticum* isolates analysed. Underlined bases in the sequence logo represent nucleotide changes that have an effect on the amino acid sequence.

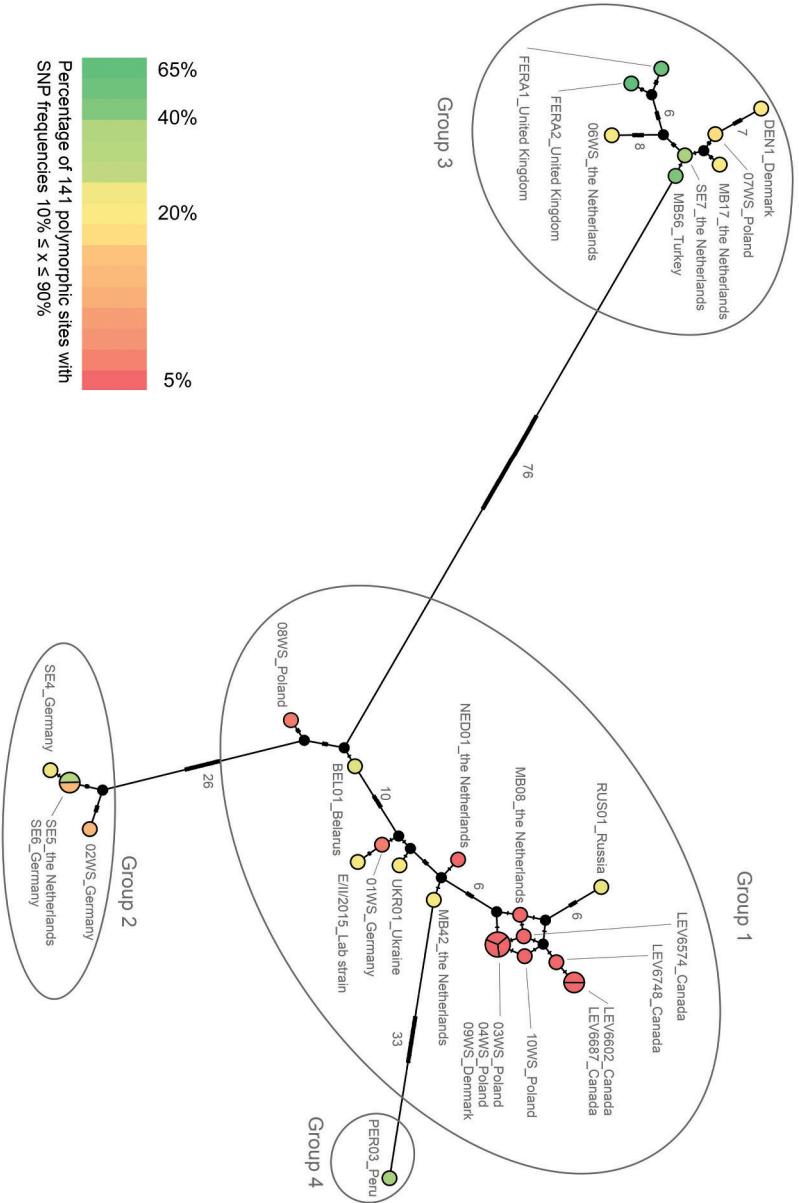


Figure S6. Within-isolate diversity in haplotype network based on polymorphic sites of the *S. endobioticum* mtDNA. Median joining haplotype network based on 141 polymorphic sites of the mtDNA. Nodes in the network are coloured based on the within-isolate diversity expressed as the percentage of polymorphic sites with intermediate ($10\% \leq x \leq 90\%$) frequencies. Black nodes represent hypothetical ancestors. Isolates from group three and the Peruvian isolate (group four) show high within-isolate diversity while several isolates from group one are scarce in diversity.

***Synchytrium taraxaci* Taqman assay**

Thirty-nine *Synchytrium* spp. sequences of the (partial) ribosomal DNA cassette (18S, ITS1, 5.8S, ITS2, and 28S) were downloaded from NCBI GenBank: *Synchytrium athyrii* (KF160863), *Synchytrium aureum* (KF160864), *Synchytrium australe* (KF160865), *Synchytrium collapsum* (KF160866), *Synchytrium cupulatum* (KF160867), *Synchytrium decipiens* (AY997094, DQ273819, DQ536475, KF160868), *Synchytrium endobioticum* (AJ784274, AY854021, JF795579, JF795580, KF160869, KF160870, KF160871, KF160872), *Synchytrium epilobii* (KF160873), *Synchytrium fulgens* (KF160874), *Synchytrium macrosporum* (AY997095, DQ273820, DQ322623, NG_017170, NG_027565), *Synchytrium minutum* (HQ324138, HQ324139), *Synchytrium papillatum* (JX093568, KF160875), *Synchytrium perforatum* (KF160876), *Synchytrium plantagineum* (KF160877), *Synchytrium puerariae* (EF053261,

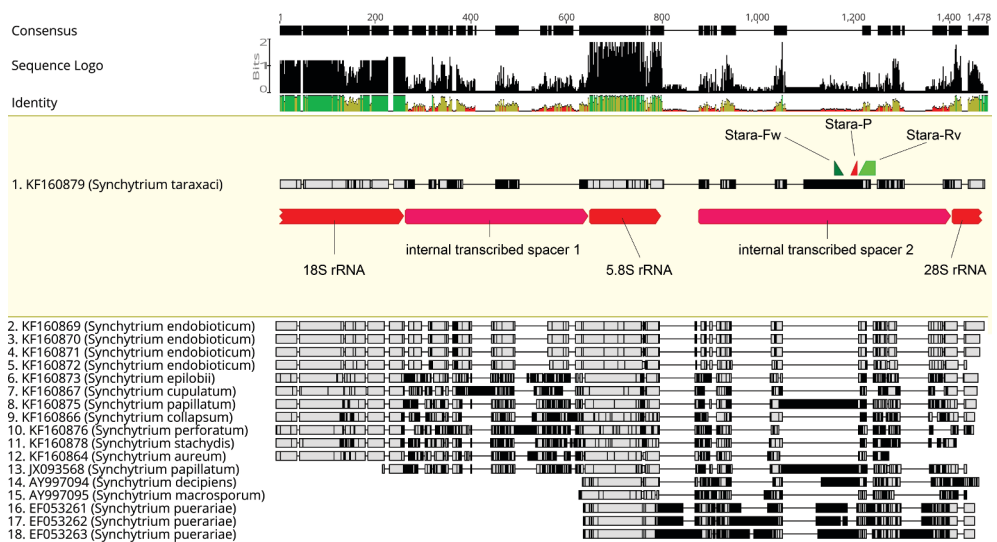


Figure S7. *Synchytrium taraxaci* TaqMan design. MAFFT alignment of 18 partial ribosomal cassette sequences covering 13 *Synchytrium* species: *S. aureum*, *S. collapsum*, *S. cupulatum*, *S. decipiens*, *S. endobioticum*, *S. epilobii*, *S. macrosporum*, *S. papillatum*, *S. perforatum*, *S. plantagineum*, *S. puerariae*, *S. stachydis*, and *S. taraxaci*. Sequence KF160879 (*S. taraxaci*) serves as reference sequence and is displayed on top of the alignment, remaining sequences are sorted by length. Bases in alignment similar to the reference are shown in grey and differences to the reference are highlighted (black). Ribosomal RNA genes (red) and the internal transcribed spacers (purple) are annotated on the reference sequence, as well as the annealing sites for the *S. taraxaci* specific primers Stara-Fw (dark green), Stara-Rv (light green) and probe Stara-P (Red).

EF053262, EF053263), *Synchytrium* sp. CAL-2013 (KF160861, KF160862), *Synchytrium* sp. DAOM 240977 (HQ317546), *Synchytrium* sp. DUH0009364 (DQ536476), *Synchytrium stachydis* (KF160878), and *Synchytrium taraxaci* (KF160879). Alignments of ITS1, 5.8S and ITS2 allowed inclusion of the highest number of specimens in the analysis, and the ITS2 region was chosen for the primer and probe design (Fig. S7). *In silico* analysis of primer and probe specificity showed that no non-specific amplification is to be expected from the species included in the design (Geneious R10, mismatches allowed = 3).

Real-time PCR reactions targeting the *S. taraxaci* ITS2 region were performed in a total volume of 30 μ L containing 1x Premix Ex Taq (TaKaRa), 250 nM of each primer (Stara-fw: 5'-GTTGTTTGGACCTTTTGTCCG-3', Stara-rv: 5'GAGTGGAAGATCAAAGGCATCTC-3'), 83 nM of probe (Stara-p: 5' FAM- CTGCATGACAGGACTGT -MGB 3'), and 3 μ L genomic DNA. Real-time PCR reactions were carried out in an ABI PRISM 7500 Sequence detector (ThermoFisher) using the following cycling conditions: 2 min at 94 °C, 40 cycles of 15 sec at 95 °C and 1 min at 60 °C. A synthetic construct (IDT, United States of America) of the *S. taraxaci* ITS accession KF160879 was used as positive amplification control to monitor the amplification efficiency. Cq values obtained for *S. taraxaci* spores extracted from dandelion leaves ranged from 19.9 to 23.3. No false positive results were obtained when testing *S. endobioticum* or *S. microbalum* DNA.

TdT-tailing controls

Positive and negative controls were included to monitor the effectiveness of the TdT tailing. The TdT tailing control consists of a 3,747 bp mtDNA insert cloned in a pUCIDT (IDT) high copy vector. The mitochondrial insert sequence contains the terminal inverted repeat (TIR) sequence, with modifications to allow synthesis, containing primer sites mtDNA_00007-fw and mtDNA00787-rv, the 5' specific sequence with primer site mtDNA_03691-rv, and the 3' specific sequence with primer site mtDNA_69209-fw. After digestion with *BsmBI* (New England Biolabs), a linear 6,041 bp fragment containing the mtDNA insert is obtained together with two smaller linear fragments (416 bp and 42 bp).

In its circular form the plasmid acts as a negative control for the TdT-tailing reaction, and after digestion the linear plasmid fragments are used as positive control. Efficiency of digestion is monitored using primer pair M13F/mtDNA_00787-rv: without digestion an amplicon is expected with this primer pair, and after full digestion no amplicon is expected since the plasmid is cleaved between primer site M13F and mtDNA_00787-rv (Fig 8a). Primer M13F is

not to be confused with primer M13FpolyG. The first anneals to a primer site on the circular TdT control, and forms a pair with primer mtDNA00787-rv: i.e. a control for digestion. The latter anneals to the digested, linear fragment after TdT tailing (poly-C addition), and forms a pair with primer mtDNA00787-rv: i.e. a control for TdT tailing.

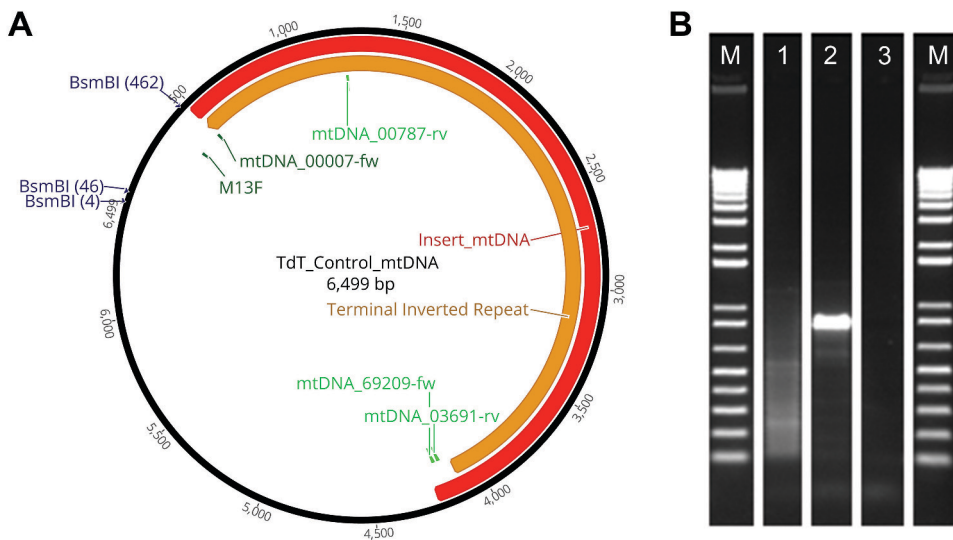


Figure S8. TdT tailing control. **A.** Circular conformation of the TdT control annotated with the mtDNA insert sequence (red), the modified TIR (orange), forward orientated primers (dark green), reverse orientated sequences (light green), and *BsmBI* restriction sites. **B.** Gel image of PCR reactions after TdT tailing using primer pair M13F-polyG x mtDNA_00787-rv. The following samples were tested: 1. undigested TdT control, 2. *BsmBI* digested TdT control, 3. molecular grade water. Estimation of amplicon size was done based on the Invitrogen 1kb plus maker (M).

The *S. endobioticum* mitochondrial genome

Table S1. Mitogenomes and mitochondrial genes included in the Bayesian inference (BI) of phylogeny. Genes indicated with a blank box were not detected on the mitochondrial genome or mitochondrial scaffold(s). Genes indicated with a “+” were detected, aligned and used to construct the BI phylogeny.

Species	Isolate	mtDNA scaffolds	Length (bp)	<i>atp6</i>	<i>atp8</i>	<i>atp9</i>	<i>cob</i>	<i>cox1</i>	<i>cox2</i>	<i>cox3</i>	<i>nad1</i>	<i>nad2</i>	<i>nad3</i>	<i>nad4</i>	<i>nad4L</i>	<i>nad5</i>	<i>nad6</i>
<i>Synchytrium endobioticum</i>	MB42	1	72,865	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>Chytridium confervae</i>	CBS 675.73	1	225,604	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>Powellomyces hirtus</i>	CBS 809.83	3	298,453	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>Spizellomyces pilustris</i>	CBS 455.65	1	126,474	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>Synchytrium microbalum</i>	JEL517	1	23,811	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>Synchytrium toraxaci</i>	Star13	14	39,246	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>Allomyces macrogynus</i>	ATCC 46923	1	57,473	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>Batrachochytrium dendrobatidis</i>	JEL423	1	175,284	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>Harpochytrium</i> sp.	JEL94	1	19,473	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>Harpochytrium</i> sp.	JEL105	1	24,169	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>Hyaloraphidium curvatum</i>	SAG 235-1	1	29,593	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>Monoblepharella</i> sp.	JEL15	1	60,432	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>Rhizophyidium brooksianum</i>	JEL136	1	68,834	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>Spizellomyces punctatus</i>	-	3	61,347	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Chapter 3

Table S2. Mitochondrial haplotype classification of *S. endobioticum* isolates. Assignment to mitochondrial lineage based on the clustering analysis of the entire mtDNA genome is indicated as mtDNA haplogroup, whereas clustering based on informative sites in coding sequences of seven mitochondrial genes is indicated as haplotype.

Haplotype (mtDNA CDS)	mtDNA haplogroup (lineage)	Isolate	Origin
1	1	1(D1)_MB42	the Netherlands
		1(D1)_NED_01	the Netherlands
2	1	2(Ch1)_03WS	Poland
		3(M1)_04WS	Poland
		P40_09WS	Denmark
		P41_10WS	Poland
		unknown_UKR_01	Ukraine
3	1	2(G1)_MB08	the Netherlands
		6(O1)_LEV6574	Canada
		6(O1)_LEV6602	Canada
		6(O1)_LEV6687	Canada
		6(O1)_LEV6748	Canada
4	1	unknown_RUS_01	Russian Federation
5	1	1(D1)_01WS	Germany
		39(P1)_08WS	Poland
		unknown_BEL_01	Belarus
		6(O1)_E/II/2015	Isolate obtained after two multiplications of 01WS on cultivar Erika
6	2	2(G1)_02WS	Germany
7	2	2(G1)_SE4	Germany
		6(O1)_SE5	the Netherlands
		6(O1)_SE6	Germany
8	3	8(F1)_DEN_01	Denmark
		18(T1)_07WS	Greece
		18(T1)_MB17	the Netherlands
		18(T1)_SE7	the Netherlands
		38(Nevsehir)_MB56	Turkey
		unknown_FERA_01	United Kingdom
		unknown_FERA_02	United Kingdom
9	3	8(F1)_06WS	the Netherlands
10	4	unknown_PER_03	Peru

The *S. endobioticum* mitochondrial genome

Table S3. Fungal materials and NextGen sequence information

<i>Synchytrium endobioticum</i>									
Pathotype	Isolate	Origin	Sequence data generated	SRA accession	Reads mapped to mtDNA	Mean coverage (x)	Length mtDNA (bp)		
1(D1)	MB42	Langenboom, the Netherlands	HiSeq PE100, 6.18 Gb MiSeq PE250, 5.12 Gb PacBio SMRT, 65.3 Mb	ERR2286952 ERR2286953 ERR2286954 ERR2698824	2,848,801	4.360	72,865		
1(D1)	1/2007/D1 (01WS)	Germany	HiSeq PE125, 7.08 Gb	ERR2286106	1,080,757	1.760	72,938		
1(D1)	NED01	Brabant, the Netherlands	HiSeq PE100, 7.56 Gb	ERR2286957	3,740,284	4.905	72,868		
2(G1)	4/2005/G1 (02WS)	Germany	HiSeq PE125, 29.9 Gb	ERR2286913 ERR2286914	1,450,581	2.387	72,824		
2(G1)	MB08	Mussel, the Netherlands	HiSeq PE125, 44.4 Gb	ERR2286949 ERR2286950	14,230,974	23.485	72,888		
2(G1)	BBA 2(G1) 09-04 (SE4)	Germany	HiSeq PE125, 51.1 Gb	ERR2286960 ERR2286961	5,368,430	8.837	72,839		
6(O1)	PL28/2007/2 (04WS)	Poland	HiSeq PE125, 30.3 Gb	ERR2286917 ERR2286918	1,121,611	1.862	72,889		
6(O1)	DK17/2015 (09WS)	Denmark	HiSeq PE125, 49.4 Gb	ERR2286930 ERR2286931 ERR2286932	344,557	1.25	72,888		
6(O1)	PL2/2015 (10WS)	Poland	HiSeq PE125, 45.8 Gb	ERR2286933 ERR2286934 ERR2286935	226,664	1.25	72,889		
6(O1)	E/II/2015	Isolate obtained after two multiplications of 01WS on variety Erika	HiSeq PE125, 7.18 Gb	ERR2286939	46,294	75	72,908		
6(O1)	LEV6574	St. Eleanor's, Prince Edward Island, Canada	MiSeq PE300, 38.9 Gb	ERR2286942 ERR2286943 ERR2286944 ERR2286945 PacBio SMRT, 35.5 Mb ERR2698823	65,232,021	163.200	72,889		

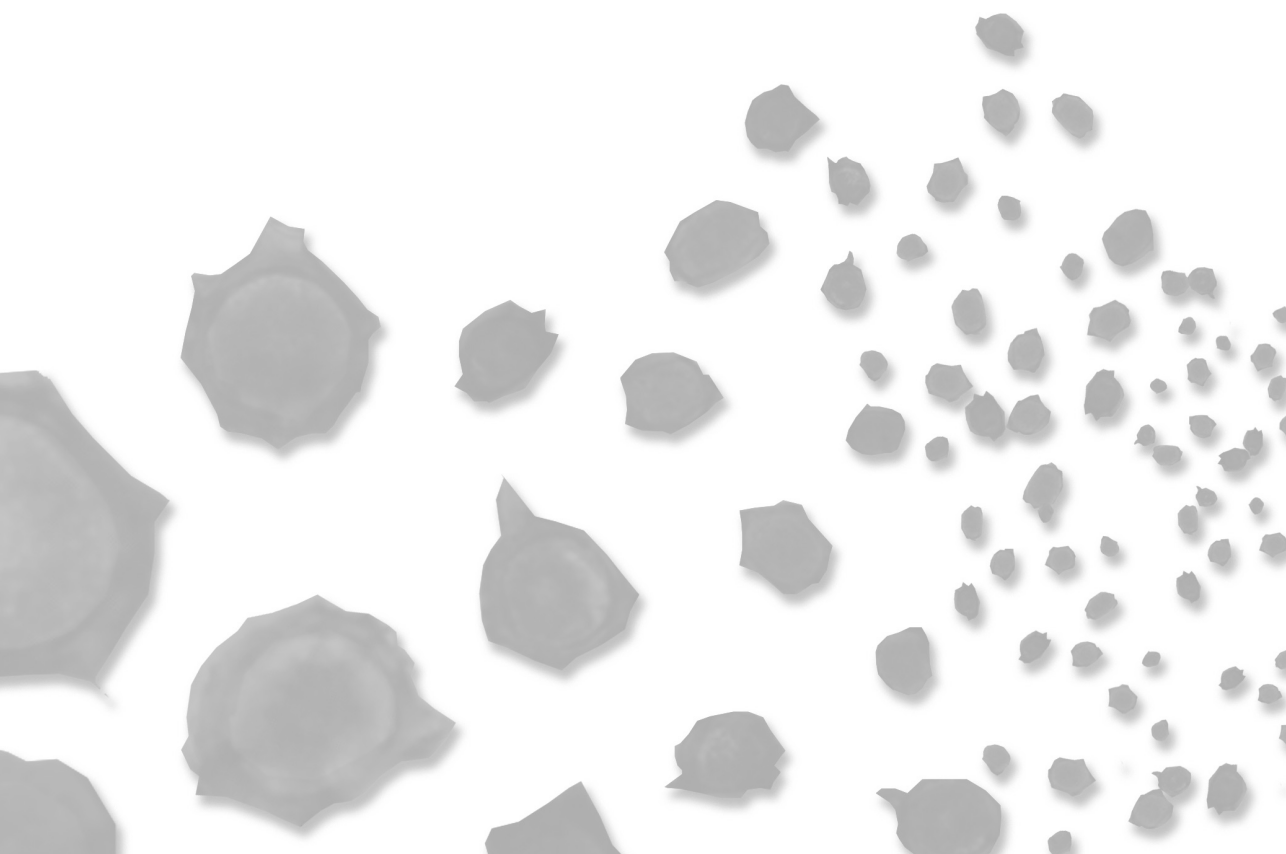
Table S3 (continued). Fungal materials and NextGen sequence information

<i>Synchytrium endobioticum</i>		Origin	Sequence data generated	SRA accession	Reads mapped to mtDNA	Mean coverage (x)	Length mtDNA (bp)
Pathotype	Isolate						
6(O1)	LEV6602	Augustine Cove, Prince Edward Island, Canada	MiSeq PE300, 7.54 Gb	ERR2286946	12,497,168	41,913	72,889
6(O1)	LEV6687	New Annan, Prince Edward Island, Canada	MiSeq PE300, 12.1 Gb	ERR2286947	10,504,480	35,926	72,891
6(O1)	LEV6748	New Glasgow, Prince Edward Island, Canada	MiSeq PE300, 12.0 Gb	ERR2286948	3,242,947	251	72,894
6(O1)	HLB 6(O1) 02-06 (SE5)	the Netherlands	HiSeq PE100, 6.05 Gb	ERR2286962	2,025,653	2,689	72,884
6(O1)	BBA 6(O1) 05_8.3 (SE6)	Germany	HiSeq PE100, 42.6 Gb	ERR2286963 ERR2286964	20,702,761	33,795	72,900
8(F1)	DEN01	Jylland, Denmark	HiSeq PE125, 62.1 Gb	ERR2286937 ERR2286938	4,197,193	6,938	72,872
8(F1)	3/2005/F1 (06WS)	the Netherlands	HiSeq PE125, 17.3 Gb	ERR2286919 ERR2286920 ERR2286921 ERR2286922 ERR2286923 ERR2286924 ERR2286925	1,332,731	2,653	72,941
8(F1)	2/2005/Ch1 (03WS)	Poland	HiSeq PE125, 31.6 Gb	ERR2286915 ERR2286916	1,533,292	2,540	72,889
18(T1)	GR2/2015 (07WS)	Greece	HiSeq PE125, 6.70 Gb	ERR2286926	1,679,744	2,763	72,931
18(T1)	MB17	Borgercompagnie, the Netherlands	HiSeq PE100, 6.94 Gb	ERR2286951	1,821,241	2,386	72,907
18(T1)	HLB P18(T1) -02-06 (SE7)	Borgercompagnie, the Netherlands	HiSeq PE100, 7.91 Gb	ERR2286965	2,495,395	3,291	72,896
38(N1)	MB56	Neveshir, Turkey	HiSeq PE100, 34.5 Gb	ERR2286955 ERR2286956	19,870,357	32,255	72,910
39(P1)	PL69/2009 (08WS)	Plekielnik, Poland	HiSeq PE100, 48.4 Gb	ERR2286927 ERR2286928 ERR2286929	1,679,744	4,037	72,874

The *S. endobioticum* mitochondrial genome

Table S3 (continued). Fungal materials and NextGen sequence information

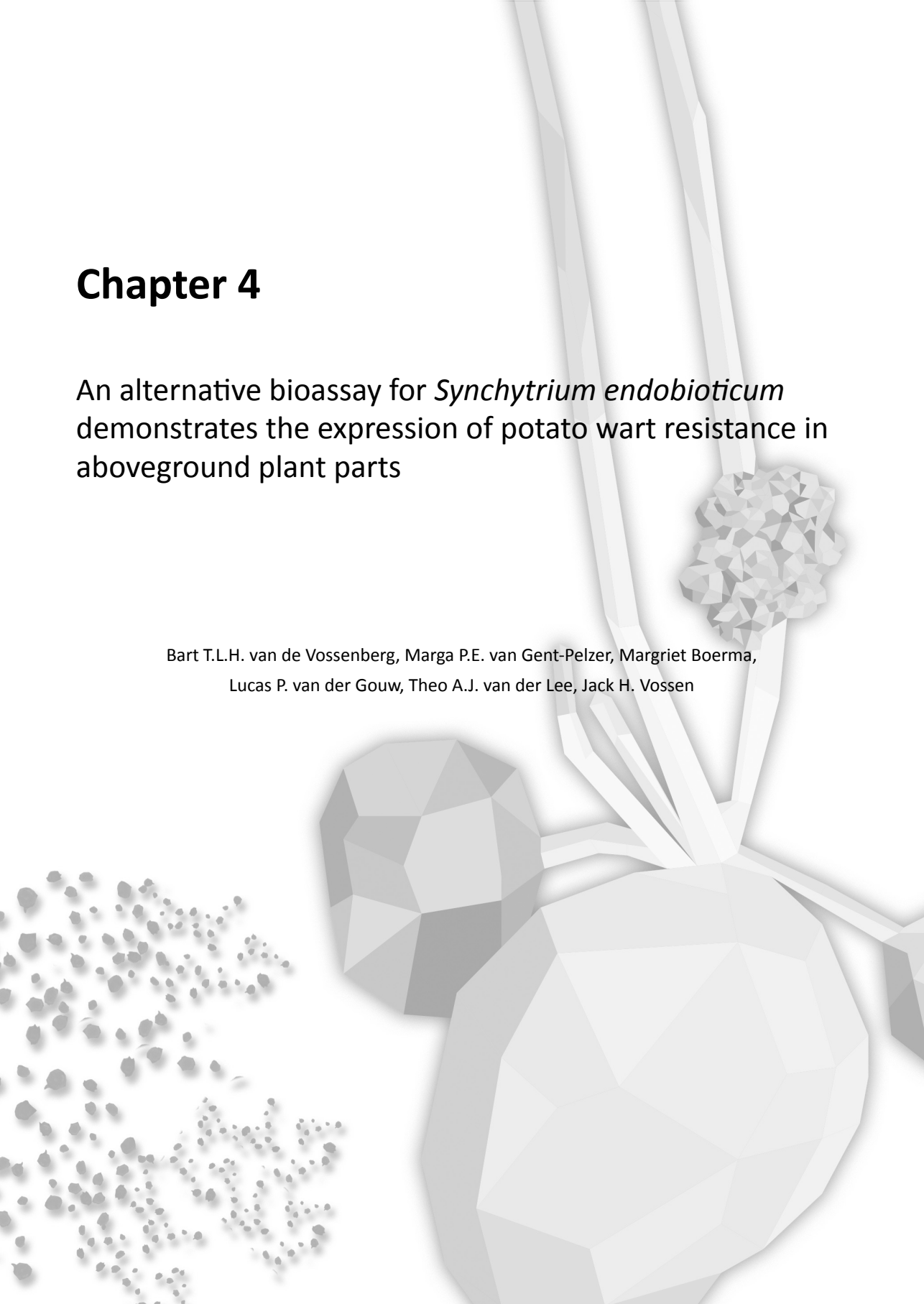
<i>Synchytrium endobioticum</i>							
Pathotype	Isolate	Origin	Sequence data generated	SRA accession	Reads mapped to mtDNA	Mean coverage (x)	Length mtDNA (bp)
unknown	BELO1	Belarus	HiSeq PE100, 1.31 Gb	ERR2286936	43,067	68	72,874
unknown	FERA1	United Kingdom	HiSeq PE100, 32.6 Gb	ERR2286940	387,230	493	72,775
unknown	FERA2	United Kingdom	HiSeq PE100, 33.6 Gb	ERR2286941	400,095	510	72,810
unknown	PER03	Aco Paucartambo, Pasco, Peru	HiSeq PE100, 6.57 Gb	ERR2286958	10,418,036	17,309	72,814
unknown	RUS01	St. Petersburg, Russian Federation	HiSeq PE100, 1.94 Gb	ERR2286959	123,602	202	72,879
unknown	UKR01	Ukraine	HiSeq PE100, 2.25 Gb	ERR2286966	31,611	50	72,879
Other Chytridiomycota							
Species	Isolate	Origin	Sequence data generated	SRA accession	Reads mapped to mtDNA	Mean coverage (x)	Length mtDNA (bp)
<i>Chytridium congeriae</i>	CBS 675.73	Canada	HiSeq PE125, 3.50 Gb	ERR2451316	3,113,642	1,731	225,604
<i>Powellomyces hirtus</i>	CBS 809.83	the Netherlands	HiSeq PE125, 28.6 Gb	ERR2451317	29,555,878	6,929	298,453
<i>Spizellomyces palustris</i>	CBS 455.65	Germany	HiSeq PE125, 3.57 Gb	ERR2451319	3,538,853	3,508	126,474
<i>Synchytrium microbali</i>	JEL517	Hancock Co., Maine, USA	HiSeq PE125, 4.23 Gb	ERR2451318	247,495	1,300	23,811
<i>Synchytrium toraxaci</i>	Star13	Bennekorn, the Netherlands	HiSeq PE125, 27.0 Gb	ERR2451320	7,656,249	26,745	39,246



Chapter 4

An alternative bioassay for *Synchytrium endobioticum* demonstrates the expression of potato wart resistance in aboveground plant parts

Bart T.L.H. van de Vossenberg, Marga P.E. van Gent-Pelzer, Margriet Boerma,
Lucas P. van der Gouw, Theo A.J. van der Lee, Jack H. Vossen



Chapter 4

This chapter was published as “van de Vossen B.T.L.H., van Gent-Pelzer M.P.E., Boerma M., van der Gouw L.P., van der Lee T.A.J., and Vossen J.H. (2019) An alternative bioassay for *Synchytrium endobioticum* demonstrates the expression of potato wart resistance in aboveground plant parts, *Phytopathology*, doi: 10.1094/PHYTO-01-19-0024-R” with additional supplementary materials. Slight formatting changes were made to facilitate thesis consistency.

Abstract

The obligate biotrophic chytrid species *Synchytrium endobioticum* is the causal agent of potato wart disease. Currently thirty-nine pathotypes have been described based on their interaction with a differential set of potato varieties. Wart resistance and pathotyping is performed using bioassays in which etiolated tuber sprouts are inoculated.

Here we describe an alternative method in which aboveground plant parts are inoculated. Susceptible plants produced typical wart symptoms in developing, but not in fully expanded, aboveground organs. Colonization of the host by *S. endobioticum* was verified by screening for resting spores by microscopy and by molecular techniques using TaqMan PCR and RNAseq analysis. When applied to resistant plants, none of these symptoms were detectable. Recognition of *S. endobioticum* pathotypes by differentially resistant potato varieties was identical in axillary buds and the tuber-based bioassays. This suggests that *S. endobioticum* resistance genes are expressed both in etiolated “belowground” sprouts and green aboveground organs. RNAseq analysis demonstrated that the symptomatic aboveground materials contain less contaminants compared to resting spores extracted from tuber-based assays. This reduced microbial contamination in the aboveground bioassay could be an important advantage to study this obligate biotrophic plant-pathogen interaction.

As wart resistance is active in both below and above ground organs, the aboveground bioassay can potentially speed up screening for *S. endobioticum* resistance in potato breeding programs as it omits the requirement for tuber formation. In addition, possibilities arise to express *S. endobioticum* effectors in potato leaves through agroinfiltration, thereby providing additional phenotyping tools for research and breeding.

Introduction

Synchytrium endobioticum is a soilborne obligate biotrophic fungal pathogen causing potato wart disease, and belongs to the basal fungal division of Chytridiomycota. The disease is characterized by hypertrophic growth and callus-like wart structures on potato tubers [1]. Warts can vary strongly in size from less than pea-sized proliferations to the size of a fist [2]. In the infected tissue, summer sporangia are developed, leading to new infections. Belowground warts are white to brown, whereas aboveground malformed tissues are green. Both types of wart turn black upon host tissue decay which coincides with enhanced resting (or winter) spore formation. Decomposition of warted tissue allows release of these spores into the surrounding soil. Resting spores have been reported to remain viable and infectious in fields without potato cultivation for decades [3]. The original description of *S. endobioticum* by Schilberszky [4] was based on infected potato tubers, but all parts of the shoot system can be infected while the root system is immune [5].

In the early 1900s, potato varieties resistant to *S. endobioticum* were discovered [6] which represented the start of breeding programs for potato wart resistance. Until 1941 only a single pathotype was recognized but currently 39 pathotypes (races) of the pathogen are acknowledged based on their interaction using a set of differential potato varieties [7-9], of which pathotypes 1(D1), 2(G1), 6(O1) and 18(T1) are considered to be of main importance in Europe. In Canada, mainly pathotypes 2(G1) and 6(O1) have been reported [2].

Once introduced into fields with potato cultivation, there are no effective chemical agents to control the disease [10]. Strict phytosanitary measures combined with the use of resistant potato varieties are the only successful strategy to prevent introduction and spread of the pathogen. For these reasons, *S. endobioticum* has a quarantine status in most countries world-wide [11], and is included on the HHS and USDA select agent list.

Studying biotrophic pathogens heavily depends on the quality of the bioassays to study the interaction between pathogens and their hosts. In addition, soilborne obligate biotrophic organisms represent a unique challenge, as the complex microbiome of the soil is an obscure and possibly disturbing component in the interaction. Bioassays are also exploited in numerous pathosystems to determine pathogen virulence or to group pathogens in races or pathotypes. Grouping of *S. endobioticum* isolates into pathotypes is of major importance and fundamental to phytosanitary measures to restrict the spread of the disease [12]. Two different bioassays are currently used for pathotype grouping; Spieckermann [13] and

Glynne-Lemmerzahl [14-16], of which the latter is most frequently used in Europe [17]. In Spieckermann assays, composts containing resting spores are used as inoculum, whereas in the Glynne-Lemmerzahl assays, fresh wart cuttings, which mainly contain summer sporangia, serve as inoculum. Apart from pathotype identification, these bioassays are used in breeding programs to determine the resistance or susceptibility of new potato genotypes to selected *S. endobioticum* pathotypes. Even though the Spieckermann and Glynne-Lemmerzahl bioassays are routinely applied for pathotyping and resistance screening, these assays remain challenging even though international programs were established to harmonize and optimize these bioassays. Alternatives to these tuber-based assays have been described. Hampson used micro-propagated plantlets and showed that these produced more consistent infection percentages and were more sensitive than the tuber based assays [18, 19]. In 1962, Ullrich used a leaf-based bioassay to study the morphology of malformed host tissues [20]. Nevertheless, these methods were never adopted by the research community.

Without culturing options, and since no transformation protocols are described for the obligate biotrophic *S. endobioticum*, heterologous systems need to be used to study the plant-pathogen interaction on a molecular level. Transient expression of pathogen genes in plants with the leaf-based agroinfiltration system has been used in other pathosystem to elucidate their role in pathogenicity [21]. Since resistance to plant pathogens in potato can occur in an organ-specific manner [22], similar expression of potato wart resistance in aboveground plant parts compared to the tuber-based bioassays is key to use the leaf-based agroinfiltration system for *S. endobioticum*.

Here we describe a bioassay targeting aboveground plant parts (i.e. axillary buds) that was developed to determine if: 1. potato resistance genes are active in aboveground plant parts producing the same differential responses as the tuber-based assays; and 2. to eliminate the complex microbial contamination that is associated with the soil matrix from the potato- *S. endobioticum* interaction to generate more suitable material to study the interaction on a molecular level. Finally, we discuss the potential use of the aboveground bioassay to speed up screening for *S. endobioticum* resistance in breeding programs.

Results

Optimization experiments

A total of 153 aboveground plant parts were treated, of which 101 were inoculated with fresh wart cuttings, 31 were inoculated with compost containing resting spores, and 21 inoculations did not have any inoculum (i.e. mock treatments) (Fig. 1). From the combinations with inoculum, 53 were on variety Kuba which is resistant to pathotype 1(D1), whereas 79 were performed with the varieties Deodara and Tomensa which are susceptible to pathotype 1(D1). None of the mock treated or inoculated resistant plants produced any symptoms, whereas 14 inoculated aboveground plant parts of the susceptible plants (18%) produced malformed tissue. When further differentiating the formation of malformed tissues according to their respective treatments, wart inoculated axillary buds wrapped with household plastic produced the highest percentage of symptoms regardless of their

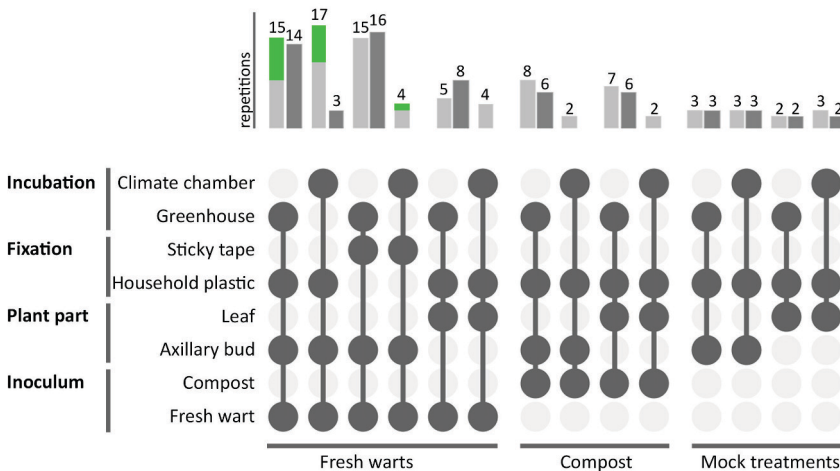


Figure 1. Optimization of the aboveground bioassay. The different treatments are divided in four groups (vertical), i.e. inoculum type used (pathotype 1(D1) compost or fresh warts); plant part inoculated (expanded leaf or axillary bud); how the inoculum was fixed to the plant; and the incubation location. The different treatment combinations are sorted in three groups (horizontal), i.e. plant parts inoculated with fresh wart pieces, compost, and mock treated plants (no inoculum). Combinations tested are shown in dark grey, and the bar graphs indicate the number of repetitions performed for each combination with susceptible (light grey) and resistant (dark grey) varieties. Green bars indicate the number of symptomatic plants, and the numbers above the bars indicate the total number of repetitions performed.

incubation method: 47% and 35% of the inoculated axillary buds for plants incubated under transparent foil and in the climate chamber respectively. None of the inoculated leaves resulted in malformation of the plant tissue. Similarly, none of the compost inoculated plant parts produced any symptoms.

Proliferation of the pathogen was confirmed using a species specific TaqMan PCR, which resulted in significantly lower (general ANOVA, $p < 0.001$) Cq values for inoculated plant parts showing symptoms ($Cq_{\text{mean}} = 17.1 \pm 3.1$) compared to mock treated plants (33.2 ± 3.2) and inoculated plant parts without signs of the disease (33.1 ± 2.7) (Fig. 2). In the susceptible varieties Deodara and Tomensa, seven inoculated plant parts were scored as “unclear” (examples see Fig. S2). Two of these samples produced Cq values consistent with proliferation of the pathogen, while the others produced Cq values similar to the mock treated plants and inoculated plant parts showing no visual symptoms. A bimodal distribution is obtained for Cq values obtained from infected plant parts compared to those obtained

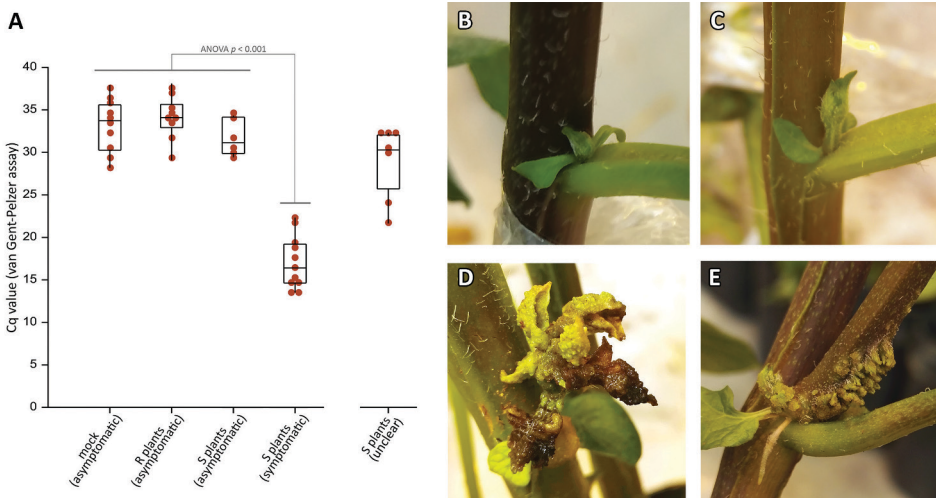


Figure 2. Observations from optimization experiments. A. Cq values obtained with the *S. endobioticum* specific TaqMan PCR described by Van Gent-Pelzer *et al.* (2010) from asymptomatic mock treated plants (Kuba, Deodara and Tomensa), asymptomatic resistant plants (Kuba), asymptomatic susceptible plants (Deodara and Tomensa), and susceptible plants with clear malformations. Inoculated plant parts for which the reaction was not clear are included as a separate group (S plants, unclear). Red dots in the boxplots indicate the individual Cq values. Examples of reaction types observed: B. asymptomatic on susceptible variety Tomensa; C. asymptomatic on resistant variety Kuba; D. symptomatic malformation of leaf tissue on susceptible variety Deodara; and E. symptomatic malformation of shoot tissue and aerial root formation on susceptible variety Deodara.

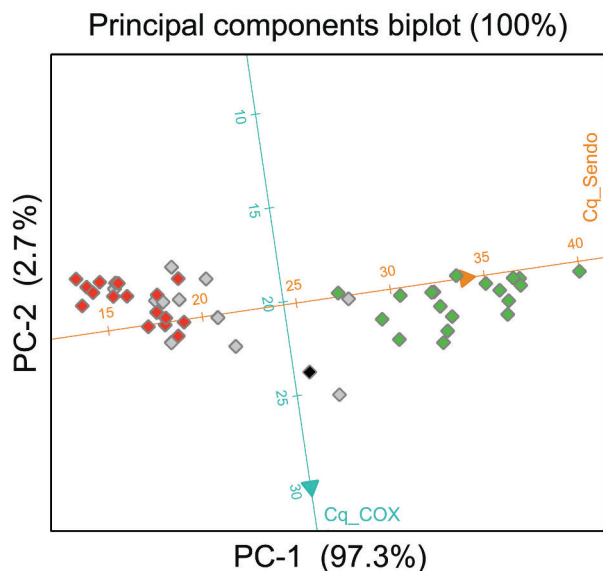


Figure 3. Biplot of Cq values from aboveground inoculated EPPO differential potato varieties obtained with the *S. endobioticum* specific TaqMan PCR (PC-1) and a generic plant *cox1* TaqMan PCR. Samples visually scored as “symptomatic” are shown in red, asymptomatic plant parts are shown in green, and samples scored as “unclear” are shown in grey. A single inoculation site where the inoculum was not removed until sampling which resulted in rot is shown in black. Almost all variation (97.3 %) in the dataset is attributed the *S. endobioticum* specific Cq values.

Table 1. Final scoring results 35 to 59 days post inoculation of EPPO differential potato varieties with *S. endobioticum* pathotypes 1(D1), 2(G1), 6(O1), and 18(T1). The susceptible (S) or resistant (R) interaction for a given variety-pathotype combination based on the Glynne-Lemmerzahl method is listed under “suspected interaction”. Susceptible interactions observed in the aboveground bioassay are shown in bold.

Potato variety	Anticipated interaction	Inoculum	Inoculum				mock
			1(D1)	2(G1)	6(O1)	18(T1)	
Deodara	S: 1, 2, 6, 18	Inoculated buds	11	15	12	13	4
		Symptomatic	8 (72%)	7 (47%)	3 (25%)	6 (46%)^c	0
Producent	S: 2, 6, 18; R:1	Inoculated buds	8	16	12	12	4
		Symptomatic	0	9 (56%)	10 (83%)	7 (58%)	0
Talent	S: 2 ^a , 6 ^a , 18; R:1	Inoculated buds	8	10	12	12	4
		Symptomatic	0	2 (20%)^b	4 (33%)^b	4 (33%)^d	0
Saphir	S: 2; R:1, 6, 18	Inoculated buds	8	12	12	12	4
		Symptomatic	0	3 (25%)	0	0	0
Belita	R:1, 2, 6, 18	Inoculated buds	8	12	12	12	4
		Symptomatic	0	0	0	0	0

a. Slightly susceptible reactions in tuber based assays: non-necrotic sori fields are observed but not full wart formation [12]; b. initially scored as “unclear” but TaqMan verified; c. four samples were initially scored as “unclear” but TaqMan verified; d. one additional sample was scored as “unclear”.

from healthy plant parts (Fig. S3). It was concluded that treatment combination “wart, household plastic, bud, greenhouse” is the most efficient one and is further referred to as the “aboveground bioassay”.

Aboveground inoculation of EPPO differential set

Five potato varieties (i.e. Deodara, Producent, Talent, Saphir and Belita) listed in the EPPO differential set for *S. endobioticum* pathotype identification with the Glynn-Lemmerzahl method [12] were inoculated using the aboveground bioassay. A total of 233 axillary buds were inoculated with fresh warts on Deodara obtained with pathotypes 1(D1), 2(G1), 6(O1), and 18(T1). Two months after inoculation, pathotype 1(D1) produced symptoms on Deodara on the majority (72%) of inoculated axillary buds but not on resistant varieties (table 1). Pathotype 2(G1) produced clear symptoms on varieties Deodara, Producent and Saphir. For Deodara and Producent the success rate of the inoculation was higher compared to Saphir: 47%, 56%, and 25% respectively. In Talent inoculated with pathotype 2(G1), two of the replicates were scored as “unclear phenotype” whereas the other eight produced no visible symptoms. Pathotype 6(O1) produced symptoms on Deodara (25%) and Producent (83%), of which the latter had the highest success rate for any of the pathotype-variety combinations tested. Similar to pathotype 2(G1), pathotype 6(O1) only produced four slight malformations on Talent. Pathotype 18(T1) produced malformed tissues on varieties Deodara (46%), Producent (58%) and Talent (33%). Inoculated plant parts sampled 35 dpi for microscopy and molecular verification showed the presence of spores in the malformed tissue (Fig. S4).

Verification of pathogen proliferation using the *S. endobioticum* specific TaqMan assay on the EPPO differential set yielded similar results compared to the Cq values obtained in the optimization experiments. The *S. endobioticum* specific TaqMan PCR could be used to distinguish healthy plant parts from infected plant parts (Fig. 3). Symptomatic plant samples from susceptible varieties produced significantly (general ANOVA, $p < 0.001$) lower Cq values compared to mock treated plants and resistant varieties with no symptoms: i.e. $Cq_{\text{means}} 16.4 \pm 1.8$ versus 35.5 ± 2.3 ; and 30.8 ± 3.1 respectively. Of the eleven samples initially scored as unclear, nine produced Cq values consistent with proliferation of the pathogen. These included both susceptible interactions (pathotype 18(T1) inoculated Deodara plants) and interactions that are formally described as slightly susceptible (pathotype 2(G1) and 6(O1) infected Talent plants).

Chapter 4

Interestingly, in some cases malformed tissue regrew after removal of the warted tissue for sampling at 35 dpi (Fig. 4). The new outgrowth of the malformed tissue appeared 24 days (59 dpi) after the first sampling on varieties Deodara and Producent. For Deodara, a single pathotype 1(D1) and a single pathotype 2(G1) inoculated site showed outgrowth of new malformed tissue, whereas for Producent three 2(G1), a single 6(O1), and three 18(T1) inoculated sites showed new outgrowth. This suggests that not all of the affected tissue was removed allowing new outgrowth of malformed tissue. In addition, the development of warts on belowground plant parts, which were not actively inoculated, of five Deodara and two Producent in combination with pathotypes 1(D1) and 2(G1) became apparent 59 dpi (Fig. S5).

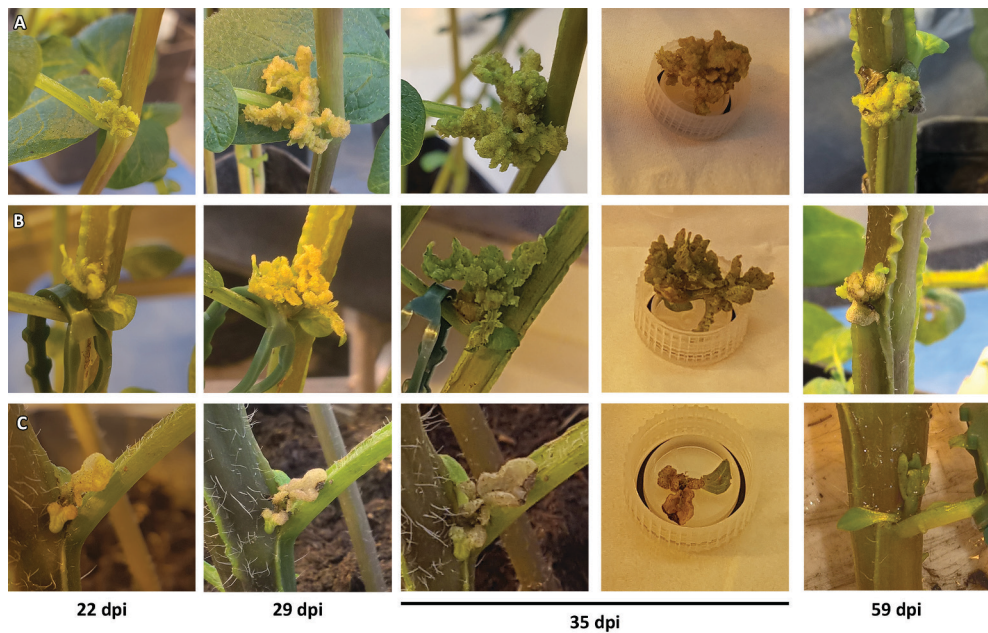


Figure 4. Development of symptoms of *S. endobioticum* inoculated axillary buds at 22 dpi until sampling at 35 dpi. All visibly malformed tissue was removed, but outgrowth of new malformed tissue was observed at 59 dpi. **A.** Producent (plant 1, bud 3) inoculated with pathotype 18(T1); **B.** Producent (plant 3, bud 2) inoculated with pathotype 2(G1); and **C.** Deodara (plant 3, bud 1) inoculated with pathotype 1(D1).

RNAseq analysis

Three samples per pathotype from Deodara axillary buds exhibiting early signs of malformation were subjected to RNAseq. In each sample expression of *S. endobioticum* genes was found (table 2, table S3). Mapping to the RNAseq reads to the MB42 reference genome shows expression for the *S. endobioticum* household gene elongation factor 1 α (*SeEF1 α* ; SeMB42_g00306) ranging from 140 to 2,429 Transcripts Per Million (TPM). In resting spores obtained from fresh warts on tuber sprouts in an independent inoculation experiment, *SeEF1 α* TPM values ranged from 408 to 917. For the malformed axillary buds, 0.4 to 6.4% of all generated reads map to *S. endobioticum*, whereas the majority of reads maps to its host potato (84.1 to 94.5%). This contrasts to resting spores extracted from fresh belowground

Table 2. Statistics RNAseq analysis for aboveground (axillary bud) and belowground samples (resting spores from tubers). The sample names of the belowground parts are abbreviated as follows: D+ DNase treatment; D- no DNA treatment; pA polyA enrichment; ribo- plant ribosomal RNA depletion.

Sample	Patho- type	Sample	Total reads (* 10 ⁶)	Reads mapped to SeMB42	Reads mapped to potato	Un- mapped reads	SeMB42 genes with TPM \geq 3	Contigs from un- mapped reads
Axillary bud	1(D1)	1	36.6	3.3%	90.1%	7%	87%	2,928
		2	37.9	1.6%	92.3%	6%	81%	2,762
		3	37.9	0.8%	94.5%	5%	73%	1,458
	2(G1)	1	37.9	6.4%	84.1%	10%	90%	10,156
		2	37.8	2.8%	93.2%	4%	56%	1,453
		3	35.8	1.5%	94.2%	4%	85%	1,720
	6(O1)	1	39.6	1.4%	93.8%	5%	61%	2,208
		2	37.8	1.0%	90.4%	9%	60%	3,952
		3	38.2	3.4%	91.7%	5%	89%	1,784
	18(T1)	1	38.5	0.4%	92.8%	7%	4% ^a	3,784
		2	35.7	1.0%	93.5%	6%	68%	2,752
		3	39.9	5.2%	85.2%	10%	91%	1,919
Resting spores from tubers	1(D1)	D+,pA	95.8	41.8%	0.6%	58%	88%	10,328
		D-,pA	99.3	42.5%	0.6%	57%	91%	40,935
		D-,ribo-	122.3	16.1%	2.0%	82%	91%	85,248
		D-,ribo-	102.2	18.1%	9.2%	73%	92%	96,796

a. this number is an overestimation of the true expression as a result of the bias introduced by TPM in low expressed datasets.

Chapter 4

warts which have more reads mapping to *S. endobioticum* (16.1 to 42.5%) and very few reads mapping to potato (0.6 to 9.2%). In the malformed axillary buds 4.0 to 9.6% of all reads remain unmapped. In contrast, the resting spores extracted from fresh warts on tubers that have 56.9 to 81.9% of all reads remaining unmapped. The unmapped reads were believed to be derived of other biological agents represented in the aboveground samples and belowground samples To putatively identify these agents, *de novo* assemblies were performed of the unmapped reads, which resulted in a significantly (T-test, $p < 0.001$) higher number of *de novo* assembled transcripts for the samples derived from resting spores from fresh warts.

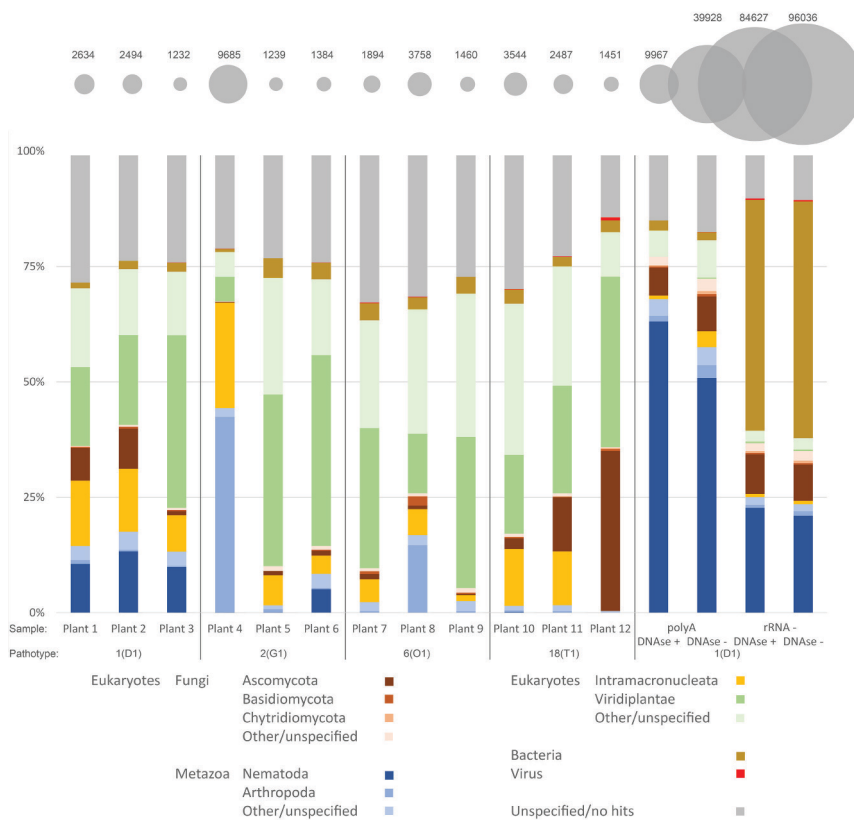


Figure 5. Relative abundance of taxa other than *S. endobioticum* or potato in aboveground plant parts and resting spores extracted from fresh tuber warts. Bubble plots represent the number of contigs included in the blast analysis. Plants 1-3 were inoculated with pathotype 1(D1), plants 4-6 with 2(G1), plant 7-9 with 6(O1), and plants 10 -12 with 18(T1). The last four samples represent the belowground derived RNAseq datasets with different treatments (polyA = polyA enrichment, rRNA- = ribosomal plant RNA depletion, DNase + or - treatment).

Of the *de novo* assembled transcripts, still 2.2% produced significant (e-value < 1e-20) hits with either the *S. endobioticum* pathotype 6(O1) LEV6574 genome or the SolTub v3.0 potato genome, and were removed from the analysis. For the remaining 264,128 contigs, approximately 90% could be identified as eukaryotic (further specified as fungal, metazoan, ciliates and plant), bacterial or viral (Fig. 5, table S4). For the aboveground derived samples, on average a quarter of all contigs were believed to originate from the host plant. This in contrast to the belowground derived samples which had very low percentages (<0.5%) of plant hits. Species from the ciliate subphylum Intramacronucleata were abundantly present in both aboveground and belowground samples, but in the latter category they represented only a very small portion of all transcripts. Almost no non-*S. endobioticum* Chytridiomycota hits were obtained with the aboveground plant parts. In the belowground plant parts however, 23 to 500 non-*S. endobioticum* Chytridiomycota transcripts were detected. Ascomycota hits were obtained in both aboveground and belowground plant parts representing species of *Fusarium* and *Verticillium*, be it with significantly higher numbers of transcripts in the latter group (T-test, $p < 0.001$). Insect and nematode transcripts were identified in some of the aboveground samples, and in all the belowground plant parts. Particularly in the belowground samples, the number and proportion of nematode hits is high, and were believed to originate from the soilborne bacteria-feeding *Caenorhabditis elegans*. Finally, on average a quarter of all contigs obtained from the aboveground plant parts did not produce significant hits, compared to approximately 12% in contigs from the belowground plant parts.

Discussion

Bioassays are important tools to study the interaction between pathogens and their hosts. In many pathosystems they are exploited on a differential set of genotypes to determine pathogen virulence or to group pathogens in races or pathotypes. For *S. endobioticum*, the tuber-based Spieckermann and Glynne-Lemmerzahl bioassays are widely used for potato resistance screening and pathotype identification. Infection of micro-propagated plantlets has been described as an alternative to the belowground tuber-based assays [18, 19], but these methods were not adopted by the research community. With the publication of the genomes of two *S. endobioticum* isolates, representing different pathotypes and geographical origins (Van de Vossenbergh & Warris *et al.*, unpublished), these alternatives are of increasing interest if these alternative bioassays result in lower microbiological contaminations, or when monitoring of the infection process is easier or more synchronous. Such

Chapter 4

high quality starting material can be instrumental in elucidating the interaction between host plant and pathogen on a molecular level. We propose that the aboveground assay described here is an improvement for two main reasons:

First, central in the plant-pathogen interaction are pathogen effector genes that facilitate infection and manipulation of the host plant. When the products of these effector genes are recognized by the plant resistance (*R*) genes, these effector genes are referred to as avirulence (*Avr*) genes. The differences in the *R* gene composition of the differential variety panel determines which *Avrs* are recognized and form the basis for pathotype grouping. Assays based on pathotype specific *Avrs* (specific DNA markers or Hypersensitivity Responses (HR)) could in time replace the time-consuming and labor-intensive bioassays currently used for pathotyping, as has been demonstrated for *Phytophthora infestans* [21]. One way to demonstrate that a pathogen *Avr* gene product is recognized by a host *R* gene is by *Agrobacterium tumefaciens* mediated transient transformation (agroinfiltration) [23]. In this system, plant leaves are infiltrated with *A. tumefaciens* suspensions carrying the gene of interest on a plant transformation vector. When expressed and recognized by plant receptors, defense responses are triggered culminating in a visual HR. Agroinfiltration is highly efficient in leaves but much more complicated in other organs like tubers and stems. To apply agroinfiltration as a verification method for *S. endobioticum* resistance, expression of plant *R* genes in aboveground plant parts is required.

Second, in pathosystems where a multitude of species are present next to the studied host and pathogen, it is difficult to determine which responses originate from the pathosystem and which are related to other microbial factors. Particularly on the molecular level of the host responses can be masked or obscured by the microbiome. This is even more the case for the soilborne pathogens as the rhizosphere has been shown to be much more complex than the phyllosphere in terms of microbial organisms [24] and could, in case of potato wart disease, promote accelerated decay of infected potato tissue. Hence, infected axillary buds could represent a more suitable material to study the interaction between *S. endobioticum* and its host on a molecular level.

Resistance in aboveground plant parts

Susceptible and resistant interactions with *S. endobioticum* in the EPPO differential set are identical in the aboveground bioassay and the tuber-based Spieckermann assays (table S2). None of the inoculated axillary buds of presumed resistant varieties produced symptoms. In

contrast when susceptible varieties were inoculated, wart-like symptoms were produced in 25 to 83% of the inoculations. This showed that the *S. endobioticum* resistance genes from the EPPO differential set plants are expressed in axillary buds similar to the resistance in belowground organs. In addition, the percentage of symptom expression of susceptible interactions was similar to those observed in the Spieckermann assay which ranged from 23% to 93% (table S2). Strongly varying percentages of symptom expression that can be visually observed on potato tubers infected with *S. endobioticum* under controlled conditions have been reported as early as 1922, and have been persistent over time complicating the analysis [18]. These inconsistencies have been attributed to a number of factors such as spore age and density, tuber and sprout age, potting medium or soil types, their porosity and texture, and watering regime [10]. For the aboveground bioassay, properly fixing the inoculum and maintaining the moist environment at the inoculation site were regarded to be crucial for successful infection.

In the tuber-based bioassays, HR can be observed on etiolated sprouts in resistant interactions in the form of ladder-shaped necrotic spots. In our aboveground assay, similar spots were observed on shoots of the resistant potato variety Kuba that could be a manifestation of HR (Fig S6). However, the frequency at which these “laddered HR” were observed was too low to serve as a hallmark for resistance scoring. When malformation was observed on the aboveground plant parts of susceptible plants, two main types of symptoms could be differentiated, i.e. leaf malformation (Fig. 2D) and shoot malformations (Fig. 2E), of which the leaf malformation was encountered most frequently. It appears that normal development is halted in one case (leaf malformation), where it is more or less continued in the other (shoot malformations). We hypothesize these differences could be caused by a disruption of apical dominance. The mechanism of apical dominance ensures the growth of only one meristem ensuring clear features, and is based on auxins [25]. Also, the formation of aerial roots was observed for several infected axillary buds (Fig. 2E). This aerial root formation could also indicate a change in the auxin transport. In the tomato “aerial roots” (aer) mutant, auxin (mis-)accumulation caused root formation along the stem [26].

After the end of the experiment (59 dpi), warts were discovered on belowground plant parts while these were not inoculated nor were resting spores added to the potting soil. Similarly, malformations of non-inoculated aboveground plant parts were observed. These ectopic proliferations of warted tissue occurred at low frequency and are believed to be caused by zoospores being transported down by with water dripping or flowing down the stem.

Chapter 4

Pathogen proliferation was confirmed using a species-specific real-time PCR. Some background reaction with the plant matrix was observed, as was previously noted in an interlaboratory comparison (ILC) study by Van de Vossen *et al.* [27]. In the ILC, Cq values obtained with healthy plant samples were used to establish a Cq cut-off value over 30 to eliminate false positive results. Typically, infected potato tubers produced Cq values ranging from 15 to 25 in the ILC, which is similar to the range that was found for the symptomatic aboveground plant parts. The distribution of Cq values obtained from symptomatic and asymptomatic plant parts allowed to distinguish between background reactions and pathogen proliferation, and were used to determine the pathogen proliferation of the plant parts scored as “unclear”. The majority of inoculated axillary buds initially scored as “unclear” could be identified as infected (i.e. proliferation of the pathogen). These were mainly attributed to pathotype 2(G1) and 6(O1) inoculations on potato variety Talent. In one specific case, a relatively low Cq value (25.2) was obtained for a sample taken from a resistant interaction (18(T1) on variety Belita. In this particular case, the inoculum was not removed until sampling which resulted in rotting of the inoculum. In this particular case the inoculum may have resulted in a lower Cq value, and not the proliferation of the pathogen.

In the EPPO diagnostic protocol, potato variety Talent is listed as resistant to pathotype 1(D1), but weakly resistant to pathotypes 2(G1) and 6(O1). With this variety, no full wart formation is observed in the Glynne-Lemmerzahl bioassay [12]. In our bioassay, slight malformations were observed, but not the severe malformations that were observed in susceptible interactions (Fig S7). The use of semi-resistant varieties in differential sets poses the risk of ambiguous scoring of reactions types. Currently the Glynne-Lemmerzahl bioassay is used most often for pathotyping and potato wart resistance determination. Five different reaction types are distinguished ranging from extremely resistant to extremely susceptible. However, differences in the scoring and interpretation exist in different laboratories working with the bioassay [17]. In this study we demonstrate that verification of proliferation using molecular test may aid the assessment based on phenotypic/visual methods providing a clearer distinction between the resistance or susceptibility of a variety for a given isolate of the pathogen. The use of semi-resistant varieties should also be avoided for the production of starting inoculum as this may result in population shift as a result of virulent genotype selection. For *S. endobioticum* it was demonstrated that a pathotype 1(D1) isolate produced a pathotype 6(O1) phenotype after two multiplications on a semi-resistant potato variety [28], and it was suggested both cognate (*Avr*) and non-cognate (*avr*) genotypes co-exist within the pathogen population. *S. endobioticum* growth on the semi-resistant variety

caused a selection against the *Avr* genotypes in the population, and the corresponding change in phenotype. In agronomical practices, plants that do not provide full resistance against certain isolates could also reshape the *S. endobioticum* population, selecting for genotypes with increased virulence. This could be the case for Talent and related genotypes in their interactions with pathotypes 2(G1) and 6(O1). We recommend additional research on this topic and caution in the use of semi-resistant varieties in practice as they could be stepping stones for the pathogen.

With the bioassay presented in this paper, the resistance and susceptibility responses of differentially resistant potato varieties to *S. endobioticum* pathotypes was equal in aboveground plant parts and the tuber-based bioassays. This suggests that *S. endobioticum* resistance genes are expressed both in etiolated belowground sprouts and green aboveground organs. Our findings open up possibilities to test candidate *Avr* genes using leaf-based *A. tumefaciens* mediated transient transformation assays.

Aboveground versus belowground

Almost all (90 - 96%) RNAseq data obtained from the aboveground bioassay belongs either to *S. endobioticum* or potato. This in contrast to the belowground derived samples in which only 18 to 43% of the data belonged to the pathogen or its host. In the belowground derived samples the majority of reads were from other soilborne (micro-)organisms. Even though in the aboveground plant parts transcripts of other prokaryotic and eukaryotic biological agents were identified, they were significantly less compared to the belowground derived RNAseq data. The absence of soilborne saprobic pathogens could be an advantage of using the aboveground bioassays as these organisms influence the plant and potentially also the pathogen transcriptome. The majority of reads in the aboveground samples represent potato transcripts, but still more than half of these samples show expression at TPM > 3 in more than 70% of all *S. endobioticum* genes. Resting spores extracted from fresh warts have relative high percentages of pathogen specific reads but have almost no potato derived reads. The proportion of potato derived reads in the aboveground samples enables identification of differentially expressed plant hormone pathways, such as auxin production and transport, as a result of pathogen proliferation.

In one of the samples (pathotype 18(T1), plant 10), the total number of reads mapping to *S. endobioticum* was very low, resulting in a bias in the normalized read coverage score giving an overestimation of the expression of *S. endobioticum* genes. It is known that normalization

Chapter 4

methods such as TPM and RPKM can introduce biases in such situations [29]. Improvement could be made by sampling also at a later stage when the symptoms have advanced further and more *S. endobioticum* biomass is accumulating. The biomass ratio between *S. endobioticum* and potato can be assessed by the use of the TaqMan PCR of *S. endobioticum* and the plant *cox1* that we applied in this study.

Studying the molecular plant-pathogen interaction of soilborne obligate biotrophic organisms represent a unique challenge as the microbiome potentially interacts with both host and pathogen. This can obscure the interaction of the pathosystem. In general, infected tissues from the aboveground bioassay were less prone to rot compared to tissue from the tuber-based bioassay. Using the aboveground bioassay, the interaction can be studied assessing both pathogen and host derived transcripts with significantly fewer and less diverse contaminants allowing studies of transcriptional responses of the plant to *S. endobioticum* infection and to distinguish reactions to other microbes.

Closing remarks

Bioassays for *S. endobioticum* pathotyping and resistance determination are time-consuming and labor-intensive. Also, they require expertise that has to be developed over many years to visually assess the reaction types. Even though they have their challenges in terms of standardization, harmonization and repeatability of test results obtained, they are fit for purpose of the laboratories working with these methods. Implementing new bioassays implies experts have to familiarize themselves with the art of the new method. Laboratories may not have the time or expertise to implement new methods, and to test different methods in parallel. This possibly explains why the alternative bioassay described by Hampson [18, 19] was not adopted despite advantages such as increased sensitivity, higher success rates, year-round application. Nevertheless, bioassays such as the one described by Hampson, and the one from our study could have a potential use in conventional breeding programs. Their adoption by the community will depend on advantages that are not provided by other assays. We see clear advantages in breeding as our assay can be directly used on potato plants. Currently it takes at least two years to obtain sufficient tubers of a new variety for *S. endobioticum* pathotype resistance screening. With our aboveground bioassay this could be reduced to one year as clonally propagated cuttings from seedlings can be used in this assay. Also, other Solanaceous species could represent new potato wart resistance sources. Non-tuber bearing species or species with poor tuber formation cannot be tested with the

current tuber-based bioassays. With our assay these now can be screened for potato wart resistance as no tubers are required, unlocking the potential of these species for new potato wart resistance.

Material and Methods

Fungal materials

Fresh warts of pathotypes used as inoculum in the aboveground bioassay were obtained from pot tests with the susceptible potato variety Deodara. Several 16 x 16 cm pots filled with potting soil amended with composts of pathotypes 1(D1) isolate HLB 11-11 (~1,200 spores/g), 2(G1) isolate HLB 03-08 (~1,500 sp/g), 6(O1) isolate HLB 10-11 (~600 sp/g), and 18(T1) isolate HLB 05-08 (~1,800 sp/g) in a 1:10 ratio for compost:potting soil. These isolates are also used in national resistance testing programs from the Netherlands, and reactions on differential potato varieties Deodara, Producent, Saphir, Miriam, Delcora and Belita were consistent with those described in EPPO standard PM7/28(1) [30] (table S2). Pots were kept separated per pathotype in a quarantine glasshouse in which temperatures were kept below 25 °C with a light regime of 16 h light /8 h darkness. Shoots were cut back regularly to keep the plants below approximately 50 cm height until harvesting the warts (typically 8 to 12 weeks after inoculation).

Aboveground bioassay

Ten Deodara and eight Tomensa plants, both susceptible to pathotype 1(D1) and eleven Kuba plants resistant to pathotype 1(D1) were tested with fourteen different treatment combinations (including mock treatments) to optimize the procedure. Optimization parameters included inoculum type (fresh wart cuttings and compost), inoculation sites (axillary buds and leaves), fixing the inoculum (household plastic and sticky tape) and incubation conditions (in the greenhouse with temperatures kept below 25 °C, inoculated plants were placed in a closed tent of transparent foil to ensure maximum humidity, as described below; or in a climate chamber maintained at ~16 °C and 80% RH, without illumination). The overview of varieties and parameters used (Fig. 1) was created with the UpSet visualization tool [31].

The following procedure is the result of optimization experiments performed in early spring 2016. Potato plants were grown from tubers in 16 x 16 cm pots with potting soil containing Osmocote^(R) Exact slow release grains for nutrients (ICL Specialty Fertilizers, the Netherlands)

Chapter 4

in March to April under greenhouse conditions mentioned under “fungal materials”. When plants were 20 to 40 cm tall, typically after three to four weeks, they were inoculated with small (5 to 10 mm) cuttings of fresh warts. Precautions were taken to ensure that the inoculum was as “clean” as possible. First, only young fresh warts, that not yet showed any signs of decay, were taken from the Spieckermann assay, which were thoroughly rinsed with tap water to remove any remaining soil. Finally, visually clean cuttings were taken from the rinsed fresh warts and used as inoculum. Mock treated plants were included for each of the varieties tested.

Axillary buds were moistened with a fine mist after which wart cuttings were placed with the uncut surface facing towards early developing bud ensuring that they were in close proximity to the plant tissue. Inoculated buds were moistened again and wrapped in household plastic to hold them in place and to maintain a moist environment to facilitate infection. Inoculated buds were marked with an unique label, and after inoculation, plants were placed on a greenhouse table with moisture retaining cloth under a bespoke 100 cm x 96 cm x 210 cm (height x depth x width) frame fitted with 0.5 mm transparent foil (Fig. S1). The cloth and pots were watered intensively, and the plastic was lowered over the frame to create a closed environment with RH up to 100%. The plants remained untouched for a week after which the inoculum was removed and apical shoot and non-inoculated axillary shoots were cut to stimulate growth of the inoculated axillary buds. Plants were watered regularly to maintain a 100% RH environment under the plastic, and were regularly inspected for signs of malformation of the plant tissue.

After four weeks the first symptoms became apparent, and plants were kept until 8 weeks after inoculation. Selected symptomatic and asymptomatic buds for all variety pathotype combinations, including mock treated plants, were sampled 30 to 59 days post inoculation (dpi) for microscopy and molecular testing. Plant were scored visually as symptomatic when they showed clear signs of malformation, or asymptomatic when no malformation was observed. Slight malformations that could be the result of proliferation of the pathogen or stress caused by the conditions in the greenhouse (e.g. high humidity and axillary buds being wrapped in household plastic) were scored as “unclear phenotype”.

In the following year (2017), varieties from EPPO standard PM7/28 (2) [12] Deodara, Producent, Talent, Saphir and Belita (at least three plants per variety) were used in combination with *S. endobioticum* pathotypes 1(D1), 2(G1), 6(O1), and 18(T1). Buds showing

early malformations for each pathotype tested on variety Deodara were sampled 24 dpi for RNAseq analysis. These samples were flash frozen with liquid nitrogen and stored at -80 °C until RNA extraction.

DNA extraction and *S. endobioticum* quantification

DNA was extracted from approximately 50 to 200 mg of plant material using the Wizard Magnetic DNA Purification System for Food (Promega, WI, USA) following manufacturer's instructions. Plant material was homogenized at room temperature in 400 µL Lysis Buffer A and 200 µL Lysis Buffer B with three 3.2 mm stainless steel beads using a Precellys Evolution (Bertin Instruments, France) at 5,000 bpm for 45 sec (3 times 15 sec with 10 sec pause in between) prior to DNA extraction. *S. endobioticum* DNA was quantified using a *S. endobioticum* specific real-time PCR with plant *cox1* internal control [32]. Boxplots, biplots and statistical analyses were performed with GenStat v19.1 (VSN International, Hemel Hempstead, United Kingdom).

RNA extraction and Illumina sequencing

Plant material was homogenised dry in stainless steel microvials (BioSpec, OK, USA) with 3.2 mm beads and cooled with liquid nitrogen, using a Precellys Evolution (Bertin Instruments, France) homogenizer at 8,000 rpm for 15 sec. Stainless steel microvials were cooled again with liquid nitrogen and homogenized at 8,000 rpm for another 15 sec prior to RNA extraction. Homogenised plant material was lysed in RLC buffer with β-mercaptoethanol. Total RNA was extracted using the RNeasy plant mini kit (Qiagen, the Netherlands) following manufacturer's instructions. Extracted RNA was DNase treated using The DNA-free™ DNA removal Kit (AM190, Thermo Fisher Scientific, MA, USA) and quantified with the Quant-iT RiboGreen RNA (Thermo Fisher, MA, USA) reagents on an Infinite 200 PRO (Tecan Life Sciences, Switzerland). Poly-A enriched Illumina RNAseq libraries were prepared with TruSeq RNA Library Prep Kit v2 (Illumina, CA, USA) and sequenced on a HiSeq2500 (Illumina, CA, USA).

Expression of *S. endobioticum* genes and assessment of represented taxa

RNAseq data obtained from aboveground plant parts were compared to RNAseq data derived from belowground infected plant parts. The belowground derived RNAseq data were taken from Van de Vossenbergh and Warris *et al.* (unpublished) (SRA: SRR8074754 and SRR8074755), and were obtained from resting spores which were extracted and purified from

Chapter 4

fresh wart tissue as described by Bonants *et al.* [33]. Libraries for the belowground RNAseq datasets were prepared with either a polyA enrichment or ribosomal plant RNA depletion (table S1). Reads were mapped to the coding sequences and genomes of pathotype 1(D1) *S. endobioticum* (MB42) (NCBI accession: QEAN00000000, version 1.0) in CLC genomics workbench using the RNA-seq analysis tool (length fraction = 0.8, similarity fraction = 0.9, counting paired reads as two), and expression levels were reported as TPM (Transcripts Per Kilobase Million) per gene. Unmapped reads were collected and mapped to the *Solanum tuberosum* subspecies *phureja* (DM v3.0) genome [34] using the same tool and settings, after which reads that did not map to either *S. endobioticum* MB42 v1.0 and *S. tuberosum* DM v3.0 were assembled *de novo* in CLC (word size: automatic, bubble size: automatic, scaffolding: on). Scaffolds were updated using a read mapping approach (similarity fraction = 0.8, length fraction = 0.9). Contigs smaller than 200 bp and those still producing significant megablast hits (e-value < 1e-50) with the *S. endobioticum* LEV6574 v1.0 (NCBI accession: QEAM00000000, version 1.0) or DMv3.0 were discarded, and the remaining sequences were used as input for a taxonomy identification pipeline. This pipeline consisted of a blastn, which is part of the BLAST+ v2.6.0 package [35], analysis in a local NCBI nr/nt database (Download date: 24 September 2018). Sequences producing no hit with blastn (e-value < 1e-50) were selected and were further analysed using Diamond [36]. Taxa represented in the blastn and Diamond output were visualized using Krona [37].

Acknowledgements

We would like to thank Richard Strijker (HLB) for maintaining plants. Richard G.F. Visser and Herman van Eck (Wageningen University and Research, the Netherlands) are thanked for their support and advice on the project. We thank Averis seeds BV, Böhm-Nordkartoffel Agrarproduktion GmbH & Co. OHG, Danespo, HZPC Holland BV, C. Meijer BV, SaKa Pflanzenzucht GmbH & Co. KG and Teagasc for their involvement in the public-private partnership that was formed to fund this study. Thank you Donna S. Smith (Canadian Food Inspection Agency), for critically reviewing the manuscript before submission.

Funding

This research (1406-056) is financially supported by the Dutch Topsector Horticulture & Starting Materials (<http://topsectortu.nl/nl/integrated-genomics-and-effectoromics-impulse-potato-wart-resistance-management-and-breeding>).

Author contributions

BV, JV, TL produced the research outline and wrote the manuscript. BV, JV, MB, MG, and TL performed laboratory and greenhouse experiments. LG provided bioinformatics support. All authors read, commented and approved the final manuscript.

Author affiliations

BV, JV, MG, TL: Wageningen University and Research, Droevendaalsesteeg 1, 6708 PB, Wageningen, the Netherlands; BV, LG: Dutch National Plant Protection Organization, Geertjesweg 15, 6706 EA, Wageningen, the Netherlands; MB: HLB B.V., Kampsweg 27, 9418 PD, the Netherlands.

References

1. Curtis, K.M., IX.— The life-history and cytology of *synchytrium endobioticum* (schilb.), perc., the cause of wart disease in potato. Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character, 1921. 210(372-381): p. 409.
2. Hampson, M.C., History, biology and control of potato wart disease in Canada. Canadian Journal of Plant Pathology, 1993. 15(4): p. 223-244.
3. Przetakiewicz, J., The Viability of Winter Sporangia of *Synchytrium endobioticum* (Schilb.) Perc. from Poland. American Journal of Potato Research, 2015. 92(6): p. 704-708.
4. Schilberszky, K., Ein neuer Schorfparasit der Kartoffelknollen. Ber. Deut. Botan. Ges., 1896. 14: p. 36-37.
5. Weiss, F., The conditions of infection in potato wart. American Journal of Botany, 1925. 12(7): p. 413-443.
6. Moore, W.C., The Breakdown of Immunity from Potato Wart Disease. Outlook on Agriculture, 1957. 1(6): p. 240-243.
7. Baayen, R.P., *et al.*, History of potato wart disease in Europe - a proposal for harmonisation in defining pathotypes. European Journal of Plant Pathology, 2006. 116(1): p. 21-31.
8. Çakır, E., *et al.*, Identification of pathotypes of *Synchytrium endobioticum* found in infested fields in Turkey. EPPO Bulletin, 2009. 39(2): p. 175-178.
9. Przetakiewicz, J., First Report of New Pathotype 39(P1) of *Synchytrium endobioticum* Causing Potato Wart Disease in Poland. Plant Disease, 2015. 99(2): p. 285-285.
10. Hampson, M.C., Control of potato wart disease through the application of chemical soil treatments: a historical review of early studies (1909-1928). EPPO Bulletin, 1988. 18(1): p. 153-161.
11. Smith, I.M., *et al.*, Quarantine Pests for Europe - Data Sheets on quarantine pests for the European Union and for the European and Mediterranean Plant Protection Organization. 2 ed. 1997: CAB International, Wallingford, UK.
12. OEPP/EPPO, PM 7/28 (2) *Synchytrium endobioticum*. EPPO Bulletin, 2017. 47(3): p. 420-440.
13. Spieckermann, A.K., P., Testing potatoes for wart resistance. Deutsche Landwirtschaftliche Presse, 1924. 51: p. 114-115.

Chapter 4

14. Glynne, M.D., Infection experiments with wart disease of potatoes. *Synchytrium endobioticum* (Schilb.) Perc. Annals of Applied Biology, 1925. 12(1): p. 34-60.
15. Lemmerzähl, J., Neues vereinfachtes Infektionsverfahren zur Prüfung von Kartoffelsorten auf Krebsfestigkeit. Der Züchter, 1930. 2: p. 288-297.
16. Noble, M.G., M.D., Wart disease of potatoes. FAO Plant. Protection Bulletin, 1970. 18: p. 125-135.
17. Flath, K., et al., Interlaboratory tests for resistance to *Synchytrium endobioticum* in potato by the Glynne-Lemmerzähl method. EPPO Bulletin, 2014. 44(3): p. 510-517.
18. Hampson, M.C., A bioassay for *Synchytrium endobioticum* using micropropagated potato plantlets. Canadian Journal of Plant Pathology, 1992. 14(4): p. 289-292.
19. Hampson, M.C., J.W. Coombes, and S.C. Debnath, Dual Culture of *Solanum tuberosum* and *Synchytrium endobioticum* (Pathotype 2). Mycologia, 1997. 89(5): p. 772-776.
20. Ullrich, J., Morphologische und anatomische Untersuchungen über die durch *Synchytrium endobioticum* (Schilb.) Perc. bei der Kartoffel ausgelöste Gallbildung. Journal of Phytopathology, 1962. 44(1): p. 57-75.
21. Vleeshouwers, V.G.A.A., et al., Understanding and Exploiting Late Blight Resistance in the Age of Effectors. Annual Review of Phytopathology, 2011. 49(1): p. 507-531.
22. Millett, B.P., et al., Potato Tuber Blight Resistance Phenotypes Correlate with RB Transgene Transcript Levels in an Age-Dependent Manner. Phytopathology, 2015. 105(8): p. 1131-6.
23. Van der Hoorn, R.A., et al., Agroinfiltration is a versatile tool that facilitates comparative analyses of *Avr9/Cf-9*-induced and *Avr4/Cf-4*-induced necrosis. Mol Plant Microbe Interact, 2000. 13(4): p. 439-46.
24. Lindow, S.E. and M.T. Brandl, Microbiology of the phyllosphere. Applied and environmental microbiology, 2003. 69(4): p. 1875-1883.
25. Kebrom, T.H., A Growing Stem Inhibits Bud Outgrowth - The Overlooked Theory of Apical Dominance. Frontiers in plant science, 2017. 8: p. 1874-1874.
26. Mignolli, F., et al., Differential auxin transport and accumulation in the stem base lead to profuse adventitious root primordia formation in the aerial roots (aer) mutant of tomato (*Solanum lycopersicum* L.). J Plant Physiol, 2017. 213: p. 55-65.
27. van de Vossenbergh, B.T.L.H., et al., Euphresco Sendo: An international laboratory comparison study of molecular tests for *Synchytrium endobioticum* detection and identification. European Journal of Plant Pathology, 2018. 151(3): p. 757-766.
28. van de Vossenbergh, B.T.L.H., et al., The linear mitochondrial genome of the quarantine chytrid *Synchytrium endobioticum*; insights into the evolution and recent history of an obligate biotrophic plant pathogen. BMC Evolutionary Biology, 2018. 18(1): p. 136.
29. Dillies, M.-A., et al., A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. Briefings in Bioinformatics, 2013. 14(6): p. 671-683.
30. OEPP/EPPO, PM 7/28 (1) *Synchytrium endobioticum*. EPPO Bulletin, 2004. 34(2): p. 213-218.
31. Lex, A., et al., UpSet: Visualization of Intersecting Sets. IEEE Transactions on Visualization and Computer Graphics, 2014. 20(12): p. 1983-1992.
32. van Gent-Pelzer, M.P.E., M. Krijger, and P.J.M. Bonants, Improved real-time PCR assay for detection of the quarantine potato pathogen, *Synchytrium endobioticum*, in zonal centrifuge extracts from soil and in plants. European Journal of Plant Pathology, 2010. 126(1): p. 129-133.

33. Bonants, P.J.M., *et al.*, A real-time TaqMan PCR assay to discriminate between pathotype 1 (D1) and non-pathotype 1 (D1) isolates of *Synchytrium endobioticum*. *European Journal of Plant Pathology*, 2015. 143(3): p. 495-506.
34. The Potato Genome Sequencing, C., *et al.*, Genome sequence and analysis of the tuber crop potato. *Nature*, 2011. 475: p. 189.
35. Camacho, C., *et al.*, BLAST+: architecture and applications. *BMC Bioinformatics*, 2009. 10: p. 421.
36. Buchfink, B., C. Xie, and D.H. Huson, Fast and sensitive protein alignment using DIAMOND. *Nat Methods*, 2015. 12(1): p. 59-60.
37. Ondov, B.D., N.H. Bergman, and A.M. Phillippy, Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics*, 2011. 12(1): p. 385.

Supplementary material

Supplementary figures

- Fig. S1 Selected steps of the inoculation procedure in the aboveground bioassay
- Fig. S2 Symptomatic and asymptomatic interactions by *S. endobioticum* pathotype 1(D1) on susceptible and resistant cultivars as observed in the aboveground bioassay
- Fig. S3 Distribution of Cq values obtained for inoculated plant parts tested in the optimization experiments
- Fig. S4 Microscopy on symptomatic aboveground plant parts
- Fig. S5 Wart formation on belowground plant parts
- Fig. S6 Possible manifestations of a Hypersensitive Response (HR)
- Fig. S7 Reaction types on the slightly susceptible variety Talent

Selected supplementary tables

- Table S1 Overview of RNAseq data used in this study
- Table S2 Spieckermann results for composts used in this study to generate warts on *Deodara* which are used as inoculum in the aboveground bioassay.
- Table S4 Blastn and DIAMOND based identification of *de novo* assembled contigs generated with reads not mapping to *S. endobioticum* or potato.

Supplementary tables S3 was not included in this thesis on account of its size. This file is available online.

Figure S1. Selected steps of the inoculation procedure in the aboveground bioassay. **A.** Fresh warts obtained from pot tests were used as inoculum source. Potato warts prior to washing to remove remaining potting soil are shown. **B.** Potato plants were grown from tubers in the quarantine greenhouse until they were 20 to 40 cm tall. **C.** Axillary buds were sprayed with water to create a moist environment, and small cuttings of the washed wart tissue were placed in axillary buds (arrow) with the malformed tissue facing towards the developing plant tissue. **D.** Inoculated axillary buds were sprayed again with water and wrapped in household plastic (arrow) to hold the inoculum in place, ensuring close proximity to the developing plant tissue, and to maintain a moist environment to facilitate infection. **E.** Inoculated plants were placed on a greenhouse table with moisture retaining cloth under a bespoke 100 cm x 96 cm x 210 cm (height x depth x width) frame fitted with 0.5 mm transparent plastic. The greenhouse table and potting soil of the inoculated plants was extensively watered before lowering the transparent foil allowing incubation of the inoculated plants for one week. Conditions in the quarantine glasshouse were kept below 25 °C with a light regime of 16h light/8h darkness.





Figure S2. Symptomatic and asymptomatic interactions by *S. endobioticum* pathotype 1(D1) on susceptible (S) and resistant (R) cultivars as observed in the aboveground bioassay. Plants with symptomatic tissues were the result of wart inoculated axillary buds with the inoculum held in place with household plastic unless stated otherwise. **A.** Deodara (S) plant 2 bud 1, 29 days post inoculation (dpi), malformation on leaves and shoot, Cq-value: 13.8; **B.** Deodara (S) plant 2 bud 2, 29 dpi, malformation on leaves and shoot, Cq-value: 17.5; **C.** Deodara (S) plant 1 bud 1, 29 dpi, severe malformation on leaves and shoot with rotting of the malformed tissue, Cq-value: 13.5; **D.** Tomensa (S) plant 3, bud 1, inoculum held in place with sticky tape, 29 dpi, malformation of the developing plant tissue with abortion of further growth or development, Cq-value: 14.6; **E.** Tomensa (S) plant 5, bud 2, 29 dpi, malformation of the shoot, Cq-value: 14.7; **F.** Deodara (S) plant 14, bud 1, 30 dpi, slight malformation of the developing tissue, scored as “unclear”, Cq-value: 24.2; **G.** Deodara (S) plant 7, bud 2, 45 dpi, malformation of leaves, scored as “unclear”, Cq-value: 21.6; **H.** Tomensa (S) plant 3, bud 2, mock treated plant part, 29 dpi, no malformation observed, Cq-value: 28.2; **I.** Kuba (R) plant 3, bud 1, 29 dpi, possible hypersensitive response (dying of the developing tissue) as a result of inoculation with the pathogen, Cq-value: 34.7.

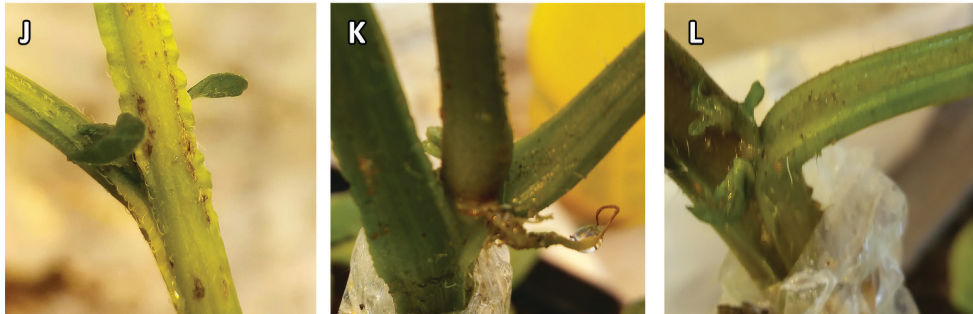


Figure S2 (continued). J. Kuba (R) plant 3, bud 2, 29 dpi, no malformation observed; K. Deodara (S) plant 10, bud 2, 30 dpi, scored as “unclear” based on thickening of base of the shoot and the formation of an aerial root, Cq-value: 32.0, no increase in *S. endobioticum* biomass and therefore considered not infected; L. Deodara (S) plant 10, bud 1, 30 dpi, scored as “unclear”, Cq-value: 30.3, not infected.

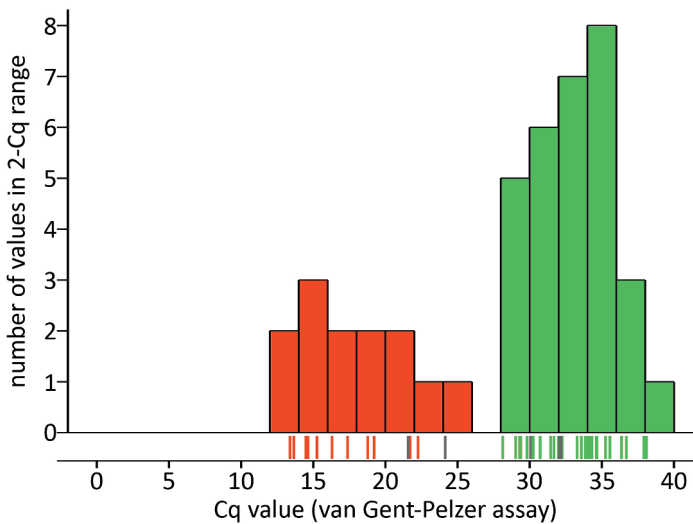


Figure S3. Distribution of Cq values obtained for inoculated plant parts tested in the optimization experiments. Cq values are obtained with the *S. endobioticum* specific TaqMan PCR described by Van Gent-Pelzer *et al.* (2010). Marks on the x-axis indicate the individual Cq values, where values obtained from symptomatic material are shown in red, Cq values from asymptomatic material are shown in green, and Cq values from material with no clear phenotype are shown in grey. Bars in the histogram of infected plant parts are colored red, whereas healthy plant are colored green.

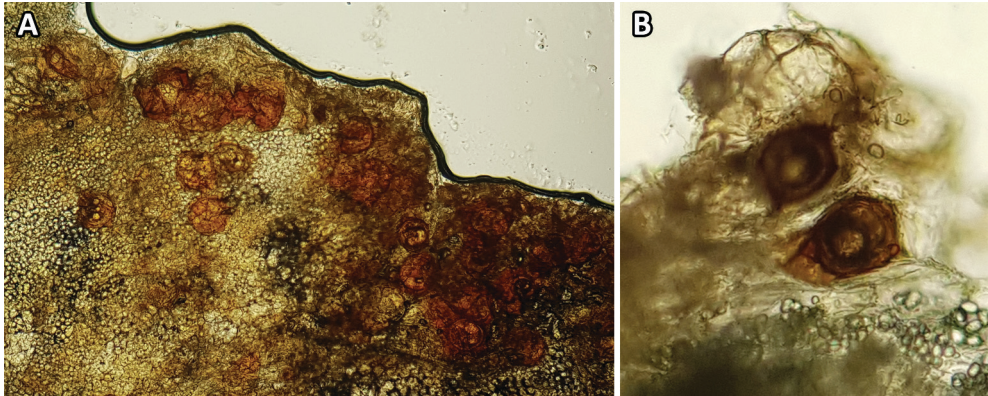


Figure S4 Microscopy on symptomatic aboveground plant parts. A+B. Resting spores in malformed leaf tissue of Producent (plant 3, bud 2) inoculated with pathotype 2(G1) as observed under a stereo microscope

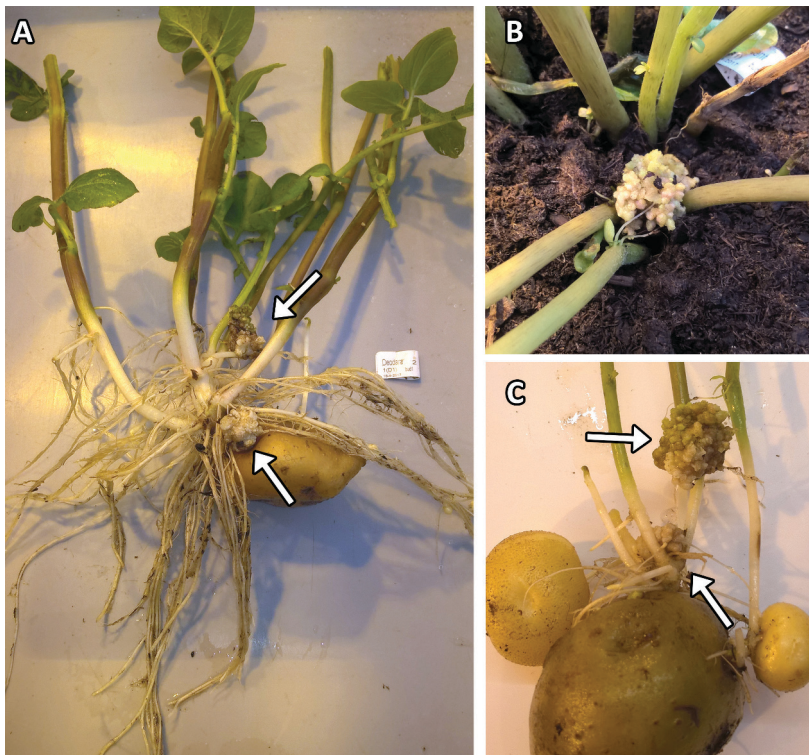


Figure S5. Wart formation on belowground plant parts. Development of symptoms on belowground plant parts which are the result of transport of zoospores, to belowground plant parts with irrigation water. **A.** Wart formation on Deodara plant 2 inoculated with pathotype 1(D1). **B.** Producent plant 2 inoculated with pathotype 2(G1) with wart formation visible breaking the surface of the potting soil, and **C.** on the tuber and stem after removal of potting soil. Arrows indicate the position of malformed tissues, and images were captured 59 dpi.



Figure S6. Possible manifestations of a Hypersensitive Response (HR). A+B. Ladder-shaped necrotic spots on shoots (arrow) of the resistant potato variety Kuba inoculated with *S. endobioticum* pathotype 1(D1) that could be a HR as is known to occur on etiolated sprouts.

A bioassay targeting aboveground plant parts

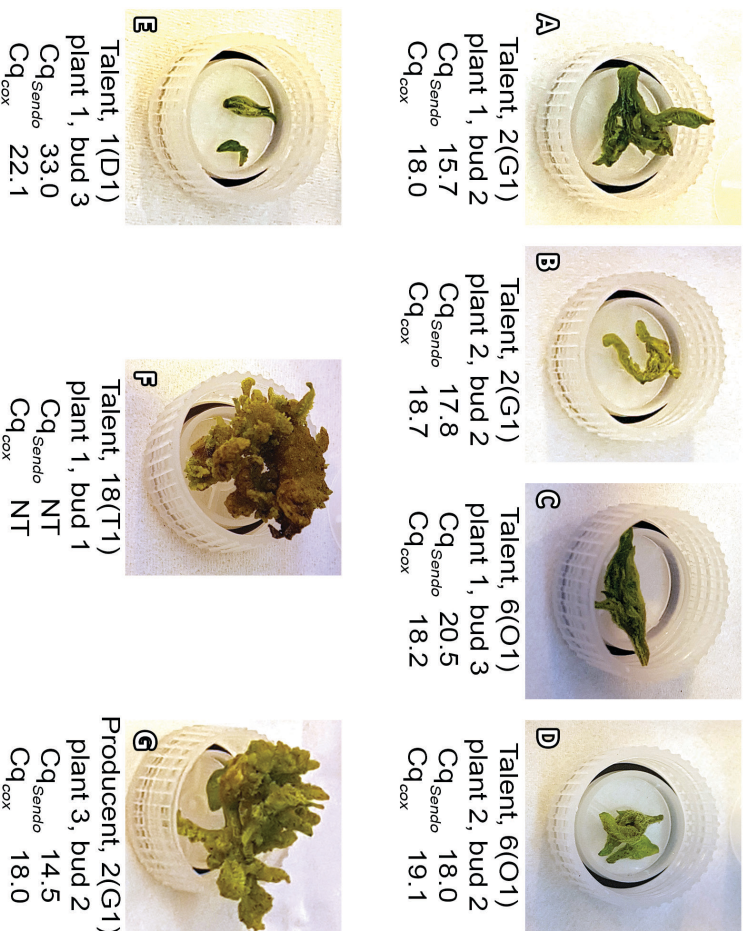


Figure S7. Reaction types on the slightly susceptible variety Talent. Selected reaction types of aboveground inoculated plant parts of potato cultivar Talent to different pathotypes. Talent is slightly susceptible to pathotypes 2(G1) and 6(O1) in tuber based bioassays. Also in the aboveground bioassay, axillary buds of Talent inoculated with pathotypes 2(G1) and 6(O1) were scored as “unclear” (A-D) as they did not produce the typical malformation associated with pathogen proliferation as was the case with Talent buds inoculated with pathotype 18(T1) (F). Pathogen specific Cq values consistent with pathogen proliferation were found in the samples scored as “unclear”. This was not the case for the asymptomatic Talent material sampled from a pathotype 1(D1) inoculated bud (E). The symptomatic material from talent (F) was not tested by qPCR. Similar symptomatic material shown in G. shows strong proliferation of the pathogen in the qPCR (Cq value 14.5). Tube caps measure 2 cm in diameter and are indicative for the sample size.

Table S1. Overview of RNAseq data used in this study

<i>S. endobioticum</i> pathotype	Isolate	Sample accession	Origin	Sub-Sample	SRA accession	Material	Treatments
1(D1)	MB42, HLB11-11	ERS3016529	Langenboom, the Netherlands	Sendo_plant_1	ERX3121142	Deodara axillary bud, 24dpi	DNase treatment, PolyA enriched libraries
				Sendo_plant_2	ERX3121143	aboveground bioassay	riched libraries
				Sendo_plant_3	ERX3121144		
2(G1)	HLB 03-08	ERS3016530	Germany	Sendo_plant_4	ERX3121145	Deodara axillary bud, 24dpi	DNase treatment, PolyA enriched libraries
				Sendo_plant_5	ERX3121146	aboveground bioassay	riched libraries
				Sendo_plant_6	ERX3121147		
6(O1)	HLB 10-11	ERS3016531	Smilde, the Netherlands	Sendo_plant_7	ERX3121148	Deodara axillary bud, 24dpi	DNase treatment, PolyA enriched libraries
				Sendo_plant_8	ERX3121149	aboveground bioassay	riched libraries
				Sendo_plant_9	ERX3121150		
18(T1)	HLB 05-08	ERS3016532	Germany	Sendo_plant_10	ERX3121151	Deodara axillary bud, 24dpi	DNase treatment, PolyA enriched libraries
				Sendo_plant_11	ERX3121152	aboveground bioassay	riched libraries
				Sendo_plant_12	ERX3121153		
1(D1)	MB42	ERS2166656	Langenboom, the Netherlands	MB42+polyA	SRR8074754	Resting spores extracted from warts on Deodara tubers	DNase treatment, PolyA enriched libraries
				MB42-polyA	SRR8074755	Resting spores extracted from warts on Deodara tubers	no DNase treatment, PolyA enriched libraries
				MB42+ribo-	ERX3121154	Resting spores extracted from warts on Deodara tubers	DNase treatment, plant rRNA depleted libraries
				MB42-ribo-	ERX3121155	Resting spores extracted from warts on Deodara tubers	no DNase treatment, plant rRNA depleted libraries

Table S2. Spieckermann results for composts used in this study to generate warts on Deodara which are used as inoculum in the aboveground bioassay. For each cultivar-pathotype combination the expected interaction is given, i.e. resistant (R) or susceptible (S). The percentage of escapes for S interactions is listed. The scoring is based on the description of the Spieckermann bioassay as published in EPPO standard PM7/28 *Synchytrium endobioticum* (1), EPPO Bulletin (2004) 34, 213 - 218

Pathotype	Isolate	Spores per gram soil	Variety	R or S interaction	Escapes (%)	Reactions										
						1	2	3	4	5						
1(D1)	HLB P1 (D1) 11-11	1200	Deodara	S	13%	7			2	2	2			5	35	
			Producent	R	-	53	1									
2(G1)	HLB P2(G1) 03-08	1500	Deodara	S	21%	11			3		4	2	1	5	26	
			Producent	S	33%	15	2	1	2	2	3	7	0	6	16	
			Saphir	S	50%	22	1	2	4	5	5	7	2	2		
			Delcora	R	-	48	2	1								
			Miriam	R	-	46	2	2								
			Belita	R	-	49	2	1								
6(O1)	HLB P6(O1) 10-11	800	Deodara	S	15%	8				2		2		2	40	
			Miriam	R	-	50	2	2								
			Saphir	R	-	52	2									
			Producent	S	28%	15			3	1	2	2	4	2	25	
			Delcora	R	-	48	2	4								
			Belita	R	-	54										
18(T1)	HLB P18(T1) 05-08	1800	Deodara	S	7%	4			1		2	2		5	40	
			Belita	R	-	53										
			Saphir	R	-	50	1									
			Miriam	S	31%	16	1		1		2		4	30		
			Producent	S	20%	9	2				3		3	10	30	
	Delcora	S	67%	32	2	2	2	1	2	3	3	1	3	5		

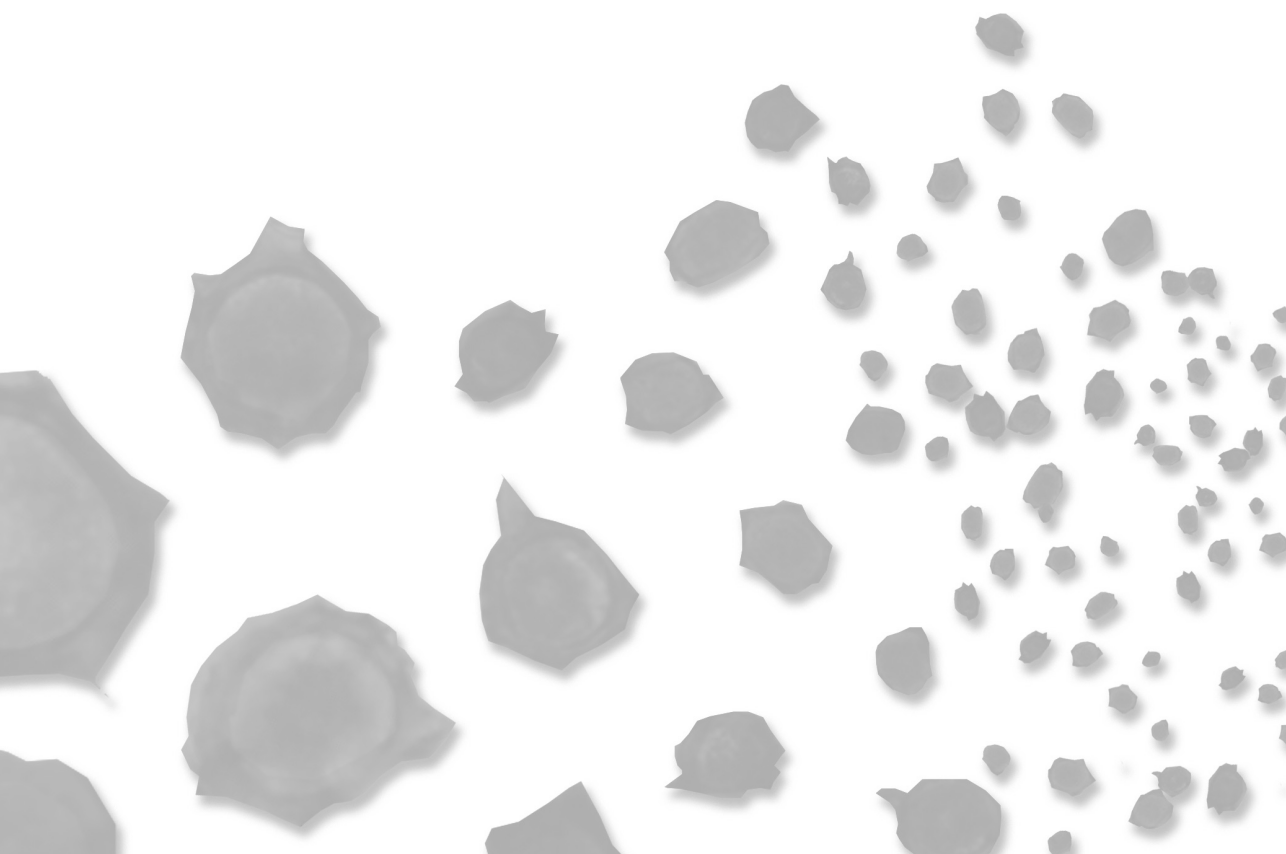
Table S4. Blastn and DIAMOND based identification of *de novo* assembled contigs generated with reads not mapping to *S. endobioticum* or potato.

	Plant 1	Plant 2	Plant 3	Plant 4	Plant 5	Plant 6	Plant 7	Plant 8
Eukaryotes								
Metazoa	278	332	122	6	1	71	0	0
Nematoda								
Arthropoda	22	7	3	4,102	8	3	7	549
other/unspecified	81	99	38	185	11	43	36	83
Intramacronucleata (ciliates)	372	340	97	2,209	80	54	94	212
Fungi								
Ascomycota	188	217	11	22	12	15	22	27
Basidiomycota	5	10	3	0	0	3	11	76
Chytridiomycota	3	0	0	1	0	0	0	3
other/unspecified	3	9	6	7	13	11	13	22
Viridiplantae	449	485	460	512	461	572	574	485
other/unspecified	450	357	170	523	313	228	443	1,013
Bacteria								
	32	44	24	66	53	49	69	97
Virus								
	0	0	1	7	0	1	3	6
Unspecified/No hits	751	594	297	2,045	287	334	622	1,185
Total <i>de novo</i> assembled contigs	2,928	2,762	1,458	10,156	1,453	1,720	2,208	3,952
Total contigs in blastn and DIAMOND analysis	2,634	2,494	1,232	9,685	1,239	1,384	1,894	3,758
% contigs of assembled contigs in analysis	90%	90%	84%	95%	85%	80%	86%	95%
% plant hits	17.0%	19.4%	37.3%	5.3%	37.2%	41.3%	30.3%	12.9%
% no hits	28.5%	23.8%	24.1%	21.1%	23.2%	24.1%	32.8%	31.5%

A bioassay targeting aboveground plant parts

Table S4 (continued). Blastn and DIAMOND based identification of *de novo* assembled contigs generated with reads not mapping to *S. endobioticum* or potato.

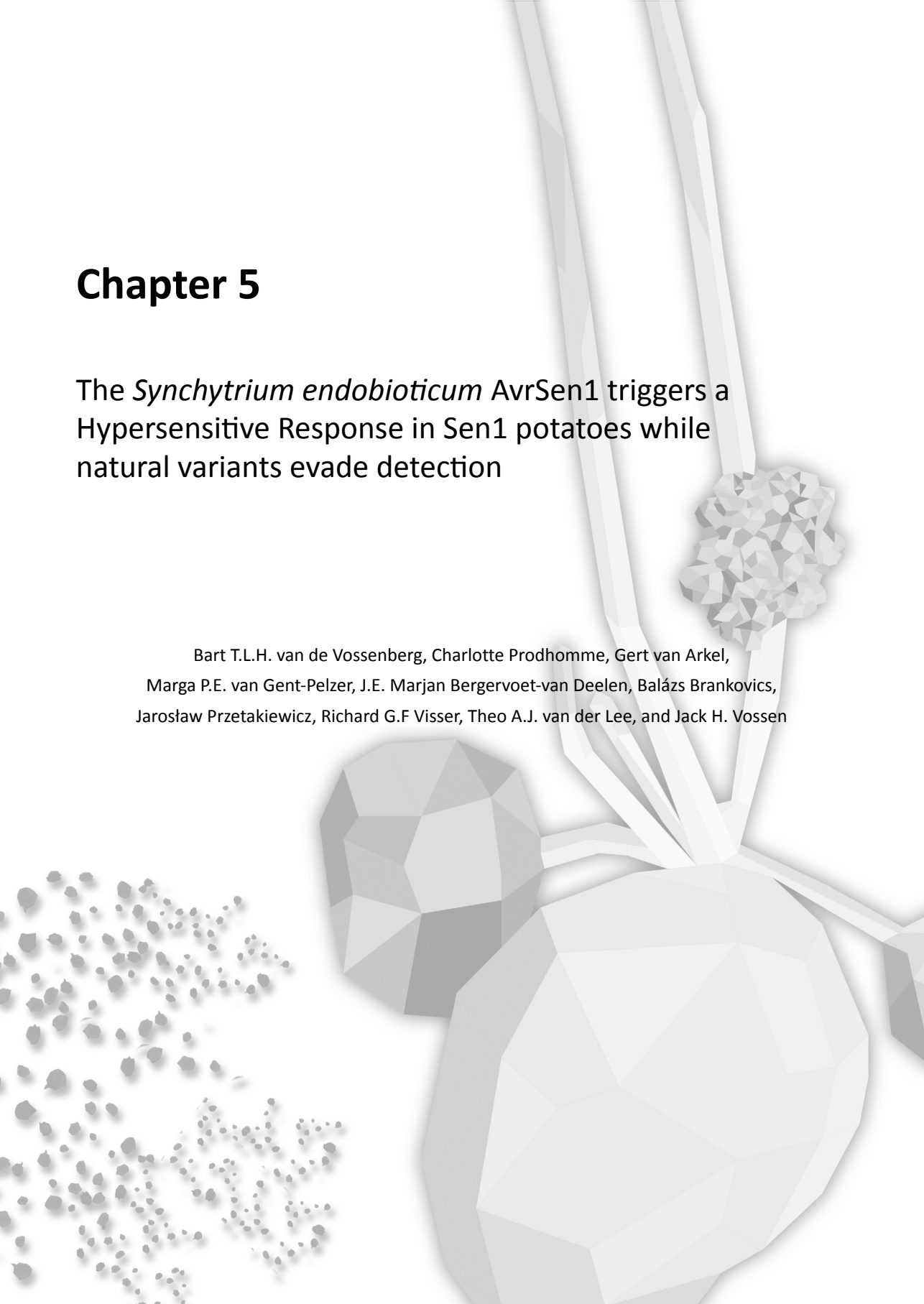
	Plant 9	Plant 10	Plant 11	Plant 12	DNase + polyA	DNase - polyA	DNase + rRNA -	DNase - rRNA -
Eukaryotes								
Metazoa								
Nematoda	0	6	0	0	6,290	20,317	19,239	20,160
Arthropoda	5	11	8	0	120	1,092	519	965
other/unspecified	32	36	33	6	365	1,564	1,452	1,446
Intramacroucleata (ciliates)	19	437	290	0	73	1,371	559	729
Fungi								
Ascomycota	4	81	291	500	604	3,021	7,268	7,512
Basidiomycota	3	9	3	6	28	189	247	300
Chytridiomycota	1	1	2	4	23	257	331	500
other/unspecified	14	25	16	2	184	1,085	1,491	2,050
Viridiplantae	478	606	580	534	0	78	255	302
other/unspecified	453	1,159	643	139	566	3,249	2,000	2,336
Bacteria	53	109	51	37	213	686	42,317	49,282
Virus	0	5	3	9	1	12	259	245
Unspecified/No hits	398	1,059	567	208	1,500	7,007	8,690	10,237
Total <i>de novo</i> assembled contigs	1,784	3,784	2752	1,919	10,328	40,935	85,248	96,796
Total contigs in blastn and DIAMOND analysis	1,460	3,544	2487	1,445	9,967	40,214	84,627	96,064
% contigs of assembled contigs in analysis	82%	94%	90%	75%	97%	98%	99%	99%
% plant hits	32.7%	17.1%	23.3%	37.0%	0.0%	0.2%	0.3%	0.3%
% no hits	27.3%	29.9%	22.8%	14.4%	15.0%	17.4%	10.3%	10.7%



Chapter 5

The *Synchytrium endobioticum* AvrSen1 triggers a Hypersensitive Response in Sen1 potatoes while natural variants evade detection

Bart T.L.H. van de Vossenbergh, Charlotte Prodhomme, Gert van Arkel,
Marga P.E. van Gent-Pelzer, J.E. Marjan Bergervoet-van Deelen, Balázs Brankovics,
Jarosław Przetakiewicz, Richard G.F Visser, Theo A.J. van der Lee, and Jack H. Vossen



Chapter 5

This chapter was published on the BioRxiv preprint server as “van de Vossen B.T.L.H., Prodhomme C., van Arkel G., van Gent-Pelzer M.P.E., Bergervoet-van Deelen J.E.M., Brankovics B., Przetakiewicz J., Visser R.G.F, van der Lee T.A.J., and Vossen J.H. (2019) The *Synchytrium endobioticum* AvrSen1 triggers a Hypersensitive Response in Sen1 potatoes while natural variants evade detection, doi.org/10.1101/646984”. Slight formatting changes were made to facilitate thesis consistency.

Abstract

Synchytrium endobioticum is an obligate biotrophic fungus of the phylum Chytridiomycota. It causes potato wart disease, has a world-wide quarantine status and is included on the HHS and USDA Select Agent list. *S. endobioticum* isolates are grouped in pathotypes based on their ability to evade host-resistance in a set of differential potato varieties. So far, thirty-nine pathotypes are reported.

A single dominant gene (*Sen1*) governs pathotype 1 resistance and we anticipated that the underlying molecular model would involve a pathogen effector (*AvrSen1*) that is recognized by the host.

The *S. endobioticum* specific secretome of fourteen isolates representing six different pathotypes was screened for effectors specifically present in pathotype 1(D1) isolates but absent in others. We identified a single *AvrSen1* candidate. Expression of this candidate in potato *Sen1* plants showed a specific hypersensitive response, which co-segregated with the *Sen1* resistance in potato populations. No HR was obtained with truncated genes found in pathotypes that evaded recognition by *Sen1*.

These findings established that our candidate gene was indeed *Avrsen1*. *AvrSen1* is a single copy gene and encodes a 376 amino acid protein without predicted function or functional domains, and is the first effector gene identified in Chytridiomycota.

Introduction

Potato wart is a severe disease of cultivated potatoes, caused by the soil borne obligate biotrophic fungus *Synchytrium endobioticum* (Schilb.) Percival. Hypertrophic growth of the infected tissue resulting in wart-like malformations, that destroy the economic value of the potato tubers, characterize the disease [1]. Resting spores that are formed in the warted potato tissues can remain viable and infectious in soils for decades [2]. No chemical control agents are available to eradicate the pathogen from contaminated soils [3], and disease management relies on preventing its introduction and spread through the deployment of resistant potato varieties. Because of its impact on potato cultivation, *S. endobioticum* has a quarantine status in most countries with potato production and is included on the HHS and USDA Select Agent list [4]. The pathogen has been reported in potato-growing countries in Asia, Africa, Europe, North-America, South-America, and Oceania [5].

Isolates of the pathogen are grouped into pathotypes based on their interaction with a differential set of resistant potato varieties [6]. For decades after the first description of the pathogen by Schilberszky [7], only a single pathotype was recognized which is nowadays referred to as pathotype 1(D1). Emergence of a new pathotype, now known as 2(G1), was recognized when wart formation was discovered on formerly resistant potato cultivars in 1941 [8]. Currently, 39 pathotypes have been described, of which pathotypes 1(D1), 2(G1), 6(O1) and 18(T1) are most widespread and considered to be of main importance [9]. Comparative analysis of thirty *S. endobioticum* isolates using their mitochondrial genome sequences, showed that the pathogen was introduced multiple times in Europe and that several pathotypes emerged in parallel. Interestingly, single isolates were found to represent populations of distinct genotypes [10].

As other Chytridiomycota, *S. endobioticum* does not form hyphae or mycelia but produces summer and resting sporangia that contain motile zoospores [11, 12]. Zoospores encyst on the potato host cell and the content of the spore penetrates the host cell leaving the empty cyst wall outside the host. After penetration, the fungal thallus is separated from the point of infection and migrates to the host nucleus. The intracellular lifecycle is completed by forming summer sporangia which give rise to new zoospores that either re-infect the host or conjugate to produce biflagellate zygotes that give rise to resting spores after host penetration [1, 13, 14]. Even in incompatible interactions, zoospores penetrate host cells after which an immune response is triggered resulting in a localized cell death [15].

Plant-pathogen interactions have evolved over millions of years, generating a broad range of diversity on both sides of the interaction. Weapons in this arms race are pathogen effector proteins and plant resistance (*R*) genes. The molecular mechanisms involved in plant-fungi interactions have been reviewed by various authors [16-19]. In short, the first line of defense developed by plants is based on the recognition of broadly conserved pathogen/microbe associated molecular patterns (PAMPs/MAMPs) by pattern-recognition receptors (PRRs) on the host's cell surface. The fungal cell wall component chitin functions as a PAMP, which triggers the basal plant defense called PAMP-triggered immunity (PTI). To prevent PTI, pathogens have evolved effectors to shield PRR ligands, suppress the host signaling and immune response, or to manipulate the host cell physiology. In turn, these effectors can be specifically recognized through receptors encoded by plant *R* genes, which results in effector-triggered immunity (ETI). ETI is typically associated with a hypersensitive response (HR) and local cell death, which restricts the spread of the pathogen at the site of infection. Pathogen genes coding for effectors that are recognized by a plant *R* gene and trigger ETI are called avirulence (*Avr*) genes. In agricultural systems the arms race model [20], in which both the pathogen and the host develop in continuous cycles causing temporary fixation of new effector and *R* gene alleles, has been suggested to be the main driving force in pathogen effector and plant defense evolution [21].

Potato is host to many pathogens from diverse taxonomical groups, such as oomycetes (e.g. *Phytophthora infestans*), bacteria (e.g. *Ralstonia* species), nematodes (e.g. *Globodera* species), but also viruses (e.g. PVY). Resistance in most pathosystems is governed by *R* gene recognition of specific effector molecules. The best elaborated examples are the gene-for gene [22] interactions in the *P. infestans* pathosystem [23]. In potato, several quantitative resistance loci (QRLs) for *S. endobioticum* resistance have been identified, and two of these give resistance to pathotype 1(D1) isolates only: *Sen1* [24] and *Sen1-4* [25] which reside on chromosomes 11 and 4 respectively. Recently, *Sen2* and *Sen3* were described which give broad resistance to multiple pathotypes, including 1(D1) for *Sen2* [26, 27].

The recently annotated genomes of the pathotype 1(D1) isolate MB42 (QEAN00000000 v.1) and pathotype 6(O1) isolate LEV6574 (QEAM00000000 v.1) [28] open up the possibility to identify *Avr* gene candidates using a comparative genomic approach. Availability of *S. endobioticum* *Avr* genes will greatly advance our understanding and management of this challenging obligate biotrophic soilborne pathogenic fungus. Comparative studies of genome sequences from *S. endobioticum* isolates most frequently found in Europe and

Canada, revealed an *AvrSen1* gene candidate which was present in single copy in pathotype 1(D1) genomes. The gene showed different variants in pathotypes that are not recognized by Sen1. This gene represents the first avirulence gene reported from a pathogen in the fungal phylum of Chytridiomycota. We discuss the potential applications of the applied comparative genomic strategy and the identified *AvrSen1* gene for potato wart disease resistance and management.

Results

Screening for *AvrSen1* candidates

Out of sequence data generated by Van de Vossenber *et al.* [10], 14 isolates were selected because of their >10 x median sequence coverage and known pathotype identity (Fig. S1). The *S. endobioticum* pathotype 1(D1) isolate MB42 genome contains 8,031 protein coding genes of which 477 (5.9 %) were regarded as the MB42 secretome on account of the presence of a signal peptide and absence of transmembrane domains or GPI anchors.

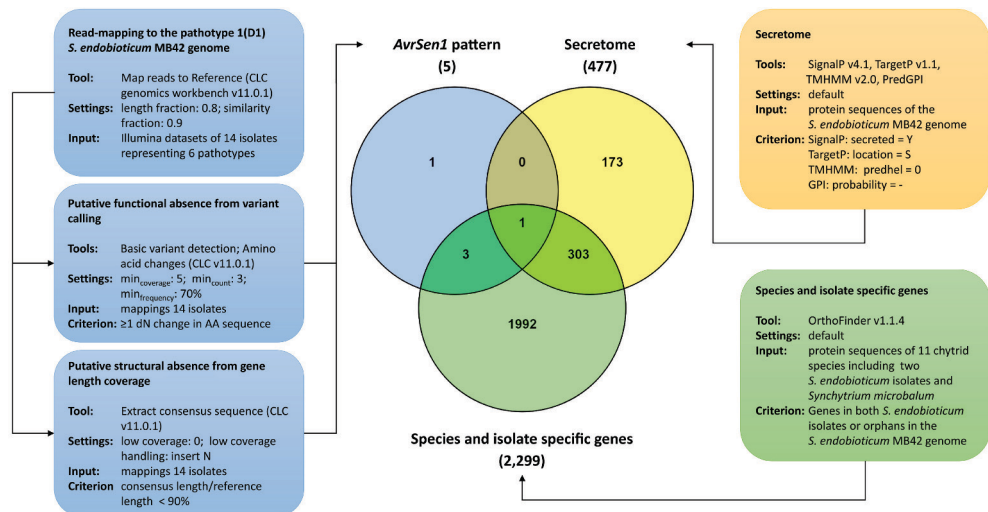


Figure 1. Converging approaches for the identification of *AvrSen1* candidates. The pipeline contains three main elements indicated in blue (putative presence and absence of genes), yellow (determination of the secretome) and green (species and isolate specific genes). Bioinformatic tools, their settings, input and output criteria are shown. Intersections of these three elements resulted in the identification of a single candidate *AvrSen1* gene from a total of 8,031 genes.

Almost two-thirds of the secretome (n =304, 64 %) consists of species specific genes (Table S3).

To determine polymorphisms at the encoded protein sequence level (loss of function) and the DNA level (gene loss), reads of the 14 isolates were mapped to all nuclear genome scaffolds of pathotype 1(D1) isolate MB42. For the different isolates, dN substitutions were observed in 171 to 1,126 gene models at a minimal frequency of 70% (Table S4). Pathotype 8(F1), 18(T1) and 38(N1) isolates had significantly higher numbers of genes with dN substitutions (general ANOVA, $p < 0.001$). From the structural absence analysis of the different isolates, 41 to 206 gene models had less than 90% coverage relative to the MB42 reference genome and were predicted to be structurally absent.

Five gene models showed the hypothesized *AvrSen1* pattern, i.e. present in pathotype 1(D1) isolates and absent in the higher pathotypes. Four of these genes were species specific, and only one belonged to the secretome (Fig. 1). Manual verification of gene prediction, functional annotations, and weighing of the significance of polymorphisms (i.e. leading to conservative or non-conservative non-synonymous substitutions), only a single gene remained as a *AvrSen1* candidate: SeMB42_g04087. Opposed to the other genes with the hypothesized *AvrSen1* pattern, all polymorphisms in SeMB42_g04087 led to the introduction of stop codons resulting in truncated gene models.

Gene SeMB42_g04087 is unique to *S. endobioticum* genomes, even in the closely related species *Synchytrium microbalum* no orthologs were found. The gene is 1,360 nucleotides long, consists of four exons and three introns, and encodes a 376 amino acid protein containing a signal peptide with cleavage site at amino acid position 30. No predicted functions, transmembrane domains, or other functional annotations such as nuclear localization signals, chloroplast targeting peptides or mitochondrial targeting peptides were identified. The protein contains two cysteine residues, and compared to the entire *S. endobioticum* MB42 proteome it has a relative high percentage of tyrosine residues (MB42 proteome: 2.8%; SeMB42_g04087: 5.1%), while other amino acid residues show more average frequencies.

***AvrSen1* and its variants**

The candidate *AvrSen1* gene was found to be present in the two pathotype 1(D1) isolates, and structurally absent in one of two pathotype 2(G1) isolates (i.e. MB08) and in both pathotype 18(T1) isolates (Fig. 2). In the other isolates two forms of functional absence were observed: a G insertion at position 769 of the coding sequence (CDS) causing a frameshift, thereby introducing a stop codon at position 256 of the amino acid sequence (*avrSen1*:Asp256>stop256), and a C>T substitution at position 916 of the CDS causing the introduction of a stop codon at position 306 of the amino acid sequence (*avrSen1*:Gln306>stop306). The truncated *avrSen1*:Asp256>stop256 and *avrSen1*:Gln306>stop306 are referred to as variant 1 and variant 2 respectively (Fig. 2).

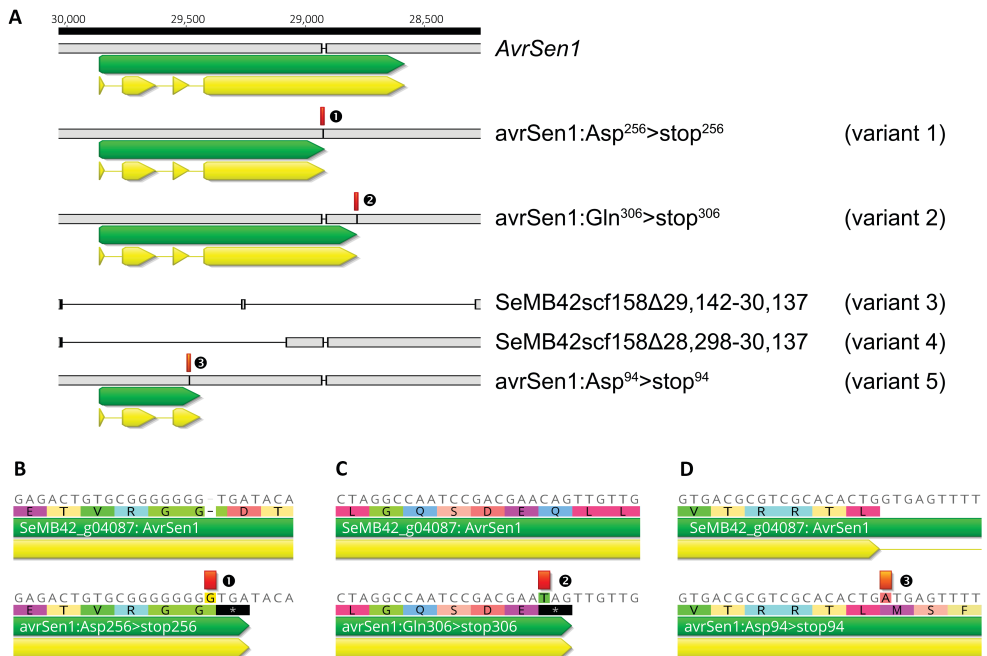


Figure 2. The *AvrSen1* gene and its variants. **A.** Sequences identical to the MB42 *AvrSen1* gene are presented in light grey, and variants to the *AvrSen1* gene are annotated in red (**1**, **2** and **3** for the respective truncated variants). Gene annotation and coding sequence (CDS) annotation are in green and yellow respectively. **B.** Detail of the G insertion in the genomic sequence of variant 1 isolates relative to isolate MB42 at position 769 of the CDS causing a frameshift and introducing a stop codon. **C.** Detail of the C>T substitution in the genomic sequence of variant 2 isolates relative to isolate MB42 at position 916 of the CDS introducing a stop codon. **D.** Detail of the G>A substitution in the genomic sequence of the variant 5 isolate on the first base of the third intron as present in isolate MB42 which results in a loss of the splice site. The numbering above the alignment indicates the original position on SeMB42_scf158 as it is represented in reverse complement orientation.

Variant 1 was detected in pathotype 2(G1) isolate SE4, 6(O1) isolates SE5 and SE6, 8(F1) isolate DEN01, and 38(N1) isolate MB56. Variant 2 was exclusively found in the Canadian pathotype 6(O1) isolates LEV6574, LEV6602, LEV6687, and LEV6748. Also, two forms of structural absence were identified: i.e. a deletion ranging from position 29,142 to 30,137 (SeMB42scf158 Δ 29,142-30,137) which was found in the two pathotype 18(T1) isolates (MB17 and SE7), and a deletion found only in one pathotype 2(G1) isolate (MB08) (SeMB42scf158 Δ 28,298-30,137). These two forms of structural absence are referred to as variants 3 and 4 respectively.

Specific recognition of *AvrSen1* in potato clones carrying the *Sen1* locus

Eight progeny plants were selected from the Aventura x Desiree and Aventura x Kuras crosses (Table S5). Together with the parents of these populations, seven genotypes possessed

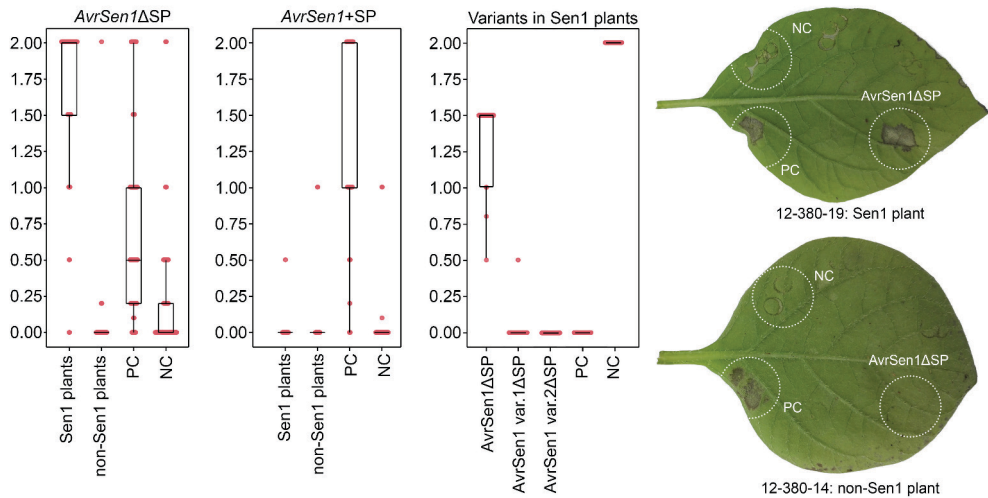


Figure 3. Agroinfiltration of *AvrSen1* variants in potato plants. Boxplots of Agroinfiltration scores obtained with constructs *AvrSen1* Δ SP (left panel), *AvrSen1*+SP (center panel), and the truncated *AvrSen1* variants var.1 Δ SP and var.2 Δ SP (right panel). Co-infiltration of *Avr8/R8* served as positive control, and the negative control consisted of GUS or *R8*. Individual scores are represented as dots. Agroinfiltration results obtained in two progeny plants of the Aventura x Kuras population represent typical reactions observed for *AvrSen1* Δ SP in *Sen1* containing plants (top leaf) and plants without *Sen1* (bottom leaf). PC represents the positive control which consisted of co-infiltration of the *P. infestans* *Avr8* and *R8*, and NC represents the negative control.

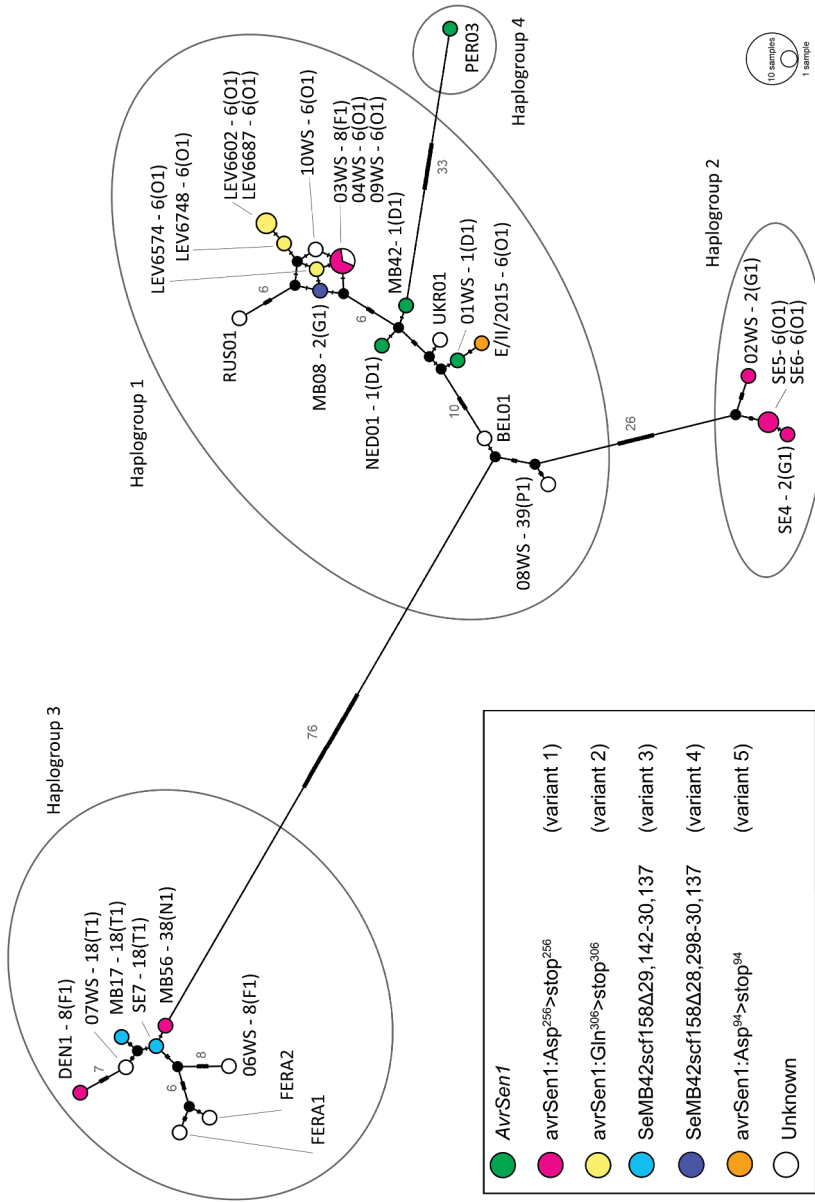


Figure 4. Distribution of *AvrSen1* and its variants among *S. endobioticum* isolates. *AvrSen1* and its variants are visualized in the mitochondrial haplotype network from Van de Vossenberget al. [10]. Colors represent the presence of *AvrSen1* or its variants for a given isolate as identified from read-mapping to the MB42 genome or through PacBio amplicon sequencing. Pathotype identities are shown when available. Black nodes represent hypothetical ancestors, and marks on the branches indicate the number of mutations, which are shown as numerical values on branches with > 5 mutations.

the *Sen1* markers and were resistant to *S. endobioticum* pathotype 1(D1) whereas the remaining four genotypes lacked the *Sen1* markers and were susceptible to pathotype 1(D1). In total 141 agroinfiltrations were performed with the *AvrSen1* construct without signal peptide (*AvrSen1*ΔSP) in three different experiments, and 45 with the *AvrSen1* construct with signal peptide (*AvrSen1*+SP), and 24 for each of the truncated variants without signal peptide in two different experiments. When plants possessing *Sen1* were agroinfiltrated with the *AvrSen1*ΔSP construct, HR-like cell death was visible in 95 % of all infiltrated sites, this in contrast to agroinfiltration of *AvrSen1*ΔSP in plants lacking *Sen1*, where 98 % produced no visible reaction. This clearly showed that *AvrSen1* was recognized in a *Sen1* dependent way. Interestingly, the *AvrSen1*+SP construct did not produce an HR in *Sen1* plants suggesting cytoplasmic recognition of the gene product. Also, both truncated variants *avrSen1*:Asp256>stop256 and *avrSen1*:Gln306>stop306 were not recognized by *Sen1* plants and produced no visible reaction (Fig. 3).

Presence of *AvrSen1* and its variants in *S. endobioticum* isolates

The presence of the *AvrSen1* gene or its variants could be determined in seventeen of thirty *S. endobioticum* isolates sequenced by Van de Vossenber *et al.* [10] by mapping of Illumina sequence reads to the MB42 genome. In addition, PacBio SMRT CCS reads generated from the *AvrSen1* amplicon allowed the detection of *AvrSen1* gene or its variants for three additional isolates. Totally, for 20 isolates the major *AvrSen1* variant was determined and superimposed to the mitochondrial haplotype network that was previously generated (Fig. 4). Based on variation of the mitochondrial genome, *S. endobioticum* isolates clustered in four major groups (haplogroups), representing separate introductions from which pathotypes 2(G1) and 6(O1) emerged multiple times independently [10].

The *AvrSen1* gene was identified in both pathotype 1(D1) isolates from the Netherlands (MB42 and NED1), the German pathotype 1(D1) isolate 01WS, and in the Peruvian isolate without known pathotype identity. Variant 1 of *AvrSen1* was found most frequently among the remaining isolates, and was identified in haplogroups 1, 2 and 3. Interestingly, pathotype 6(O1) samples from Canada possess a different *AvrSen1* variant compared to the European pathotype 6(O1) isolates, i.e. variant 2 versus variant 1 respectively. Also, pathotype 2(G1) isolate MB08 from haplogroup 1 possesses a different variant compared to the pathotype 2(G1) isolates from haplogroup 2, i.e. variant 4 compared to variant 1 respectively.

Co-occurrence of *AvrSen1* and its variants within *S. endobioticum* isolates

Read mappings generated of several isolates for the identification of the *AvrSen1* candidates showed that the *AvrSen1* locus was polymorphic in some isolates, and low percentages of the *AvrSen1* wild-type were present in these isolates. To further quantify the presence of *AvrSen1* and its variants within isolates, the gene was PCR amplified from genomic DNA (Fig. 5A, B) and subjected to PacBio SMRT sequencing which produced 8,793 to 32,336 CCS for the selected isolates (Table S1). PacBio CCS reads were generated from a median number of 36 or 37 passages for all isolates included (Fig. S3). In pathotype 1(D1) isolate MB42, all reads represented the *AvrSen1* wild-type (Fig. 5C). PacBio SMRT sequencing confirmed the presence of variant 1 as the dominant haplotype in isolates SE4, SE5, SE6, and MB56 that were also included in the screening to identify *AvrSen1* candidates. In these isolates, representing three higher pathotypes, the G insertion at position 769 of the coding sequence was observed in 93% of all reads. Interestingly, the *AvrSen1* haplotype was also found in these samples with 7% of the reads lacking the insertion that leads to the truncation of the gene. Pathotype 2(G1) isolate O2WS, which was not included in the screening for *AvrSen1*

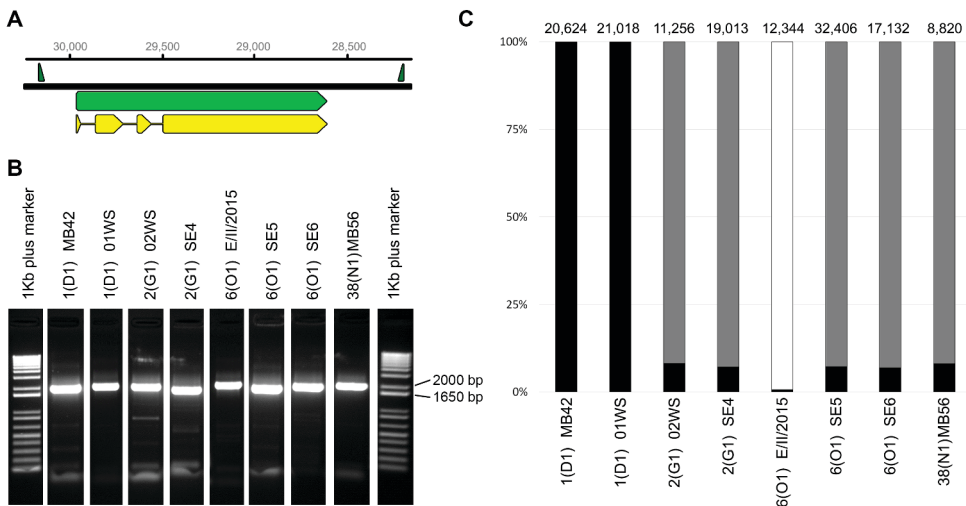


Figure 5. *AvrSen1* amplicon sequencing with PacBio SMRT. A. Primer sites flanking the *AvrSen1* gene resulting in a 1,983 nt amplicon. Primers are annotated as green triangles, the *AvrSen1* gene sequence is annotated green, and the CDS is annotated in yellow. Numbers indicate the position of the gene on SeMB42_scf158 which is presented in reverse complement orientation to present the gene in 5' to 3' direction. **B.** Amplicons obtained from selected isolates representing multiple pathotypes and mitochondrial haplogroups. The 1 kb-plus size marker was used for amplicon size estimation. **C.** Proportion of PacBio CCS reads representing *AvrSen1* (black), *avrSen1* variant 1 (grey) and *avrSen1* variant 5 (white). The number of PacBio CCS reads generated per sample are shown above the bars.

candidates because a lack of read coverage, showed presence of the *AvrSen1* variant 1 insertion in similar percentages (92%) as the other isolates carrying the *AvrSen1* variant 1 sequenced with PacBio.

In addition to the reads mapping to the MB42 genomic scaffold containing *AvrSen1*, low numbers (1 to 57) of small (95 - 640 nt) CCS reads were obtained for the tested isolates. In many cases these small CCS reads could be identified as non-specific amplification of bacteria or potato. However, a particular 466 to 470 bp amplicon representing a deletion variant was found to be present in six of the analyzed isolates in 2 to 5 copies per isolate (on average 1.6 in 10,000 reads) (Fig. S4). The short sequences resemble the *AvrSen1* variant 3 deletion, and were found in both 1(D1) isolates, both 2(G1) isolates, and the 6(O1) isolates SE5 and SE6. The short sequence was not identified in 38(N1) isolate MB56 nor in 6(O1) isolate E/II/2015.

Disruptive selection

After two multiplication rounds of pathotype 1(D1) isolate 01WS on the semi-susceptible variety Erika, a pathotype 6(O1) phenotype was obtained for the resulting isolate E/II/2015 [10]. Both isolates 01WS and E/II/2015 were PacBio sequenced to determine if a shift in the genetic population had resulted in the loss of the *AvrSen1* gene.

Almost all (>99.9%) of the 20,986 PacBio CCS reads obtained for pathotype 1(D1) isolate 01WS represented the *AvrSen1* haplotype. The *AvrSen1* haplotype in E/II/2015 was almost completely lost with only 0.7% of all 12,322 reads producing the wild-type sequence. The remaining 99.3% of the reads showed a G>A substitution on position 29,183 of scaffold SeMB42_scf0158 (Fig. S5). In the original 01WS isolate, seven CCS reads had the G>A substitution (0.0003%) observed in E/II/2015. This substitution is positioned on the first base of the third intron and affects the GU dinucleotide of the third exon-intron boundary in the pre-mRNA. The conserved GU dinucleotide is required for correct splicing of the pre-mRNA and the G>A substitution in the pre-mRNA results in a loss of the splice site. As a consequence, a stop codon is introduced on position 94 of the amino acid sequence resulting in a third truncated variant (*avrSen1*:Asp94>stop94); i.e. variant 5 (Fig. 2).

Discussion

Close to forty different *S. endobioticum* pathotypes have been reported based on bioassays using a set of differentially resistant potato varieties. We hypothesized that the phenotypic differences between pathotypes are the result of different (combinations of) *R* genes present in potato varieties and presence or absence of their cognate *Avr* genes in *S. endobioticum* isolates. At present, several QRLs providing resistance to *S. endobioticum* pathotypes have been described of which two provide resistance specifically to pathotype 1(D1) isolates: i.e. *Sen1* and *Sen1-4*. To identify the cognate *Avr* for the potato *Sen1*, which we refer to as *AvrSen1*, we exploited the recently assembled and annotated pathotype 1(D1) genome of isolate MB42 in combination with Illumina datasets generated for 30 isolates representing pathotypes that are of current importance in Europe and Canada.

***AvrSen1* identification strategy**

As potato varieties possessing the *Sen1* locus are resistant to pathotype 1(D1) isolates but are susceptible to the “higher” pathotypes, we hypothesized that pathotype 1(D1) isolates possess the *AvrSen1* gene while higher pathotypes lost it on a functional or structural level. For the comparative genomic approach, two scenarios of loss of the *AvrSen1* gene were considered: mutations in the gene sequence leading to a loss of recognition due to the change in amino acid sequence of the *AvrSen1* protein, and (partial) deletion of the gene from the genome in higher pathotypes. We did not include criteria such as molecular weight, size and cysteine-richness which have been attributed to be effector signatures [29]. Indeed, effectors frequently are reported to be cysteine-rich and relatively small, but large effector proteins have also been reported [17]. Also selection for other effector signatures based on specific motifs (e.g. the RxLR motif in oomycetes [30] or Crinkler motifs which were also found identified in the amphibian decline chytrid *Batrachochytrium dendrobatidis* [31]), were not pursued in our approach.

Only five candidates were obtained from the presence/absence analysis, of which only one was predicted to be secreted. Manual verification showed that the gene models of the other candidates were correct and that these indeed did not belong to the secretome. Otherwise, manual verification was found to be essential to assess the reliability of read-mapping and the resulting variant calling. Also, the significance of the dN changes were manually assessed with conservative dN changes, resulting in an amino acids with the same characteristics (e.g. Alanine to Valine), being regarded as insignificant as their influence to the tertiary structure

of the protein was believed to be minimal. Surprisingly, all single nucleotide polymorphisms observed in the *AvrSen1* gene of non-pathotype 1(D1) isolates led to the introduction of a stop codon resulting in a truncated gene model.

Pathotype grouping of *S. endobioticum* isolates is based on tuber-based bioassays but agroinfiltration experiments are typically performed on leaves. We previously demonstrated that plant *R* genes for the EPPO differential set [6] were equally expressed in aboveground plant parts compared to the tuber-based assays [32]. This justifies the use of the leaf-based agroinfiltration assay to test the interaction between candidate *Avr* genes and *R* genes.

AvrSen1

As the interaction between *S. endobioticum* and its host is specific, the *AvrSen1* gene was hypothesized to be present in the species specific secretome. Indeed the identified *AvrSen1* gene had no orthologs in the closely related species *S. microbalum* which has a non-pathogenic saprobic lifestyle.

Cartwright (1926) reported that in incompatible interactions, zoospores encyst on the host cell and penetration occurs. Initially the fungal thallus develops normally, but eventually an immune response is triggered resulting in localized cell death [15]. This observation suggests that the strategy of overcoming PTI by *S. endobioticum* may be identical in the different pathotypes and that the differentiation between pathotypes is the result of effectors secreted in the cytoplasm after host penetration. The strong response of *Sen1* plants to the cytoplasmic *AvrSen1*ΔSP construct, and absence of a response when the *AvrSen1* was expressed with signal peptide, supports the intracellular recognition of the effector. Also, no nuclear localization signals, chloroplast or mitochondrial targeting peptides were present, suggesting that the protein is not localized in these subcellular compartments, but has a cytoplasmic target. Similar observations were made for the obligate biotrophic flax rust (*Melampsora lini*) effectors *AvrM* and *AvrL567* [33, 34]. In flax it are cytoplasmic NLR immune receptors that recognize the cognate *M. lini* Avr. We hypothesize that also in case of wart resistance genes NLR immune receptors are involved. Many of the described fungal *Avr* genes lack functional domains and have no predicted function [35]. Similarly, no Pfam domains, GO-terms or protein family membership were predicted for *AvrSen1*, and its function in virulence remains elusive.

The *AvrSen1* gene (SeMB42_g04087 from the MB42 genome sequence) has a single ortholog in the Canadian pathotype 6(O1) isolate LEV6574 genome, i.e. SeLEV6574_g04683. Both genomes were independently sequenced, assembled and annotated [28], and the gene prediction in the LEV6574 genome reflects the truncated *AvrSen1* variant 2 that was identified from the *AvrSen1* prediction pipeline. This indicates that not only the particular SNP is present in the Canadian genome, but also that the truncated gene model is expressed (Fig. S6). Being a single copy ortholog in both *S. endobioticum* isolates is atypical for an *Avr* gene as many effectors belong to multi-gene families and have diversified from a common ancestor [17]. In addition, the number of cysteine residues ($n = 2$) is lower, and the protein length (376 AA) is larger compared to the features generally attributed to effectors. We would have not been able to detect the *AvrSen1* gene when applying these features as selection criteria as was suggested by others [29].

***AvrSen1* and its variants in *S. endobioticum* isolates**

Five variants of the *AvrSen1* gene were identified, which suggests that the *AvrSen1* gene is under strong *Sen1* selection pressure that is mainly exerted in potato cultivation. Indeed *Sen1* is widely deployed in current potato varieties [36]. In contrast, 79% of all 8,031 *S. endobioticum* genes included in the screening for *AvrSen1* candidates do not display any dN changes or reduced gene coverage for any of the fourteen isolates included in this study. A similar observation was made by Huang *et al.* [37] who analyzed genetic variations of six *Avr* genes in the rice blast fungus *Magnaporthe oryzae*, and compared these to seven randomly selected non-*Avr* control genes. In *M. oryzae*, *Avr* genes frequently show deletions and high levels of nucleotide variation leading to (shared) non-synonymous substitution in the diversified rice blast strains. Of the five *AvrSen1* variants, three are truncated gene models as the result of single nucleotide insertions or substitutions. Similarly, in the oomycete *P. infestans* a truncated version of the Avr4 protein remains unrecognized by plants with the *R4* gene [38].

Variant 1 (avrSen1:Asp256>stop256) was found most frequently in higher pathotypes representing different mitochondrial haplogroups which are believed to be independently introduced in Europe from the Andes [10]. Hence it is likely that the mutation leading to avrSen1:Asp256>stop256 emerged early in the evolution and spread of the pathogen, which would suggest that the *Sen1* locus should also be present in (wild) potato varieties in the native range of *S. endobioticum*. Variant 2 (avrSen1:Gln306>stop306) was found exclusively

in the four Canadian pathotype 6(O1) isolates. *S. endobioticum* was introduced from Europe to Canada in the early 1900s to Newfoundland [39] from which it spread to Saint Edwards Island from where the sequenced isolates were obtained. The variant 2 haplotype was not found in the European isolates, which could be an effect of sampling, or it could indicate that the SNP leading to variant 2 is recent and occurred in the Canadian *S. endobioticum* isolates *de novo*.

In South America, the presumed center of origin of potato, *S. endobioticum* interacts with many other *Solanum* species and their respective resistance genes. In this respect it is interesting to note that the Peruvian isolate has retained the complete *AvrSen1*. *AvrSen1* is found in most isolates tested, although sometimes in a minor haplotype. Apparently, it is not easily lost or even maintained, which suggests it has a virulence function.

Disruptive selection

After two multiplications of pathotype 1(D1) isolate 01WS on potato cultivar Erika, which does not provide full resistance to 1(D1), wart formation was obtained. The fungal isolate from the warted tissues (E/II/2015) was pathotyped, and produced a pathotype 6(O1) phenotype. This sample could be multiplied on (among others) Producent which contains *Sen1* (Prodhomme *et al.*, unpublished). From the amplification of the *AvrSen1* gene followed by PacBio SMRT sequencing we expected to observe a loss of the *AvrSen1* gene as a result of a disruptive selection induced by the partially resistance cultivar Erika. Indeed we observed a loss of *AvrSen1* in isolate E/II/2015, but to our surprise a novel variant was found: variant 5, which is the result of a substitution on the first base of the third intron in the *AvrSen1* gene sequence. The first two bases of exon-intron boundaries are highly conserved and are required for splicing of the introns out the pre-mRNA [40]. Polymorphisms in this sequence results in a read-through of the third exon resulting in the introduction of a stop codon at the 94th amino acid position. Low levels (0.0006%) of the SNP leading to *AvrSen1* variant 5 in isolate E/II/2015, were identified in isolate 01WS. However, at these low percentages it is impossible to differentiate between sequence errors introduced by PacBio SMRT sequencing, errors introduced by amplification errors, or a true genotype. Nonetheless, the CCS reads with the variant 5 haplotype in isolate 01WS were of high quality and were built from 4 to 125 passages with an average of 39 passages.

Where in isolate E/II/2015 almost all variation (99.3%) was selected against the *AvrSen1* allele, low frequencies of *AvrSen1* (7.0 - 8.2%) were identified in variant 1 isolates O2WS, MB56, SE4, SE5, and SE6. These variant 1 isolates were multiplied on the susceptible variety Deodara that does not contain *Sen1*. *AvrSen1* is found in most isolates tested, although sometimes in a minor haplotype. Both alleles could be maintained in the population on account of balancing selection, and could have a positive effect on the overall fitness of the population.

Closing remarks and outlook

Here we report the identification of the first *Avr* gene in Chytridiomycota which demonstrates that the gene-for-gene model applies to the potato-*S. endobioticum* pathosystem. Our strategy proved to be effective in identifying the *AvrSen1* gene and a similar approach may result in uncovering more *Avr* genes. This is particularly true when it is specifically matched to QRL or R gene based predicted absences and presences. *AvrSen1* could also be instrumental in finding *Avr* genes in other chytrid pathogens such as *B. dendrobatidis*, and may also help to identify the potato *Sen1* gene.

The *AvrSen1* gene is under strong selective pressure and several forms of loss of the *AvrSen1* locus were observed. Given the broad presence of variant 1, and to a lesser extent variant 2, in non-pathotype 1(D1) isolates, these could potentially be recognized by other *R* genes, making them avirulence factors. Since most mutations resulted in a C-terminal truncation, this suggests that this part is essential for recognition by *Sen1*. Additional research is required to show which part of *AvrSen1* is indeed recognized.

We previously observed that pathotypes 2(G1) and 6(O1) isolates that produced the same phenotype were present in different mitochondrial haplogroups [10], and concluded that these phenotypes could have emerged independently from different genetic backgrounds. Our results regarding the different types of loss of *AvrSen1* in the pathotype 2(G1) and 6(O1) isolates further strengthens this hypothesis (Fig. 4). Additionally, when regarding all species specific secreted proteins as potential effectors, different patterns of predicted loss of function can be seen between the two pathotype 2(G1) isolates, and the Canadian and European pathotype 6(O1) isolates. This analysis also allows further differentiation between the two pathotype 1(D1) isolates (Fig. S7). These different types of pathotype 2(G1) and 6(O1) could have evolved their own set of *Avr* genes which are not detected in the current pathotyping methods as the cognate *R* genes are not included in the differential potato

panel. The *AvrSen1* gene can be used to screen isolates to identify these possible different genotypes in isolates phenotypically identified as pathotype 1(D1). In time, functional markers such as the *AvrSen1* gene can contribute in the design of alternative pathotyping assays.

Materials and methods

Identification of *AvrSen1* candidates

Whole genome sequence data used in this study were generated by Van de Vossenbergh *et al.* [10] and were obtained from resting spores extracted from fresh warts (table S1). These datasets comprised 30 isolates that were grouped into seven different pathotypes. Datasets with >10x median coverage and known pathotype identity were included in this study (Fig. S1). Sequence reads were mapped (length fraction: 0.8, similarity fraction: 0.9) to the *S. endobioticum* pathotype 1(D1) isolate MB42 reference genome [28] in CLC genomics workbench v11.0.1 (Qiagen, the Netherlands), and mappings were improved with the local realignment tool (default settings) to better resolve the mapping in areas around insertions and deletions.

To determine putative amino acid sequence polymorphisms in different isolates, variants were called using the basic variant detection tool (mincoverage: 5, mincount: 3, Minfrequency: 70%, broken reads: ignore), and non-synonymous (dN) changes were identified with the Amino Acid Changes tool. Using this tool, the number of dN changes relative to the MB42 genome were determined per gene for each of the isolates tested. Genes with one or more dN change in an isolate were regarded to be putatively functionally absent (i.e. loss of function). To determine (partial) deletions of genes in different isolates, the percentage of length coverage relative to the MB42 reference genome was determined. Per mapped isolate, a consensus sequence was created for each gene, and the number of nucleotides with no coverage were determined. Genes with less than 90% coverage in an isolate were regarded to be putatively structurally absent. Genes present in pathotype 1(D1) isolates, but absent in higher pathotypes were regarded as candidate *AvrSen1* genes.

These candidate genes were individually inspected to verify if they were legitimate candidates by determining: a. correctness of gene prediction by checking the MB42 RNAseq [28] read mapping to the MB42 reference genome; b. significance of dN change and variants present in other pathotypes by differentiating between conservative and non-conservative

dN changes ; c. significance of <90% length coverage, and d. functional predictions with InterProScan. Presence of nuclear localization signals, chloroplast or mitochondrial targeting peptides in *AvrSen1* were determined with LOCALIZER v1.0.4 [42]. A graphical summary of detection pipeline is presented in Figure 1.

***S. endobioticum* specific secretome**

The secretome was defined as proteins possessing a secretion signal as predicted with SignalP v4.1 [43], absence of a mitochondrial targeting peptide as determined with TargetP v1.1 [44] and absence of transmembrane helices and/or GPI anchors as determined with TMHMM v2.0 [45] and PredGPI [46] respectively. Next, OrthoFinder v1.1.4 [47] output generated by Van de Vossen and Warris *et al.* [28] was used to identify *S. endobioticum* specific genes. In short, protein sequences of the *S. endobioticum* pathotype 1(D1) MB42 and pathotype 6(O1) LEV6574 reference genomes were compared to the proteomes of nine other chytrid isolates, including the closely related saprobic *S. microbalum* and the more distant amphibian decline pathogen *B. dendrobatidis*. From this analysis, orthologous groups unique to both *S. endobioticum* isolates or isolate MB42 were regarded *S. endobioticum* or pathotype 1(D1) specific.

Cloning of *AvrSen1* candidates and variants

Coding sequences of the *AvrSen1* candidate and its truncated variants were synthesized by GenScript (China) after codon optimization for plants and removal of *Bsal*, *Bpil*, and *BsmBI* restriction sites. Gene expression constructs were prepared with the Golden Gate Cloning system and contained the *AvrSen1*, *avrSen1*:Asp256>stop256, *avrSen1*:Gln306>stop306 coding sequences inserted between the Cauliflower mosaic virus (CaMV) pICSL13001 promoter + 5'UTR and the CaMV pICH41414 3'UTR + terminator sequences [48] in a modified pBINPLUS binary vector (pBINPLUS-GG). *AvrSen1* constructs were prepared with and without the *Nicotiana benthamiana* CRT (calreticulin) signal peptide pICH37326 to test for differential apoplastic or cytoplasmic recognition. The correctness of the constructs was verified using Sanger sequencing. pBINPLUS-GG plasmids with *AvrSen1* inserts were transformed to *A. tumefaciens* strain AGL + VirG [49] using electroporation. The presence of the intact plasmid was confirmed using PCR before using the transformed colonies for agroinfiltration.

Agroinfiltration

The verified *AvrSen1* candidate and two of the functional variants were cloned and transiently expressed in potato leaves to determine if their gene products were recognized by *Sen1*. Progeny from crosses between the pathotype 1(D1) resistant varieties Desiree and Kuras and the pathotype 1(D1) susceptible variety Avenra were selected (Prodhomme *et al.*, unpublished), based on their resistance or susceptibility to *S. endobioticum* pathotype 1(D1), on the presence or absence of *Sen1*, and on their competence for transient expression.

All progeny clones and the parent varieties were field propagated to provide tubers. In 2016, six and eight tubers respectively, were tested for pathotype 1(D1) resistance using the Glynne-Lemmerzahl and Spieckermann bioassays [50]. A subset of clones and the parents were re-phenotyped in 2017 with 6 and 12 tubers respectively. Presence or absence of *Sen1* was determined using five KASP markers specific to the *Sen1* haplotype (Prodhomme *et al.*, unpublished). The progeny clones were grown *in vitro* and tested for Agrocompetence with *A. tumefaciens* suspensions at OD 0.3 and 0.1. A positive control consisting of the co-infiltration of the *P. infestans* *Avr8* gene and the cognate *R8* gene [51] and a negative control (GUS) were used in this experiment. Non-competent lines which showed aspecific reactions to *A. tumefaciens* or did not show any HR when infiltrated with the positive control were excluded.

The transformed colonies were cultured in 5 mL of LB medium containing the appropriate antibiotics and grown overnight (28 - 30 °C, shaking 200 rpm). Between 20 to 200 µL of the LB cultures was diluted in 15 mL YEB medium containing the appropriate antibiotics, 1.5 µL acetosyringone (200 mM) and 150 µL MES (2-(N-morpholino)-ethane sulfonic acid, 1 M) and grown overnight (28 to 30 °C, shaking 200 rpm). The YEB cultures were centrifuged for 10 min at 4,000 x g. The supernatant was poured off and the pellet carefully re-suspended until the appropriate OD in freshly made MMA containing 1 mL/L of acetosyringone. The cultures were incubated 1 to 3 hours before infiltrations at room temperature in the dark.

Potato plants were clonally propagated from *in vitro* culture on Murashige & Skoogs medium supplemented with 2% Sucrose. Two weeks after incubation at 25 °C, the rooted cuttings were transferred to the greenhouse in 11 cm pots with potting soil under greenhouse conditions (constant 18 to 21 °C and 90% relative humidity with a light regime of 16 h light /8 h darkness). Using a syringe, small amounts of bacterial suspensions (OD 0.3 or 0.1) containing the pBINPLUS-GG with inserts of interest, were injected on the abaxial side

of selected leaves of 3 to 4 weeks after planting in the greenhouse. Each genotype was tested in duplicate and per plant three leaves were injected, referred to as “low”, “middle”, and “high” based on their relative position on the plant. On each leaf, a positive control (co-infiltration of *Phytophthora infestans* Avr8/R8 [51]) and a negative control (GUS or the R8 gene) was included to monitor the effectiveness of the Agroinfiltration experiments. After injection, plants were kept in the greenhouse for 2 to 5 days before scoring reactions following Rietman *et al.* [52].

AvrSen1 and its variants in *S. endobioticum* isolates

The presence of functional *AvrSen1* genes in sequenced isolates was superimposed to the *S. endobioticum* haplotype network as described by Van de Vossenberg *et al.* [10], which was based on complete mitochondrial genomes. Isolates in the network were colored according to presence of *AvrSen1* or its functional variants, which was based on the Illumina read-mappings of 30 isolates (identification of *AvrSen1* candidates) and PacBio SMRT sequencing of *AvrSen1* amplicons (see below).

Co-occurrence of *AvrSen1* and its variants within *S. endobioticum* isolates

As *S. endobioticum* isolates may contain a population of genotypes [10], and we observed allelic variation for *AvrSen1* in the Illumina sequences derived from a single isolate, we assessed the variation of *AvrSen1* within isolates by next generation sequencing of *AvrSen1* amplicons. Isolates used by Van de Vossenberg *et al.* [10] representing different pathotypes and mitochondrial haplogroups were selected for the analysis, including the German pathotype 1(D1) isolate 01WS and laboratory isolate E/II/2015 (table S2). The latter isolate was produced after two multiplications of 01WS on the semi-resistant variety Erika after which it showed a pathotype 6(O1) phenotype.

Primers were designed allowing generic amplification of the *AvrSen1* gene (Fig. S2). Amplification was performed in 20 μ L reaction mixes based on the Takara Premix HotStart (TaKaRa, Japan) reagents containing 500 nM of a tagged *AvrSen1_fw* (tag-5'-CTG GAA GCT CTA TTT CAT AGG TCA-3') primer, 500 nM of a tagged *AvrSen1_rv* (tag-5'-CAC TCA CTC GTG CCA TTT CTA-3') primer and 2 μ L target DNA. PCR thermocycler conditions were as follows: 2 min 94 °C initial denaturation followed by 35 cycles of (30 sec 94 °C, 30 sec 55 °C, and 2 min 72 °C) and with a final elongation of 5 min at 72 °C. Amplification primers were tagged to create sample specific combinations allowing selection of isolate specific sequence data

on the pooled amplicons. Amplicons were pooled in equimolar amounts, and subjected to PacBio SMRT sequencing. From the raw PacBio sequence data, circular consensus sequences (CCS) were generated and successively, CCS data were binned based on their sample specific tags using a custom script (<https://git.wur.nl/brank001/pacbio-demultiplexing>). Sample specific CCS data were mapped to the *AvrSen1* gene in Geneious Prime [53] (Biomatters Limited, New Zealand) using a reference assembly with default settings. Next, variants were detected and quantified with the Find Variations/SNPs tool (Minimum variant frequency = 0.1, Maximum variant *p*-value = $1e-4$).

Acknowledgements

We thank Averis seeds BV, Böhm-Nordkartoffel Agrarproduktion GmbH & Co. OHG, Danespo, HZPC Holland BV, C. Meijer BV, SaKa Pflanzenzucht GmbH & Co. KG and Teagasc for funding, biological material and stimulating scientific discussions.

Funding

This research (TKI-TU 1406-056) is financially supported by the Dutch Topsector Horticulture & Starting Materials (<http://topsectortu.nl/nl/integrated-genomics-and-effectoromics-impulse-potato-wart-resistance-management-and-breeding>).

Author Contributions

BV, TL, JV produced the research outline and wrote the manuscript. BV, GA, MG, MB, and CP performed laboratory and greenhouse experiments. RV provided critical comments to the manuscript and arranged critical research infrastructure. BB provided bioinformatics support, and JP provided essential research material.

Author affiliations

BB, BV, CP, GA, JV, MB, MG, RV, TL: Wageningen University and Research, Droevendaalsesteeg 1, 6708 PB, Wageningen, the Netherlands; BV: Dutch National Plant Protection Organization, Geertjesweg 15, 6706 EA, Wageningen, the Netherlands; JP: Plant Breeding and Acclimatization Institute, Radzikow, 05-870 Blonie, Poland

References

1. Curtis, K.M., IX. — The life-history and cytology of *Synchytrium endobioticum* (schilb.), perc., the cause of wart disease in potato. Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character, 1921. 210(372-381): p. 409.
2. Przetakiewicz, J., The Viability of Winter Sporangia of *Synchytrium endobioticum* (Schilb.) Perc. from Poland. American Journal of Potato Research, 2015. 92(6): p. 704-708.
3. Hampson, M.C., Control of potato wart disease through the application of chemical soil treatments: a historical review of early studies (1909-1928). EPPO Bulletin, 1988. 18(1): p. 153-161.
4. Federal Select Agent Program, Select Agents and Toxins List. 2018 [cited 2018 25 April]; Available from: www.selectagents.gov/selectagentsandtoxinslist.html.
5. EPPO. European and Mediterranean Plant Protection Organization (EPPO) global database. 2018 [cited 2018 12 September]; Available from: <https://gd.eppo.int>.
6. OEPP/EPPO, PM 7/28 (2) *Synchytrium endobioticum*. EPPO Bulletin, 2017. 47(3): p. 420-440.
7. Schilberszky, K., Ein neuer Schorfparasit der Kartoffelnollen. Ber. Deut. Botan. Ges., 1896. 14: p. 36-37.
8. Braun, H.C., Biologische Spezialisierung bei *Synchytrium endobioticum* (Schilb.). Zeitschrift für Pflanzenkrankheiten (Pflanzenpathologie) und Pflanzenschutz, 1942. 52(11): p. 481-486.
9. Baayen, R.P., *et al.*, History of potato wart disease in Europe - a proposal for harmonisation in defining pathotypes. European Journal of Plant Pathology, 2006. 116(1): p. 21-31.
10. van de Vossenbergh, B.T.L.H., *et al.*, The linear mitochondrial genome of the quarantine chytrid *Synchytrium endobioticum*; insights into the evolution and recent history of an obligate biotrophic plant pathogen. BMC Evolutionary Biology, 2018. 18(1): p. 136.
11. Barr, D.J.S., Chytridiomycota, in Systematics and Evolution: Part A. 2001, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 93-112.
12. Karling, J.S., *Synchytrium*. 1 ed. 1964: Academic Press, New York.
13. Lange, L. and L.W. Olson, Development of the zoosporangia of *Synchytrium endobioticum*, the causal agent of potato wart disease. Protoplasma, 1981. 106(1): p. 97-108.
14. Lange, L. and L.W. Olson, Development of the resting sporangia of *Synchytrium endobioticum*, the causal agent of potato wart disease. Protoplasma, 1981. 106(1): p. 83-95.
15. Cartwright, K., On the Nature of the Resistance of the Potato to Wart Disease. Annals of Botany, 1926. 40(158): p. 391-395.
16. Cook, D.E., C.H. Mesarich, and B.P.H.J. Thomma, Understanding Plant Immunity as a Surveillance System to Detect Invasion. Annual Review of Phytopathology, 2015. 53(1): p. 541-563.
17. Lo Presti, L., *et al.*, Fungal effectors and plant susceptibility. Annu Rev Plant Biol, 2015. 66: p. 513-45.
18. Asai, S. and K. Shirasu, Plant cells under siege: plant immune system versus pathogen effectors. Curr Opin Plant Biol, 2015. 28: p. 1-8.
19. Wang, Y. and Y. Wang, Trick or Treat: Microbial Pathogens Evolved Apoplastic Effectors Modulating Plant Susceptibility to Infection. Mol Plant Microbe Interact, 2018. 31(1): p. 6-12.
20. Dawkins, R., *et al.*, Arms races between and within species. Proceedings of the Royal Society of London. Series B. Biological Sciences, 1979. 205(1161): p. 489-511.
21. Brown, J.K. and A. Tellier, Plant-parasite coevolution: bridging the gap between genetics and ecology. Annu Rev Phytopathol, 2011. 49: p. 345-67.

22. Flor, H.H., Current Status of the Gene-For-Gene Concept. *Annual Review of Phytopathology*, 1971. 9(1): p. 275-296.
23. Haverkort, A.J., *et al.*, Durable Late Blight Resistance in Potato Through Dynamic Varieties Obtained by Cisgenesis: Scientific and Societal Advances in the DuRPh Project. *Potato Research*, 2016. 59(1): p. 35-66.
24. Hehl, R., *et al.*, TMV resistance gene N homologues are linked to *Synchytrium endobioticum* resistance in potato. *Theoretical and Applied Genetics*, 1999. 98(3): p. 379-386.
25. Brugmans, B., *et al.*, Exploitation of a marker dense linkage map of potato for positional cloning of a wart disease resistance gene. *Theor Appl Genet*, 2006. 112(2): p. 269-77.
26. Plich, J., *et al.*, Novel gene *Sen2* conferring broad-spectrum resistance to *Synchytrium endobioticum* mapped to potato chromosome XI. *Theoretical and Applied Genetics*, 2018. 131(11): p. 2321-2331.
27. Prodhomme, C., *et al.*, Comparative Subreads Sets Analysis (CoSSA) is a robust approach to identify haplotype specific SNPs; Mapping and pedigree analysis of a potato wart disease resistance gene *Sen3*. *Plant Methods*, unpublished.
28. van de Vossenbergh, B.T.L.H., *et al.*, Comparative genomics of chytrid fungi reveal insights into the obligate biotrophic and pathogenic lifestyle of *Synchytrium endobioticum*. under review.
29. Sonah, H., R.K. Deshmukh, and R.R. Belanger, Computational Prediction of Effector Proteins in Fungi: Opportunities and Challenges. *Front Plant Sci*, 2016. 7: p. 126.
30. Petre, B. and S. Kamoun, How Do Filamentous Pathogens Deliver Effector Proteins into Plant Cells? *PLOS Biology*, 2014. 12(2): p. e1001801.
31. Rosenblum, E.B., *et al.*, Substrate-Specific Gene Expression in *Batrachochytrium dendrobatidis*, the Chytrid Pathogen of Amphibians. *PLOS ONE*, 2012. 7(11): p. e49924.
32. van de Vossenbergh, B.T.L.H., *et al.*, An alternative bioassay for *Synchytrium endobioticum* demonstrates the expression of potato wart resistance in aboveground plant parts. under review.
33. Catanzariti, A.-M., *et al.*, Haustorially Expressed Secreted Proteins from Flax Rust Are Highly Enriched for Avirulence Elicitors. *The Plant Cell*, 2006. 18(1): p. 243-256.
34. Dodds, P.N., *et al.*, The *Melampsora lini* AvrL567 Avirulence Genes Are Expressed in Haustoria and Their Products Are Recognized inside Plant Cells. *The Plant Cell*, 2004. 16(3): p. 755-768.
35. De Wit, P.J.G.M., *et al.*, Fungal effector proteins: past, present and future. *Molecular Plant Pathology*, 2009. 10(6): p. 735-747.
36. Prodhomme, C., Genome Wide Association Analysis of wart disease resistance in the European potato breeding germplasm. unpublished.
37. Huang, J., *et al.*, Rapid evolution of avirulence genes in rice blast fungus *Magnaporthe oryzae*. *BMC Genetics*, 2014. 15(1): p. 45.
38. van Poppel, P.M., *et al.*, The *Phytophthora infestans* avirulence gene *Avr4* encodes an RXLR-dEER effector. *Mol Plant Microbe Interact*, 2008. 21(11): p. 1460-70.
39. Hampson, M.C., History, biology and control of potato wart disease in Canada. *Canadian Journal of Plant Pathology*, 1993. 15(4): p. 223-244.
40. Roy, S.W. and M. Irimia, Splicing in the eukaryotic ancestor: form, function and dysfunction. *Trends in Ecology & Evolution*, 2009. 24(8): p. 447-455.
41. Hampson, M.C., A bioassay for *Synchytrium endobioticum* using micropropagated potato plantlets. *Canadian Journal of Plant Pathology*, 1992. 14(4): p. 289-292.

Chapter 5

42. Sperschneider, J., *et al.*, LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. 2017. 7: p. 44598.
43. Petersen, T.N., *et al.*, SignalP 4.0: discriminating signal peptides from transmembrane regions. 2011. 8: p. 785.
44. Emanuelsson, O., *et al.*, Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol*, 2000. 300(4): p. 1005-16.
45. Krogh, A., *et al.*, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*, 2001. 305(3): p. 567-80.
46. Pierleoni, A., P.L. Martelli, and R. Casadio, PredGPI: a GPI-anchor predictor. *BMC Bioinformatics*, 2008. 9(1): p. 392.
47. Emms, D.M. and S. Kelly, OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, 2015. 16(1): p. 1-14.
48. Engler, C., *et al.*, A Golden Gate Modular Cloning Toolbox for Plants. *ACS Synthetic Biology*, 2014. 3(11): p. 839-843.
49. van der Fits, L., *et al.*, The ternary transformation system: constitutive virG on a compatible plasmid dramatically increases *Agrobacterium*-mediated plant transformation. *Plant Mol Biol*, 2000. 43(4): p. 495-502.
50. OEPP/EPPPO, PM 7/28 (1) *Synchytrium endobioticum*. *EPPPO Bulletin*, 2004. 34(2): p. 213-218.
51. Vossen, J.H., *et al.*, The *Solanum demissum* R8 late blight resistance gene is an Sw-5 homologue that has been deployed worldwide in late blight resistant varieties. *Theor Appl Genet*, 2016. 129(9): p. 1785-96.
52. Rietman, H., *et al.*, Qualitative and quantitative late blight resistance in the potato cultivar Sarpo Mira is determined by the perception of five distinct RXLR effectors. *Mol Plant Microbe Interact*, 2012. 25(7): p. 910-9.
53. Kearse, M., *et al.*, Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 2012. 28(12): p. 1647-9.

Supplementary material

Supplementary figures

- Fig. S1 Median read coverage in thirty *S. endobioticum* Illumina read sets
- Fig. S2 Design of amplification primers flanking *AvrSen1* and its variants
- Fig. S3 Number of passages in PacBio SMRT CCS sequences generated from *AvrSen1* amplicons
- Fig. S4 Presence of *AvrSen1* variant 3 sequences in several *S. endobioticum* isolates
- Fig. S5 Detail of PacBio CCS reads generated for isolates O1WS and E/II/2015
- Fig. S6 RNAseq for the *AvrSen1* gene in the pathotype 1(D1) isolate MB42 and variant 2 in the pathotype 6(O1) isolate LEV6574
- Fig. S7 Maximum parsimony tree of the amino acid sequences of 112 parsimony informative proteins from the species specific secretome

Selected supplementary tables

- Table S1 *Synchytrium endobioticum* isolates and NextGen sequence data used in this study
- Table S2 Samples included in within-isolate variation analysis using PacBio sequencing of *AvrSen1* amplicons
- Table S5 Progeny and parents of Aventura x Desiree and Aventura x Kuras crosses selected for agroinfiltration

Supplementary tables S3 and S4 were not included in this thesis on account of their size. These files are available online upon publication of this chapter.

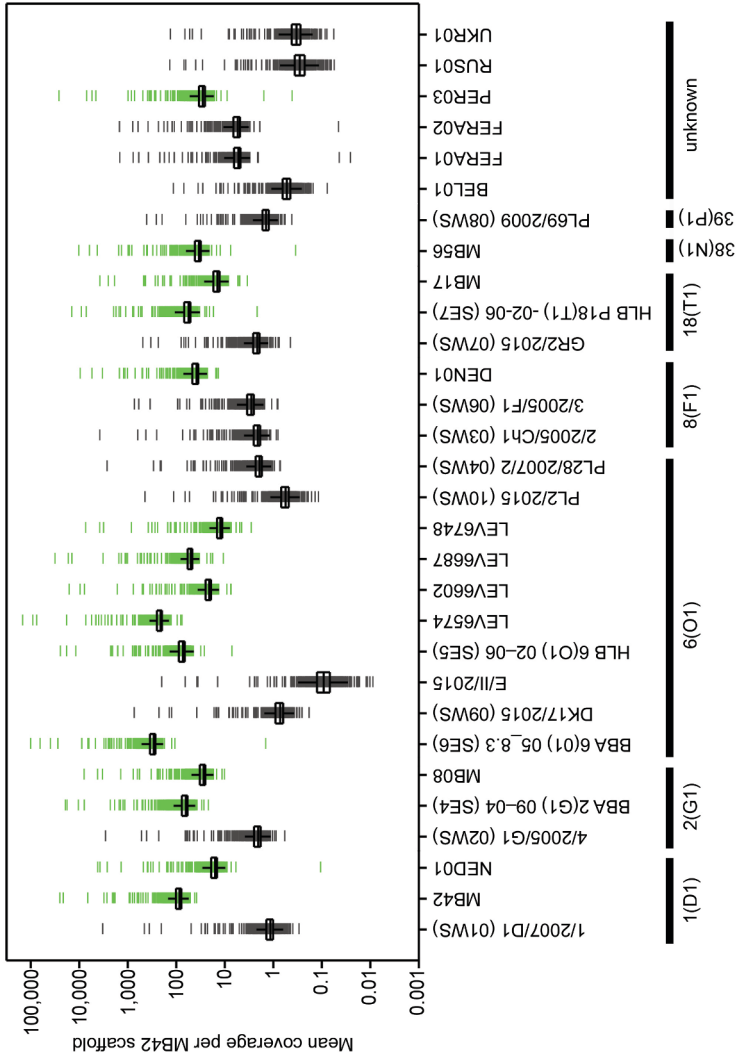


Figure S1. Median read coverage in thirty *S. endobioticum* Illumina read sets. Data generated by Van de Vossenbergh *et al.* (2018) were mapped to contigs of the *S. endobioticum* MB42 nuclear genome. Datasets were sorted based on their pathotype (x-axis) and mean coverage for each of the 786 MB42 nuclear scaffolds (y-axis) are shown on a logarithmic scale as horizontal lines. Boxplots represent the distribution of the data for the second and third quartile. Of the datasets analyzed, fifteen had >10x median coverage (represented in green). One of those did not have a known pathotype identity (PER03), and the remaining fourteen sets were selected for further analysis.

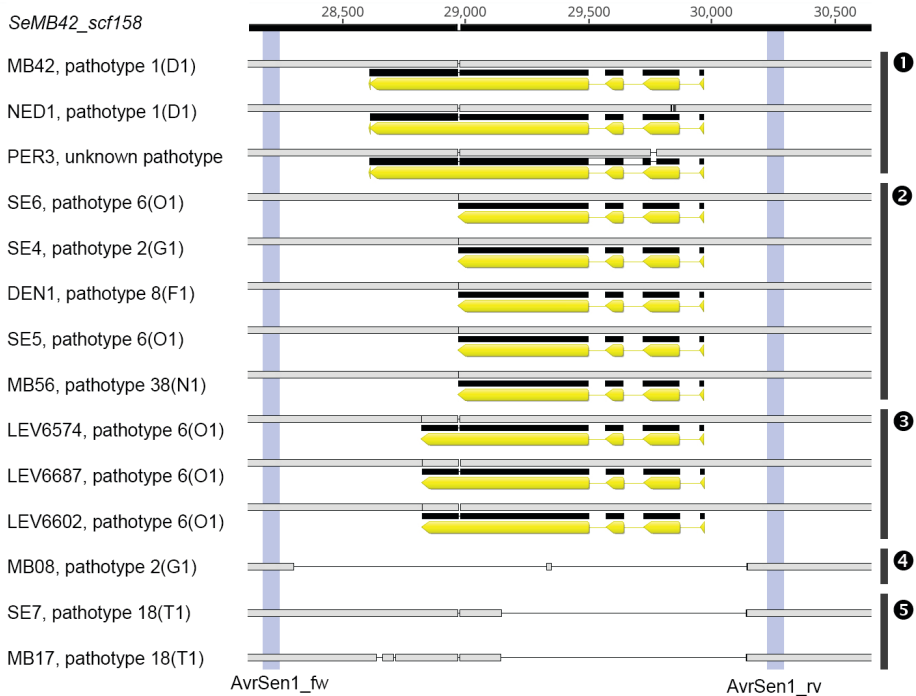


Figure S2. Design of amplification primers flanking *AvrSen1* and its variants. *AvrSen1* variants of 14 isolates were obtained after read-mapping to the MB42 reference genome. Horizontal numbering indicates the position of *AvrSen1* on *SeMB42_scf158*. Opposed to the figures in the main text, the original orientation of the genomic scaffold is retained in this image. Sequences identical to *SeMB42_scf158* are shown in light grey and differences to the reference sequence are highlighted in black. Grey bars indicate the primer annealing sites flanking the *AvrSen1* gene and its variants identified during screening for *AvrSen1* candidates (1 *AvrSen1*; 2 *avrSen1*:Asp256>stop256; 3 *avrSen1*:Gln306>stop306; 4 variant 3; and 5 variant 4). The amplicon size for *AvrSen1* MB42 is 1,983 nt.

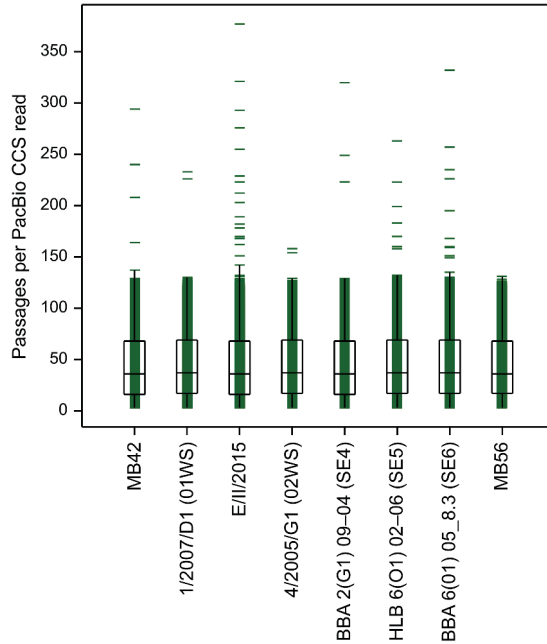


Figure S3. Number of passages in PacBio SMRT CCS sequences generated from *AvrSen1* amplicons. With PacBio, sequence fragments are sequenced multiple times (i.e. passages) to create a CCS. The number of passages is shown for CCS reads generated from eight *AvrSen1* amplicons. Boxplots represent the distribution of the data for the second and third quartile, and number of passages is shown for the individual CCS reads as green horizontal lines. The CCS reads with >150 passages represent fragments that are smaller than the 1,983 nt *AvrSen1* amplicon.

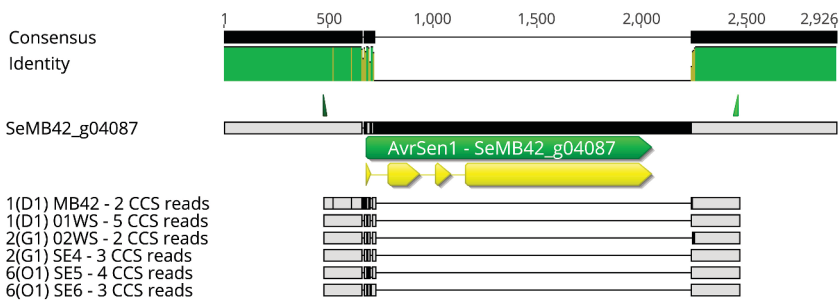
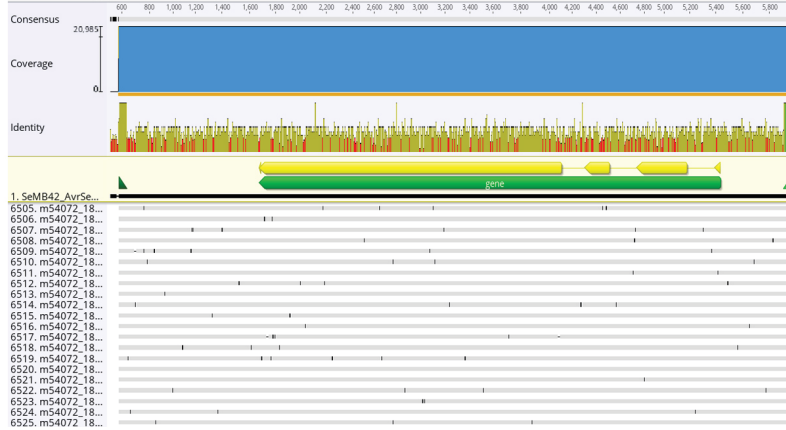
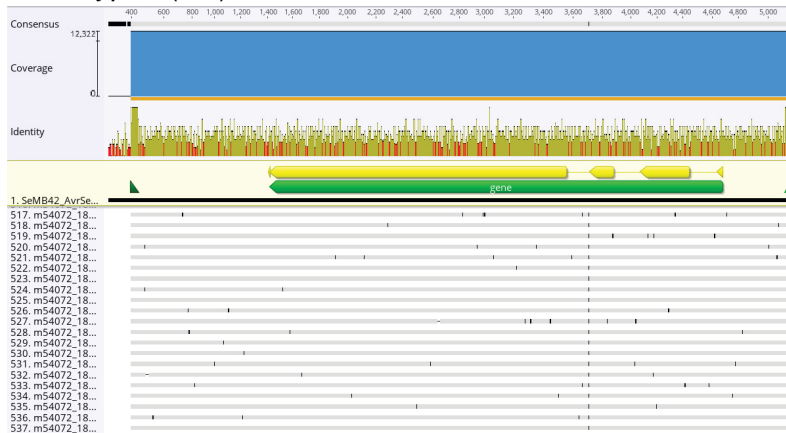


Figure S4. Presence of *AvrSen1* variant 3 sequences in several *S. endobioticum* isolates. The alignment shows an extraction of SeMB42_scf158 containing the *AvrSen1* gene and the de novo assembled short PacBio CCS reads of isolates MB42, 01WS, 02WS, SE4, SE5, and SE6 that did not produce blast hits with bacteria or potato. The number of CCS reads in the consensus sequence obtained with the de novo assembly is indicated. Sequences similar to the SeMB42_scf158 sequences are shown in grey, whereas differences are highlighted in black. Primer sites are annotated on the extraction of SeMB42_scf158.

Pathotype 1(D1) isolate 01WS



Pathotype 6(O1) isolate E/II/2015




C>T substitution on position 29,183 of scaffold SeMB42_scf0158 in 99.3% of all CCS reads 

Figure S5. Detail of PacBio CCS reads generated for isolates 01WS and E/II/2015. Isolate E/II/2015 had a pathotype 6(O1) phenotype which was obtained after two multiplications of isolate 01WS on a semi-resistant potato variety. The C>T substitution (G>A when displaying the gene sequence in 5' to 3' orientation) results in a loss of the splice site, and the introduction of a stop codon on position 94 of the amino acid sequence.

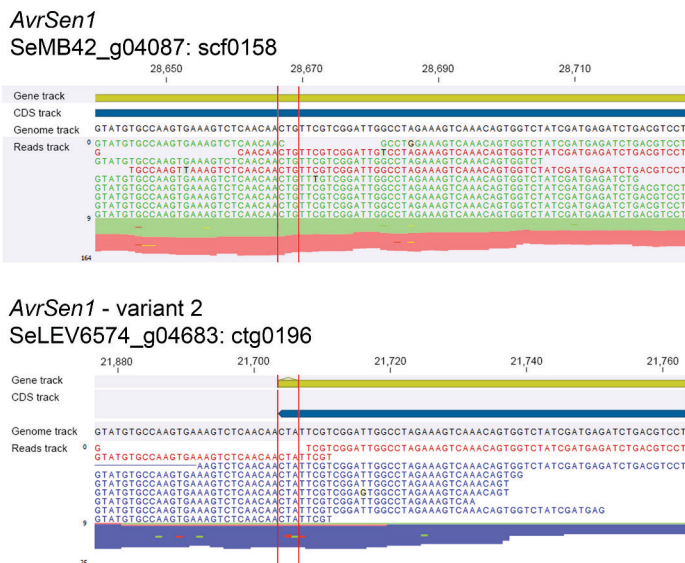


Figure S6. RNAseq for the *AvrSen1* gene in the pathotype 1(D1) isolate MB42 and variant 2 in the pathotype 6(O1) isolate LEV6574. Mapping of the RNAseq data supports the truncated gene model with the C>T substitution on position 916 of the CDS introducing a stop codon (between red lines). The read-mapping also demonstrates that the truncated gene model is expressed.

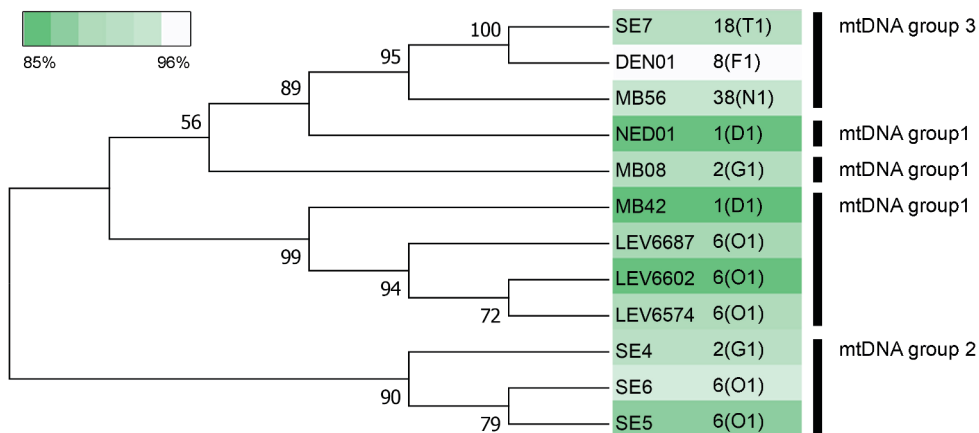


Figure S7. Maximum parsimony tree of the amino acid sequences of 112 parsimony informative proteins from the species specific secretome. Alignment of the 112 protein sequences for the twelve isolates included in the analysis resulted in a 45,506 amino acids concatenated sequence. Clustering of isolates follows the clustering as obtained with the complete mitochondrial genome as determined by Van de Vossen *et al.* (2018), opposed to a clustering based on pathotype identity. This supports the hypothesis that isolates with a similar phenotype (i.e. 2(G1) and 6(O1)) have emerged from a different genetic background. Within-isolate diversity is expressed as the percentage of polymorphic sites with intermediate SNP frequencies ($\geq 70\%$), and shown in green.

Table S1. *Synchytrium endobioticum* isolates and NextGen sequence data used in this study

Pathotype	Isolate	Origin	SRA accession(s)	Median coverage to MB42 genome
1(D1)	MB42	Langenboom, the Netherlands	ERR2286952 - ERR2286954	86
1(D1)	1/2007/D1 (01WS)	Germany	ERR2286106	1
1(D1)	NED01	Brabant, the Netherlands	ERR2286957	16
2(G1)	4/2005/G1 (02WS)	Germany	ERR2286913, ERR2286914	2
2(G1)	MB08	Mussel, the Netherlands	ERR2286949, ERR2286950	27
2(G1)	BBA 2(G1) 09-04 (SE4)	Germany	ERR2286960, ERR2286961	63
6(O1)	PL28/2007/2 (04WS)	Poland	ERR2286917, ERR2286918	2
6(O1)	DK17/2015 (09WS)	Denmark	ERR2286930 - ERR2286932	1
6(O1)	PL2/2015 (10WS)	Poland	ERR2286933 - ERR2286935	1
6(O1)	E/II/2015	Isolate obtained after two multiplications of 01WS on Erika	ERR2286939	1
6(O1)	LEV6574	St. Eleanor's, Prince Edward Island, Canada	ERR2286942 - ERR2286945	211
6(O1)	LEV6602	Augustine Cove, Prince Edward Island, Canada	ERR2286946	21
6(O1)	LEV6687	New Annan, Prince Edward Island, Canada	ERR2286947	50
6(O1)	LEV6748	New Glasgow, Prince Edward Island, Canada	ERR2286948	12
6(O1)	HLB 6(O1) 02-06 (SE5)	the Netherlands	ERR2286962	73
6(O1)	BBA 6(O1) 05_8.3 (SE6)	Germany	ERR2286963, ERR2286964	291
8(F1)	DEN01	Jylland, Denmark	ERR2286937, ERR2286938	38
8(F1)	3/2005/F1 (06WS)	the Netherlands	ERR2286919 - ERR2286925	3
8(F1)	2/2005/Ch1 (03WS)	Poland	ERR2286915, ERR2286916	2
18(T1)	GR2/2015 (07WS)	Greece	ERR2286926	2
18(T1)	MB17	Borgercompagnie, the Netherlands	ERR2286951	14
18(T1)	HLB P18(T1) -02-06 (SE7)	Borgercompagnie, the Netherlands	ERR2286965	56
38(N1)	MB56	Neveshir, Turkey	ERR2286955, ERR2286956	33

Table S1 (continued). *Synchytrium endobioticum* isolates and NextGen sequence data used in this study

Pathotype	Isolate	Origin	SRA accession(s)	Median coverage to MB42 genome
39(P1)	PL69/2009 (08WS)	Piekelnik, Poland	ERR2286927 - ERR2286929	1
unknown	BEL01	Belarus	ERR2286936	1
unknown	FERA1	United Kingdom	ERR2286940	5
unknown	FERA2	United Kingdom	ERR2286941	5
unknown	PER03	Aco Paucartambo, Pasco, Peru	ERR2286958	28
unknown	RUS01	St. Petersburg, Russian Federation	ERR2286959	1
unknown	UKR01	Ukraine	ERR2286966	1
<i>PacBio SMRT AvrSen1 amplicon sequencing</i>				
Pathotype	Isolate	Origin	SRA accession	Coverage <i>AvrSen1</i> locus
1(D1)	MB42	Langenboom, the Netherlands	ERR3176656	20,624
1(D1)	1/2007/D1 (01WS)	Germany	ERR3176657	21,018
2(G1)	4/2005/G1 (02WS)	Germany	ERR3176659	11,256
2(G1)	BBA 2(G1) 09-04 (SE4)	Germany	ERR3176660	19,013
6(O1)	E/II/2015	Isolate obtained after two multiplications of 01WS on Erika	ERR3176658	12,344
6(O1)	HLB 6(O1) 02-06 (SE5)	the Netherlands	ERR3176661	32,406
6(O1)	BBA 6(O1) 05_8.3 (SE6)	Germany	ERR3176662	17,132
38(N1)	MB56	Nevesehir, Turkey	ERR3176663	8,820

Table S2. Samples included in within-isolate variation analysis using PacBio sequencing of AvrSen1 amplicons

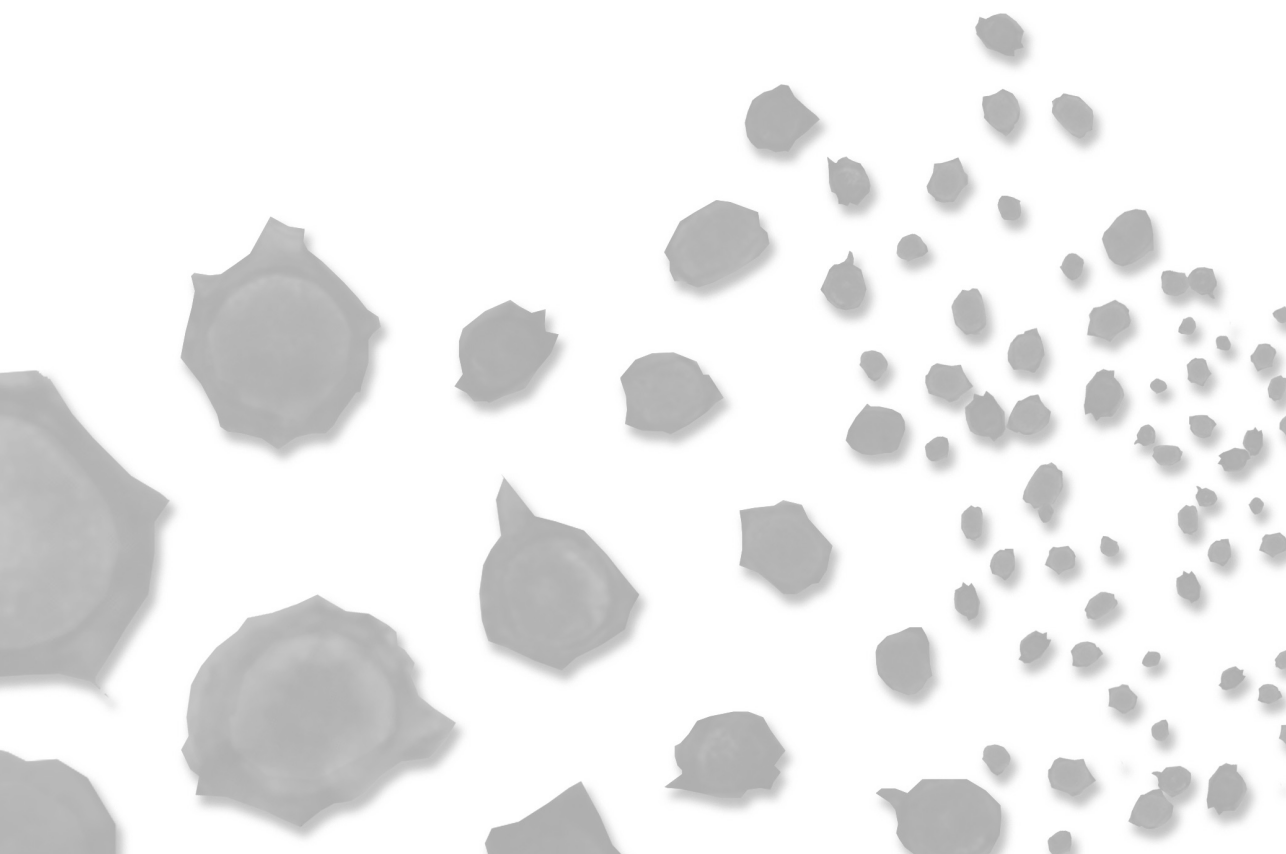
Pathotype	Isolate	Origin	Mitochondrial haplogroup ^a	Known variants ^b	Forward tag	Tag sequence	Reverse tag	Tag sequence	Coverage
1(D1)	MB42	the Netherlands	1	AvrSen1	Fw87	ATAGCGACGGGATATA	Rv83	CGTGACACTTATAGC	20,624
1(D1)	1/2007/D1 (01WS)	Germany	1	Unknown	Fw87	ATAGCGACGGGATATA	Rv84	GACTGCACATGCACGA	21,018
2(G1)	4/2005/G1 (02WS)	Germany	2	Unknown	Fw87	ATAGCGACGGGATATA	Rv86	GACGTGTCGTAGATAT	11,256
2(G1)	BBA 2(G1) 09-04 (SE4)	Germany	2	Variant 1	Fw88	ATCGCTGTGTCTATAG	Rv83	CGTGACACTTATAGC	19,013
6(O1)	E/II/2015	isolate obtained after two multiplications of 01WS on variety Erika	1	Unknown	Fw87	ATAGCGACGGGATATA	Rv85	TATGACTAGTGTACTA	12,344
6(O1)	HLB 6(O1) 02-06 (SE5)	the Netherlands	2	Variant 1	Fw88	ATCGCTGTGTCTATAG	Rv84	GACTGCACATGCACGA	32,406
6(O1)	BBA 6(O1) 05_8.3 (SE6)	Germany	2	Variant 1	Fw88	ATCGCTGTGTCTATAG	Rv85	TATGACTAGTGTACTA	17,132
38(N1)	MB56	Neusehir, Turkey	3	Variant 1	Fw88	ATCGCTGTGTCTATAG	Rv86	GACGTGTCGTAGATAT	8,820

a. as defined by Van de Vossenber *et al.* BMC Evol Biol (2018) 18: 136. <https://doi.org/10.1186/s12862-018-1246-6>

b. determined with read-mapping HiSeq data from van de Vossenber *et al.* (2018), unknown when coverage was too low to determine the AvrSen1 variant

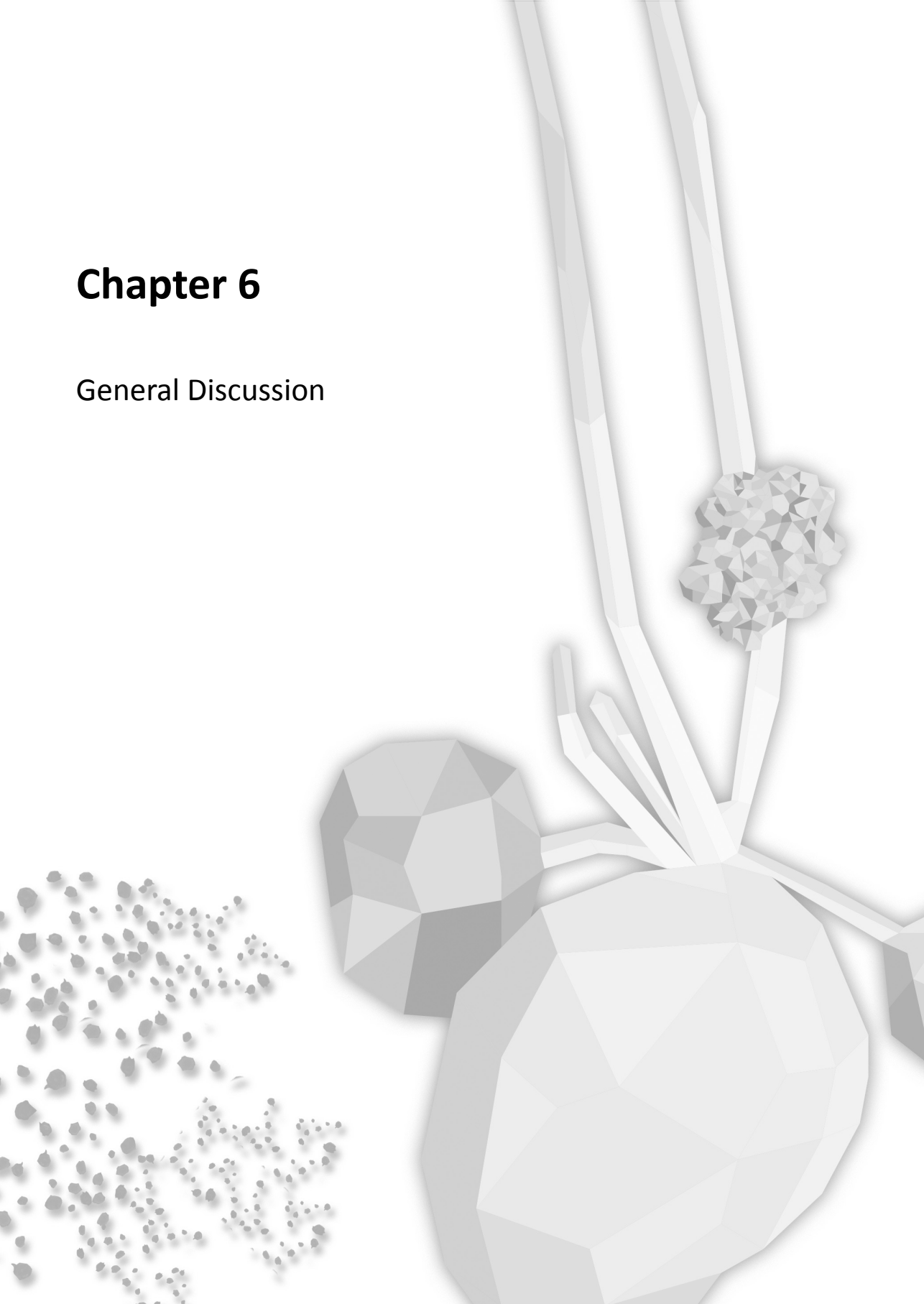
Table S5. Progeny and parents of Aventura x Desiree and Aventura x Kuras crosses selected for agroinfiltration. Results for *Sen1* haplotype specific KASP markers, and resistance screening scores bioassays are provided. Resistance screening was performed with Glynne-Lemmerzahl. Scores range from 1 (resistance, R) to 5 (susceptible, S).

Selected genotypes	Population	Sen1 KASP assay					Sen1 present	Tubers tested with Glynne-Lemmerzahl (GL)		Average GL score	Agroinfiltrated with			
		PotVar006530	PotVar006708	solcap_snp_ct_2314	PotVar0106272	PotVar0105904		2016	2017		JKI P1(D1)-2009	AvSen1ASP	AvSen1	AvSen1 variant1
12-352-01	Aventura x Desiree	0	0	0	0	0	No	6	6	4.0	9			
12-352-14	Aventura x Desiree	1	1	1	1	1	Yes			NA	9	9	9	
12-352-16	Aventura x Desiree	0	0	0	0	0	No	6	6	4.0	9			
12-352-30	Aventura x Desiree	1	1	1	1	1	Yes	6	6	1.7	21			
12-380-07	Aventura x Kuras	1	1	1	1	1	Yes	6	6	1.6	9	9	9	
12-380-14	Aventura x Kuras	0	0	0	0	0	No	6	6	5.0	21	9	9	
12-380-19	Aventura x Kuras	1	1	1	1	1	Yes	6	6	1.3	9	9	9	
12-380-37	Aventura x Kuras	1	1	1	1	1	Yes	6	6	1.3	18			
Aventura	S parent	0	0	0	0	0	No		6	2.5	12			
Desiree	R parent 1	1	1	1	1	1	Yes	6	6	1.0	12	9	24	24
Kuras	R parent 2	1	1	1	1	1	Yes	6	6	1.8	12			



Chapter 6

General Discussion



Chapter 6

Introduction

Sparrow (1958) described Chytridiomycota species (chytrids) as “recalcitrant” and “thoroughly uncooperative” organisms [1], and since then this message has been echoed by mycologists working on this basal fungal phylum. Culturing chytrids is a difficult and laborious task requiring specific baits and long incubation periods, or, in case of obligate biotrophic species such as *Synchytrium endobioticum*, is not possible at all. With the lack of culturing options, debates on aspects of its biology and pathotyping concepts, and contradicting bioassay methods, repeating Sparrow’s epithets for *S. endobioticum* is nothing less but fitting.

Early researchers studying fungi had to rely on morphological and morphometrical characteristics that could be analyzed under a microscope for classification, developmental biology and the interaction of species with their surroundings. Microscopic techniques are still valuable today, and in case of *S. endobioticum*, observations dating back to the 1920s are still among the most accurate descriptions available. With the advances in whole genome sequencing we now have new powerful tools to address fungal classification and to study molecular features underlying pathogenicity and host specificity at the molecular level. At the onset of this project we invested in collecting well-characterized biological materials and producing high quality next generation sequence data and functional annotations to allow reliable comparative genomic analyses. This, combined with integrated bioinformatic pipelines and verification of initial findings, proved to be a very successful approach.

Combining our results with the knowledge and observations from early *S. endobioticum* research, reviewed in chapter 1, allows us to lift the veil on several aspects that remained unanswered until now. The independent assembly and functional annotation of two *S. endobioticum* genomes representing different pathotypes provided the basis for our comparative genomics approach. Observations linked to chytrid biology, obligate biotrophy, genome plasticity and pathogenicity were reported from a comparison with the genomes of nine other chytrids with different lifestyles as well as six model fungi (chapter 2). Genomic diversity between and within *S. endobioticum* isolates was analyzed using the mitochondrial genome (chapter 3). We developed a new bioassay and demonstrated that potato resistance (*R*) genes are expressed in aboveground plant parts (chapter 4). This observation opened up possibilities for effector candidate screening with agroinfiltration, which eventually led to the identification and cloning of the first chytrid effector gene (chapter 5).

Chapter 6

In this general discussion, I integrate the chapters and further discuss: 1. the hypothesized model for co-evolution of *S. endobioticum* effectors and the plant immune system; 2. pathotype formation and pathotype diversity within and among fungal isolates; 3. the repeat repertoire as driving force in genome plasticity and pathogenicity; and 4. opportunities and challenges for future *S. endobioticum* research.

Model of co-evolution between *S. endobioticum* and its host

In the absence of chemical control agents, managing plant diseases in modern agricultural systems often relies on major resistance genes. In case of *S. endobioticum*, the use of resistant varieties, in combination with strict phytosanitary measures, proved very successful in limiting the spread of the disease (chapter 1). Resistance genes exert a strong selection pressure against cognate effector genes in pathogen populations. Even though fungal effectors are sometimes regarded to be dispensable due to functional redundancy, a trade-off between the need to maintain an effector as it has a role in virulence, and the need to avoid recognition by the host plant could exist [2]. Our current knowledge of the interaction of *S. endobioticum* with its host is based on light microscopy and electron microscopy studies dating from the 1920s to the 1980s [3-9]. In chapter 1 these observations were combined with our current knowledge on molecular plant-pathogen interactions [reviewed in 10-13], and a model was created to describe the interaction between *S. endobioticum* and potato at two different stages of the infection: initial invasion of the host, and proliferation inside the host cell. With the results from chapter 2 on genomic features linked to pathogenicity and chapter 5 on the identification of the first chytrid effector, we can now add further details to this model.

Invading the host

After release from either summer spores or resting spores, zoospores of *S. endobioticum* have to find a suitable host to encyst on, with the aim of penetration and infection. Zoospores of plant pathogenic fungi and oomycetes are known to direct movement in response to external stimuli [14]. In her review, Obidiegwu mentions the research performed by Esmarch (1924 - 1928) on zoospore response to chemotactic stimuli of the host, however no additional information about the nature and effects of these stimuli is presented [15]. From the analysis of the genomes of the pathogenic chytrids *S. endobioticum* and *Batrachomyces dendrobatidis*, expansion of the number of genes associated with chemotaxis were found

(chapter 2). Furthermore, zoospores can use membrane-bound G-protein-coupled receptors to sense the environment [16]. Genes coding for the G protein-coupled receptor signaling pathway were found enriched in *S. endobioticum* relative to nine other chytrid species (chapter 2), which could be involved in sensing specific host-derived signals.

When attempting to encyst on the host cell before penetration, the invading zoospore has to avoid and/or suppress the plant innate immune system. Plant membrane bound pattern recognition receptors (PRRs) trigger the immune system upon recognition of conserved pathogen associated molecular patterns (PAMPs). Chitin, a cell wall component of fungi, acts as a PAMP, and the fungus needs to secrete effector proteins that either mask the presence of chitin or suppresses the subsequent host immune response. In plants, several membrane bound chitin receptors have been identified that all contain extracellular LysM domains [17], and in fungi LysM effectors have been reported that prevent or counteract chitin-triggered immune responses [18]. In *S. endobioticum*, LysM containing proteins were found enriched relative to all nine non-plant pathogenic chytrid species analyzed (chapter 2). These LysM containing proteins could have a function in the interference of chitin-triggered immunity. Additional research is needed to determine the function of these *S. endobioticum* LysM containing proteins.

The plant innate immune system is not only triggered by conserved PAMPs. Several cell wall degradation products are categorized as damage-associated molecular patterns (DAMPs), and can lead to DAMP-triggered immunity (DTI) [19]. Reduced cell wall degrading enzyme (CWDE) content was previously found in the stealth-like pathogen *Zymoseptoria tritici* [20], and the biotrophic and symbiotic fungus *Laccaria bicolor* [21]. In plant-microbe systems it was hypothesized that a reduced CWDE repertoire may prevent DTI [22]. Relative to its saprobic sister species *Synchytrium microbalum*, *S. endobioticum* has a reduced set of CWDEs, and it is likely that *S. endobioticum* employs a stealth-like strategy to evade detection and establish its intracellular biotrophic lifestyle.

Inside the host cell

After penetration of the host, the fungal thallus secretes effector proteins to manipulate the host cell and induce hypertrophy and hyperplasia of potato tissue. For *S. endobioticum*, it is still not known if the fungal thallus is invaginated by the host plasma membrane similarly to haustoria in rust fungi and powdery mildew [23], or if it resides inside the host cytoplasm. Hyperplasia increases the amount of meristematic tissue which enhances the chance of

Chapter 6

re-infection of the host. In response to the invading pathogen, the host has evolved receptors encoded by plant *R* genes that specifically recognize the fungal effector proteins. Upon recognition, a hypersensitive response (HR) is induced resulting in local cell death, which restricts further pathogen proliferation. The genes encoding effectors that are recognized by plant *R* gene products and mediate effector triggered immunity (ETI), are referred to as avirulence (*Avr*) genes. The *AvrSen1* gene product is recognized in pathotype 1(D1) resistant potatoes carrying the *Sen1* gene. With the identification of the *AvrSen1*, we have now the first example of a *S. endobioticum* effector that is recognized by a potato *R* gene (chapter 5).

As zoospores infect resistant varieties and susceptible plants equally well, no differentiation between different pathotypes is expected in the effector repertoire used to suppress PAMP-triggered immunity in the extracellular phase. However, in an incompatible interaction (e.g. *AvrSen1/Sen1*) the pathogen dies within a few hours after penetration of the host as a result of what was described as a pathogen-induced defense necrosis [4, 24]. We hypothesized this is the effect of race-specific effectors that are secreted by the pathogen intracellularly (chapters 1, 5). Indeed, *AvrSen1* was recognized intracellularly, and was found to be present in pathotype 1(D1) isolates but functionally absent in the higher pathotypes. Furthermore, five different variants resulting in loss of function of *AvrSen1* were observed in higher pathotypes suggesting that the gene is under strong selective pressure. The observation that the fungus can penetrate host cells of resistant potato varieties dates from 1926, and at that time no differentiation between pathotypes was known. Most likely, a pathotype 1(D1) isolate of the pathogen was used for inoculation at that time. Whether higher pathotypes also have the ability to penetrate the host remains to be determined.

Mitochondrial genomic variation showed that *S. endobioticum* isolates represent communities of different mitochondrial genotypes (chapter 2). Similarly, the avirulent *AvrSen1* gene was found to be present in several higher pathotypes, be it at very low frequencies, alongside the virulent form (chapter 5). Apparently, it is maintained at low levels in the virulent fungal population, which could suggest it has a function in virulence. As many other fungal effectors [12], *AvrSen1* lacks functional domains or a predicted function. The absence of nuclear, mitochondrial or chloroplast localization signals suggests a cytoplasmic target (chapter 5).

In the Canadian pathotype 6(O1) isolate the truncated *AvrSen1* gene is expressed *in vivo* (chapter 2). The truncated *AvrSen1* gene products did not trigger an immune response

in *Sen1* plants (chapter 5). Whether the truncated gene models that are present in other isolates are also expressed, and whether they have a potential role in the plant-pathogen interaction (e.g. as *Avrs* for other *R* genes, or as functional homologs of *AvrSen1*) are currently unknown.

Pathotypes, pathotype formation and diversity

Isolates of *S. endobioticum* are grouped in pathotypes based on their response on a set of differential resistant potato varieties. Two main methods are used to assess the interaction between host and pathogen, i.e. Spieckermann [25] and Glynne-Lemmerzähl [26, 27], and both are based on the inoculation of potato tubers. In the Spieckermann assay resting spores, which are the result of a (para)sexual cycle, are used as inoculum. In contrast, fresh warts containing summer sporangia, which are produced asexually, are used as inoculum in the Glynne-Lemmerzähl assay (chapter 1). This pathotype determination of isolates is fundamental to phytosanitary measures in which resistant potato varieties are used to prevent the spread of the organism.

The use of different resistant varieties in different countries has led to a rapid increase of newly described pathotypes in the past [28]. Lack of harmonization of the differential potato varieties used, and differences in the interpretation of observed phenotypes hampered reproducibility and verification of pathotyping results [29]. With the publication of a diagnostic standard for *S. endobioticum* by the European and Mediterranean Plant Protection Organization (EPPO), a harmonized differential set of potato varieties was introduced [30]. However, the diagnostic resolution of this differential set was limited, as can be deduced from the overview presented by Baayen *et al.* (2006) in which it is shown that further differentiation among *S. endobioticum* isolates is possible when additional potato varieties are used [29]. Currently, several *R* genes are not represented in the EPPO differential set [31]. Pathotypes 1(D1), 2(G1), 6(O1), 8(F1) and 18(T1) are believed to be the most important pathotypes in Europe. However, with potato varieties described in the EPPO standard other diversity largely remains undetected (chapter 3). Furthermore, with the update of the EPPO diagnostic standard for *S. endobioticum*, potato cultivar Miriam was removed from the differential set [32]. As a result, there is now no potato variety included in the set to differentiate between pathotype 8(F1) and 6(O1). Consequently, the current pathotype identity does not sufficiently reflect the virulence potential of *S. endobioticum* isolates which can undermine the effectiveness of phytosanitary measures.

Genetic variation between isolates with the same phenotype

From the work presented in chapters 3 and 5 we now have empiric data supporting the hypothesis of genetic variation among isolates exhibiting the same phenotype. Three major *S. endobioticum* mitochondrial lineages, referred to as haplogroups, are present in Europe and Canada, and a fourth haplogroup was identified in an isolate of Peruvian origin. The occurrence of these mitochondrial lineages are likely the result of separate introductions, possibly from the Andes which is presumed region of origin. We noted that pathotypes 2(G1) and 6(O1) occurred in different haplogroups. Therefore, these higher pathotypes emerged independently from different genetic backgrounds (chapter 3). This is further corroborated by the clustering analysis of the species specific secretome and the different types of loss of the *AvrSen1* gene (chapter 5). The pathotype 2(G1) and 6(O1) isolates from the different haplogroups have evolved their own set of effectors and possibly *Avr* genes which are not detected in the current pathotyping methods as the cognate *R* genes are not included in the current differential potato panel.

Genetic variation within *S. endobioticum* isolates

S. endobioticum isolates represent communities in which multiple genotypes co-exist. This does not only holds true for mitochondrial genotypes as observed in chapter 3, but also for variation within the nuclear encoded species specific secretome and the *AvrSen1* gene as reported in chapter 5. Particularly striking is the *AvrSen1* variation found within isolates of higher pathotypes. These isolates are no longer recognized by the *Sen1 R* gene, resulting in a compatible interaction and proliferation of the pathogen. However, the avirulent form of the gene was still found to be present at low frequencies in isolates that were multiplied on the potato variety Deodara, which is susceptible to all *S. endobioticum* pathotypes (chapter 5). *Vice versa*, in pathotype 1(D1) isolates carrying the avirulent *AvrSen1* gene, virulent forms of the gene (i.e. loss of *AvrSen1*) were present (chapter 5). This was not directly apparent from genomic sequence data, but multiplications of a pathotype 1(D1) isolate on the semi-resistant potato variety Erika resulted in increased virulence as a result of selection against genotypes carrying the avirulence gene.

After two multiplications on this potato variety, the pathogen produced a pathotype 6(O1) phenotype (chapter 3). Also, the within-isolate diversity of mitochondrial genotypes increased in the 6(O1) isolate relative to the original 1(D1) isolate. Possibly this is a result of selecting against the main genotype (chapter 3). Loss of *AvrSen1* in the isolate producing the

6(O1) phenotype was verified by amplification of the *AvrSen1* gene followed by PacBio SMRT sequencing (chapter 5). Multiplications leading to the observed shift in phenotype were performed with the Glynne-Lemmerzahl bioassay, in which asexually produced zoospores from fresh warts are used as inoculum (chapters 1, 3). In contrast, the Spieckermann assay uses resting spores as inoculum that are the result of a (para)sexual cycle. The use of the Glynne-Lemmerzahl assay suggests that the diversity was already present within the initial inoculum, and was not generated by recombination in a (para)sexual cycle.

After successful deployment of resistant potato varieties and control measures, absence of the disease has been reported in several countries for periods of time, such as the Netherlands [29], Poland [33] and Denmark [34]. However, new outbreaks were reported from these countries, and where previously infections with pathotype 1(D1) were encountered, infections with higher pathotypes were found in neighboring locations after brief periods of reported absence of wart disease. The occurrence of higher pathotypes could be the result of the introduction of new pathotypes from other infested regions, as was shown in chapter 3, or the result of remodeling of the pathogen population induced by cultivation of resistant potato varieties that do not offer resistance to the entire population either by a selective sweep or by recombination in a (para)sexual cycle. Mitochondrial haplotyping of historic samples can be used to further determine the history of potato wart disease introductions and spread, but also the occurrence of selective sweeps that could have led to increased virulence in pathogen isolates.

The role of alternative hosts

An intriguing thought is the role that alternative hosts could play in the longevity of the pathogen. In absence of potato cultivation, resting spores have been reported to remain viable and infectious under field conditions up to 45 years [35]. In contrast, resting spores in composts that are maintained under (semi) controlled climate conditions in *S. endobioticum* collections for diagnostic and research purposes rarely remain viable beyond 5 to 10 years (Gerard van Leeuwen, NVWA, the Netherlands; personal communication). The question to what extent alternative hosts in field borders can act as a reservoir in absence of potato cultivation is as relevant today as when it was first suggested in 1916 [36]. If they could act as a reservoir, can they remodel the fungal population as was observed in potato variety Erika (chapter 3)?

Chapter 6

Currently, no potato cultivation is allowed on infested plots for twenty years on account of the longevity of the *S. endobioticum* resting spores. When alternative hosts prolong the viability and infectiousness of spores in the soil, eradication of these hosts could be included in phytosanitary control measures. This could potentially lead to a reduction of the period infested plots are scheduled.

The repeat repertoire as driving force in genome plasticity and pathogenicity

Several factors play a role in the speed with which pathogens can adapt to *R* genes, including population size, dissemination abilities, mode of reproduction, and effector mutation rates [37]. In several pathogenic fungi the latter has been attributed to a bipartite genome architecture, referred to as the “two-speed genome” [38]. Although appealing, a two-speed genome might be an oversimplification and the genomes of fungal plant pathogens could display more gradual levels of mutation rates [39]. In these multi-speed genomes, chromosomes consist of conserved and dynamic regions, alternatively these dynamic regions are physically separated on dispensable supernumerary chromosomes that can be lineage specific. Effector genes are often found in the dispensable parts of the genome, and are associated with transposable elements (TEs) which act as drivers of rapid relocation, duplication and subsequent diversification [40-42].

Saprobic and other pathogenic chytrid species analyzed in chapter 2, had comparable genome sizes as *S. endobioticum*, but their repeat content and the repeat repertoires differed dramatically. Indeed, both pathogenic chytrid species in our analysis, *S. endobioticum* and *B. dendrobatidis*, contained the highest repeat repertoire among the eleven chytrid genomes currently available. In addition, the repeat repertoire of both pathogens mainly consisted of complex repeats of which a small number could be identified as known class I or II transposable elements. The complex repeat repertoire of both *S. endobioticum* isolates displays the highest diversity among the chytrid species studied, and the species specific repertoire is 17 times higher compared to *B. dendrobatidis*. This suggests that the genomes of both species were invaded by transposable elements which might increase genome plasticity attributed to pathogenic species [41], but that the elements involved are different and that new TEs remain to be found in the genomes of the pathogenic chytrids.

In *Blumeria graminis*, causing powdery mildew on barley, loss of the fungal defense mechanism against TEs, i.e. repeat induced point mutations (RIP), has been suggested to be

advantageous for pathogenic species as it can speed up genome evolution [43]. However, it is unlikely that this could explain the differences in repeat repertoire between the pathogenic and saprobic chytrids, as none of the chytrid species analyzed in chapter 2 appear to possess RIP activity.

It is of interest to study the repeat repertoire of *Synchytrium* species in a phylogenetic context as all species in the genus, except for *S. microbalum*, are regarded to be obligate biotrophic parasites. If the saprobic *S. microbalum* is basal to the obligate biotrophic *Synchytrium* clade, the genome of the last common ancestor of the parasitic *Synchytrium* species could have been invaded with genetic elements such as TEs [44]. Alternatively, aneuploidy (i.e. the presence of unequal numbers of chromosomes) as result of a (para)sexual cycle could have led to the formation of supernumerary chromosomes in *Synchytrium* species. The 18S rDNA phylogeny published by Longcore *et al.* (2016) contradicts the hypothesis of such an event in the last common ancestor of the pathogenic *Synchytrium* species as it places the pathogenic *Synchytrium taraxaci* basal at the *Synchytrium* clade and not *S. microbalum* [45]. However, our phylogeny based on the protein coding mitochondrial genes (chapter 3) is in agreement with the hypothesis as it places *S. microbalum* basal among the *Synchytrium* species tested, which included *S. taraxaci*.

Opportunities and challenges for future *S. endobioticum* research

Here I present opportunities and challenges for future potato wart disease research, with emphasis on pathogen related research. Even though these activities are distinct, they are tightly interlinked (Fig. 1).

Towards a new pathotype grouping system

As already mentioned, the current *S. endobioticum* pathotyping system does not reflect the full range of host *R* genes and the cognate fungal effectors, which undermines phytosanitary control that depends on correct pathotype identity. For *S. endobioticum* pathotyping, a transition from grouping isolates based on intuitively selected potato varieties to a differential panel based on the presence of *R* genes will improve comparison and interpretation of pathotyping results, and contribute to durable potato wart resistance breeding and the efficacy of phytosanitary control measures.

Chapter 6

A similar transition has been described for *Phytophthora infestans* [46]. In this pathosystem, researchers intuitively selected differentially resistant varieties to define races and underlying *R* genes in an era where no molecular knowledge was available. Under the old system, a differential set was built based on eleven late blight resistance specificities (R_1 - R_{11}), and these were used to determine the *P. infestans* races [47]. Since, the underlying *R* genes and matching effectors are known, a single *R* gene differential set could be built for an unambiguous pathotyping system [48]. For instance, a *P. infestans* isolate that was able to infect potato clones with R_1 and R_2 but not R_3 - R_{11} was named “race 1,2”. This naming system is flexible and allows for expansion when new *R* genes are identified [49, 50]. A similar system could be envisioned for *S. endobioticum* (Fig. 1). For instance, an isolate that is able to infect potato varieties with the *Sen1* *R* gene, but not varieties with *Sen2* and *Sen3* could be named *S. endobioticum* race V1. In this naming system, “V” refers to virulence group, and numeral(s) refer to the *R* gene(s) that do not offer resistance.

Development of such a system would require a broad inventory of major *R* genes beyond the currently described *Sen1*, *Sen2* and *Sen3* [31, 51, 52], and producing a differential set with potatoes possessing a single *R* gene (Fig. 1). Pathogen isolates tested with these

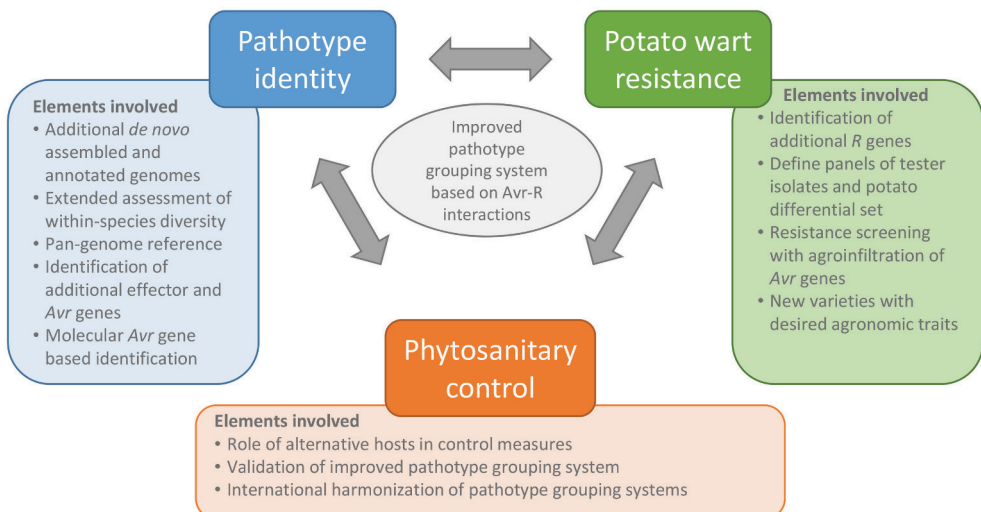


Figure 1. Suggested scheme to address potato wart research opportunities and elements involved. Central in this scheme is an unambiguous pathotype grouping system based on Avr-R interactions that will contribute to the improvement of durable potato wart resistance breeding and phytosanitary control.

potato genotypes pathogen can be grouped based on specific Avr-R interactions providing additional diagnostic resolution. With an improved, unambiguous grouping system, it will be easier to identify the cognate *Avr* genes to the potato *R* genes and *vice versa*.

The improved pathotype grouping system could also further improve international acceptance of new potato varieties. With the current bioassays, contradictory results for “potato variety-pathotype” combinations are sometimes obtained from different laboratories. This could be explained by the use of fungal isolates in the individual labs that have the same pathotype identity based on the current potato differential set, but with different underlying genotypes. These genotypic differences could result in a higher virulence potential in one of these isolates. As a consequence, a compatible interaction could be observed in one lab, whereas an incompatible interaction could be observed in the other. This implies that not only potato varieties in resistance testing need to be standardized, but also the fungal isolates used for inoculation. Producing sets of more homogeneous tester isolates requires strict regulations for potato varieties to propagate them on, and the composition of these isolates needs to be monitored regularly with next generation sequencing to detect potential disruptive selection.

As the current pathotyping systems are so embedded in phytosanitary control and legislation, it will require extensive testing and validation before an alternative pathotyping system can be accepted (Fig. 1).

Molecular pathotyping and alternative resistance screening of potato varieties

Since the introduction of molecular methods in potato wart research, authors have attempted to devise assays for molecular pathotyping based on associated markers. With the lessons learned from this thesis, we can conclude that these approaches will not work for two main reasons: 1. the current pathotyping system groups *S. endobioticum* isolates that should be further separated based on their effector repertoire, and 2. the subtle types of loss observed in *AvrSen1* will not be detected with associated markers. Even designing specific TaqMan assays for functional markers such as the *AvrSen1* gene will prove to be very difficult, and may even be impossible on account of absence of polymorphisms shared between virulent forms relative to the avirulent form. In time, a next generation sequencing approach targeting *Avr* genes as demonstrated in chapter 5 or whole genome sequencing, are likely the most robust methods for molecular pathotyping (Fig. 1).

Chapter 6

Screening for potato wart resistance with the tuber-based Spieckermann and Glynne-Lemmerzahl bioassays is a time-consuming and labor-intensive task. Because of the high number of escapes associated with these bioassays, large numbers of potato tubers need to be tested to assess their resistance for a set of *S. endobioticum* pathotypes. It takes at least two years to obtain sufficient tubers of a new variety for *S. endobioticum* pathotype resistance screening. The aboveground bioassay described in chapter 4 could reduce this to one year as clonally propagated cuttings from seedlings can be used in the assay. Also, non-tuber bearing Solanaceous species or species with poor tuber formation can be screened for potato wart resistance as no tubers are required, unlocking the potential of these species to screen for new potato wart resistance genes.

With the identification of *Avr* genes, the speed of screening progeny of potato crosses for *S. endobioticum* resistance can be even further increased by expressing pathogen effectors in potato leaves through agroinfiltration. In chapter 4 we demonstrate that plant *R* genes are equally expressed in aboveground plant parts compared to the belowground etiolated tuber sprouts. Screening for *Sen1* based pathotype 1(D1) resistance can already be performed with the *AvrSen1* gene, and with the discovery of additional *Avr* genes, this can be extended (Fig. 1).

Identification of additional effectors

By combining a clear *R* gene-based search profile with high quality datasets of fungal isolates representing multiple pathotypes and a high quality functionally annotated pathotype 1(D1) genome, we were able to identify the *AvrSen1* gene (chapter 5). This strategy can be applied to other *S. endobioticum* effectors with a hypothesized differential presence/absence pattern between isolates under effect of a *R* gene. For instance, potato variety Saphir (included in the EPPO differential set) provides resistance to all pathotypes of major importance in Europe, except for pathotype 2(G1) [32]. Even though the Saphir *R* gene is not known, the cognate *Avr* is expected to be present in the pathotype 1(D1), 6(O1), 8(F1) and 18(T1) genomes, but not in the pathotype 2(G1). Both the pathotype 1(D1) and 6(O1) genomes can serve a reference to identify the *Avr* gene cognate to the Saphir *R* gene.

Three major resistance genes have been reported providing qualitative resistance to *S. endobioticum*: *Sen1* providing resistance to pathotype 1(D1) isolates, and *Sen2* and *Sen3* offering broad resistance against pathotype 1(D1) isolates and tested higher pathotypes [31, 51, 52]. For *Sen2* and *Sen3*, the strategy mentioned above is not expected to be effective

as no differences between *S. endobioticum* isolates are anticipated as both *R* genes provide broad resistance to all tested pathotypes [31, 52]. As an alternative, the selection of “resistance breaking” isolates, in the laboratory or in the field, could be a way to identify the *Avrs* cognate to *Sen2* and *Sen3* through comparative analysis.

For the screening of *AvrSen1* candidates, the pathotype 1(D1) genome was used as reference in a read-mapping approach to detect loss of function of genes in *S. endobioticum* isolates (chapter 5). Even though the approach was successful for the identification of *AvrSen1*, it is recommended to use *de novo* assembled and annotated genomes of different pathotypes in a comparative approach for the identification of new effectors. Using *de novo* assembled genomes allows to overcome several technical difficulties associated with read-mapping approaches such as establishing detection thresholds and the identification of splice variants. Also, this allows the inclusion of pathotype-specific genes that are absent in the reference genome. Currently, over a hundred *S. endobioticum* isolates from the potato wart collection of the Dutch National Plant Protection Organization are sequenced. With the *de novo* assembled and annotated sequences of these isolates, a pan-genome reference could be build to facilitate enhanced effector and *Avr* identification.

Again, a word of caution must be uttered here as some isolates with the same phenotype emerged from different genetic backgrounds and could have a different underlying effector repertoire which is currently unnoticed (chapters 3, 5). A screening of within-species diversity using the mitochondrial genome could be instrumental to identify representative isolates to include in pathotyping assays and genome assembly. It is recommended to include at least one, but preferably multiple, isolates of each mitochondrial haplogroup-pathotype combination.

AvrSen1 candidacy was based on loss of gene function relative to the reference genome, and these absences were based on gene coverage and non-synonymous substitutions (chapter 5). In addition, alternative splicing, changes in promotor sites, and differences in gene expression could be explored to identify loss of function of effector genes in *S. endobioticum* isolates. With the aboveground bioassay described in chapter 4, cleaner starting material can be produced compared to tuber-based bioassays, with which differential expression between isolates could be studied more efficiently. In addition, plant responses to the invading pathogen can be monitored with these materials. This opens up possibilities to study the influence of the pathogen to, for instance, phytohormone pathways during proliferation in the host.

Chapter 6

Apart from major resistance genes, several quantitative resistance loci (QRLs) offering partial resistance to *S. endobioticum* pathotypes were identified in potato [53, 54]. Quantitative resistance is governed by many genes which act together in pathogen defense, however, independently these genes might provide only partial resistance. Genes underlying QRLs resemble both known components of the innate immune system, but also have functions previously not associated with disease resistance [55]. Due to its polygenic nature, quantitative resistance is in some cases regarded more durable [56]. However, as quantitative resistance does not offer full resistance, this will not provide sufficient protection against quarantine organisms (such as *S. endobioticum*) for which a zero-tolerance policy is enforced. Altogether, identifying *S. endobioticum* genes involved in triggering quantitative resistance could prove to be much more challenging compared to fungal effectors that are recognized by specific major *R* genes that follow the gene-for-gene concept described by Flor [57].

One group of genes that deserves additional attention with respect to their potential role as effectors are genes encoding RAYH-proteins (chapter 2). The name of these proteins is derived from a highly conserved 15 amino acid motif, and they display several hallmarks attributed to effectors: i.e. they reside predominantly in the species specific secretome, originate from multi-gene families with pathotype specific expansions, are found overrepresented in regions that are depleted of chytrid core genes, and mitochondrial and nuclear localization signals are overrepresented in the RAYH-proteins. The RAYH-proteins are numerous present in the two *S. endobioticum* genomes, and access to additional annotated *S. endobioticum* genomes will provide more insight in the diversity of these proteins and their association with pathotypes.

Opportunities from new sequencing technologies

On account of the high repeat content of both *S. endobioticum* genomes and the fact that multiple genotypes are sequenced to create a single consensus sequence, our current genome assemblies are still fairly fragmented (chapter 2). To better understand the genome dynamics in *S. endobioticum* isolates, and to determine if the multi-speed genome concept also applies to this plant pathogen, efforts should be made to further scaffold and complete the fragmented assemblies. Sequencing technologies such as Oxford Nanopore and 10x Illumina sequencing have proven instrumental in improving assemblies of complex and repeat-rich genomes [58, 59], and could be explored for *S. endobioticum*. However, extracting high molecular weight DNA from *S. endobioticum* has proven to be difficult, and

resting spores used for DNA and RNA extraction represent metagenomic samples with multiple *S. endobioticum* genotypes. Again, the aboveground bioassay (chapter 4) could be used to create cleaner starting material for molecular studies. Nucleic acid extraction from single resting spores followed by whole genome amplification and sequencing could reduce the complexity of the initial biological sample (zoospores from different resting spores represent different genotypes) and facilitate improvement of the genome assembly.

Opportunities from other *Synchytrium* species

In our efforts to sequence multiple chytrid genomes, we have attempted to assemble the genome of *S. taraxaci*, an obligate biotroph on dandelion and the type species to the genus. Where the assembly, selection of scaffolds from the metagenomic sequences and annotation of *S. endobioticum* genome was challenging, we were not successful in obtaining a high quality annotated genome for *S. taraxaci* on account of the highly fragmented assembly and poor gene annotation. It is likely that the challenges observed for *S. endobioticum* (i.e. repeat-richness and samples represent multiple genotypes) also apply to *S. taraxaci*. It is worthwhile to invest in generating high quality genomes of additional *Synchytrium* spp. to properly assess the phylogenetic relationships in the genus. This allows the examination of molecular features linked to pathogenicity and lifestyle in a phylogenetic context. For instance, RAYH-proteins were not found in the saprobic *S. microbalum* (chapter 2), but are they present in other pathogenic *Synchytrium* species? Do pathogenic *Synchytrium* species share the species specific repeat repertoire found in *S. endobioticum* (chapter 2)? Also, are there similarities or differences in the molecular mechanism underlying the induction of hyperplasia and hypertrophy in host plants between the pathogenic *Synchytrium* species? In the potato-*S. endobioticum* interaction, this reaction is very strong while it is mostly absent in other susceptible hosts (chapter 1).

Concluding remarks

With the work presented in this thesis we have made important contributions to potato wart research, created a solid foundation for genomic studies of *S. endobioticum*, and demonstrated several applications. Also, with our comparative genomics approach we contributed significantly to the knowledge on Chytridiomycota, a basal fungal lineage that is notoriously underrepresented in genomic studies. We have reported on several commonalities between plant pathogenic chytrids, ascomycetes, basidiomycetes and oomycetes.

Chapter 6

Through comparative genomics we have been able to study aspects linked to obligate biotrophy and pathogenicity in *S. endobioticum* where researchers in the past simply did not have the tools to do so. The functionally annotated chytrid genomes generated as part of this thesis can now be included in comparative fungal genomic studies on a systematic basis. They provide a unique phylogenetic view and can shed new light on fundamental aspects of fungal biology.

I have presented several lines of inquiry in the “opportunities and challenges” section of this chapter, and to address these, additional research on the pathogen and the host is needed, which is clearly interlinked (Fig. 1). Central in the future activities presented is an improved *Avr-R* gene based pathotyping system that will contribute to the improvement of durable potato wart resistance breeding and phytosanitary control. Also, without culturing options, the further use of heterologous systems, such as expression of *S. endobioticum* genes in *Saccharomyces cerevisiae* [60], could be explored. Maybe it will prove too challenging, but having culturing options for *S. endobioticum* would represent a major addition to our research toolbox. Perhaps the supplementation of growth media with guanine could hold the key to culturing *S. endobioticum*. In yeast, the lethal effects of knocked-out genes from the purine pathway, which were found absent in *S. endobioticum* (chapter 2), could be reversed by the addition of guanine [61].

As the potato wart research field represents only a limited number of partners compared to other fields in phytopathology, close collaboration between breeding companies, research institutes and governmental laboratories is needed to maximize the impact of potato wart research. In Europe, regional plant health reference laboratories (EURLs) are currently being established under the new EU regulation 2017/625 [62]. The to be established mycological EURL could play an important role in identifying milestones towards answering some of the big questions that remain unanswered, such as the development of new pathotyping concepts and their acceptance by private companies, the scientific community, and governments.

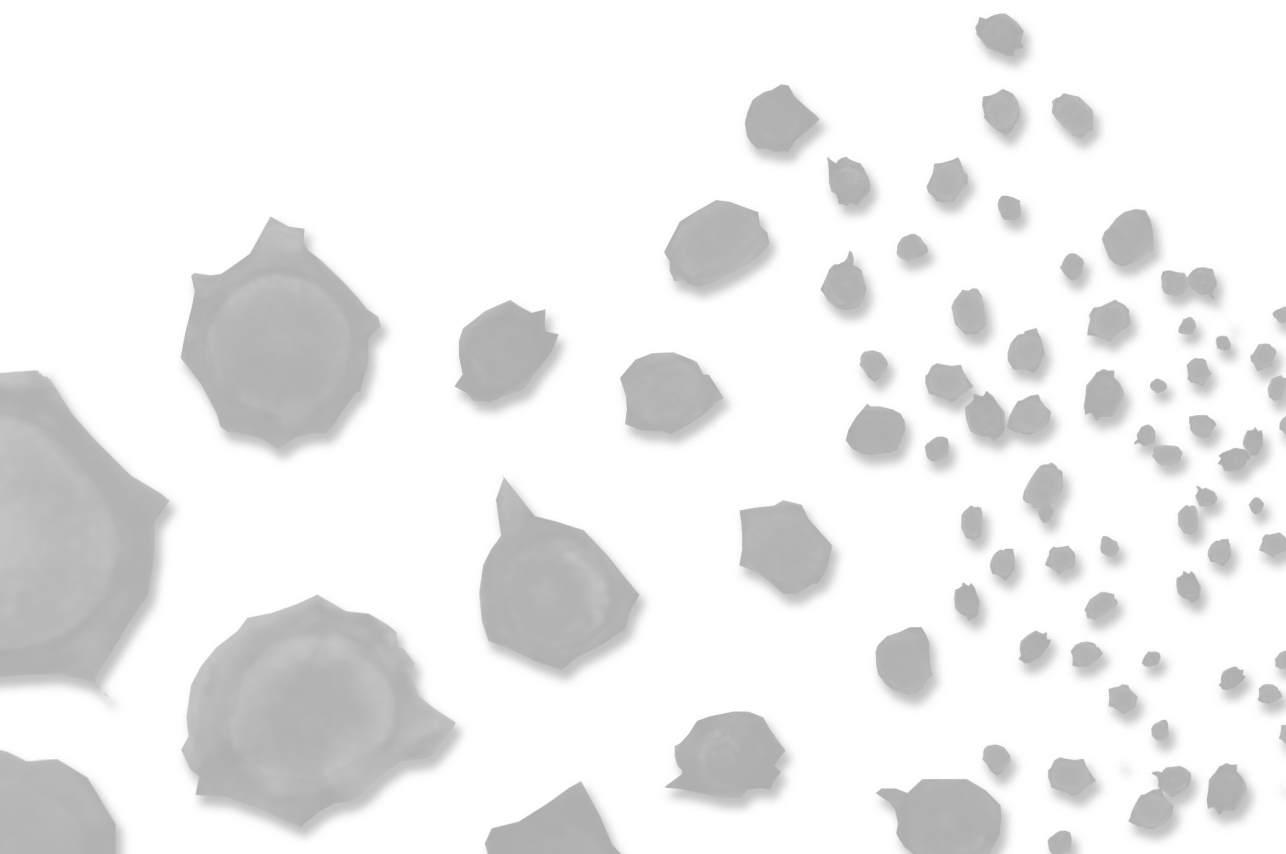
References

1. Sparrow, F.K.C.F.p.d.N.D., Interrelationships and Phylogeny of the Aquatic Phycomycetes. *Mycologia*, 1958. 50(6): p. 797-813.
2. Sharpee, W.C. and R.A. Dean, Form and function of fungal and oomycete effectors. *Fungal Biology Reviews*, 2016. 30(2): p. 62-73.
3. Curtis, K.M., IX.— The life-history and cytology of *Synchytrium endobioticum* (schilb.), perc., the cause of wart disease in potato. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character*, 1921. 210(372-381): p. 409.
4. Cartwright, K., On the Nature of the Resistance of the Potato to Wart Disease. *Annals of Botany*, 1926. 40(158): p. 391-395.
5. Lange, L. and L.W. Olson, Development of the resting sporangia of *Synchytrium endobioticum*, the causal agent of potato wart disease. *Protoplasma*, 1981. 106(1): p. 83-95.
6. Lange, L. and L.W. Olson, Germination and parasitism of the resting sporangia of *Synchytrium endobioticum*. *Protoplasma*, 1981. 106(1): p. 69-82.
7. Lange, L. and L.W. Olson, Development of the zoosporangia of *Synchytrium endobioticum*, the causal agent of potato wart disease. *Protoplasma*, 1981. 106(1): p. 97-108.
8. Lange, L. and L.W. Olson, The zoospore of *Synchytrium endobioticum*. *Canadian Journal of Botany*, 1978. 56(10): p. 1229-1239.
9. Hampson, M.C., Sequence of events in the germination of the resting spore of *Synchytrium endobioticum*, European pathotype 2, the causal agent of potato wart disease. *Canadian Journal of Botany*, 1986. 64(9): p. 2144-2150.
10. Carella, P., E. Evangelisti, and S. Schornack, Sticking to it: phytopathogen effector molecules may converge on evolutionarily conserved host targets in green plants. *Curr Opin Plant Biol*, 2018. 44: p. 175-180.
11. Selin, C., *et al.*, Elucidating the Role of Effectors in Plant-Fungal Interactions: Progress and Challenges. *Frontiers in Microbiology*, 2016. 7(600).
12. Lo Presti, L., *et al.*, Fungal effectors and plant susceptibility. *Annu Rev Plant Biol*, 2015. 66: p. 513-45.
13. De Wit, P.J.G.M., *et al.*, Fungal effector proteins: past, present and future. *Molecular Plant Pathology*, 2009. 10(6): p. 735-747.
14. Islam, M.T. and S. Tahara, Chemotaxis of fungal zoospores, with special reference to *Aphanomyces cochlioides*. *Biosci Biotechnol Biochem*, 2001. 65(9): p. 1933-48.
15. Obidiegwu, J.E., K. Flath, and C. Gebhardt, Managing potato wart: a review of present research status and future perspective. *Theor Appl Genet*, 2014. 127(4): p. 763-80.
16. Xue, C., Y.-P. Hsueh, and J. Heitman, Magnificent seven: roles of G protein-coupled receptors in extracellular sensing in fungi. *FEMS microbiology reviews*, 2008. 32(6): p. 1010-1032.
17. Kombrink, A., A. Sánchez-Vallet, and B.P.H.J. Thomma, The role of chitin detection in plant-pathogen interactions. *Microbes and Infection*, 2011. 13(14): p. 1168-1176.
18. Kombrink, A. and B.P.H.J. Thomma, LysM Effectors: Secreted Proteins Supporting Fungal Life. *PLOS Pathogens*, 2013. 9(12): p. e1003769.
19. Gust, A.A., R. Pruitt, and T. Nurnberger, Sensing Danger: Key to Activating Plant Immunity. *Trends Plant Sci*, 2017. 22(9): p. 779-791.
20. Goodwin, S.B., *et al.*, Finished Genome of the Fungal Wheat Pathogen *Mycosphaerella graminicola* Reveals

Chapter 6

- Dispensome Structure, Chromosome Plasticity, and Stealth Pathogenesis. *PLOS Genetics*, 2011. 7(6): p. e1002070.
21. Martin, F., *et al.*, The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature*, 2008. 452: p. 88.
 22. Choi, H.W. and D.F. Klessig, DAMPs, MAMPs, and NAMPs in plant innate immunity. *BMC Plant Biology*, 2016. 16: p. 232.
 23. Szabo, L.J. and W.R. Bushnell, Hidden robbers: The role of fungal haustoria in parasitism of plants. *Proceedings of the National Academy of Sciences*, 2001. 98(14): p. 7654.
 24. Karling, J.S., *Synchytrium*. 1 ed. 1964: Academic Press, New York.
 25. Spieckermann, A.K., P., Testing potatoes for wart resistance. *Deutsche Landwirtschaftliche Presse*, 1924. 51: p. 114-115.
 26. Glynne, M.D., Infection experiments with wart disease of potatoes. *Synchytrium endobioticum* (Schilb.) Perc. *Annals of Applied Biology*, 1925. 12(1): p. 34-60.
 27. Lemmerz, J., Neues vereinfachtes Infektionsverfahren zur Prüfung von Kartoffelsorten auf Krebsfestigkeit. *Der Züchter*, 1930. 2: p. 288-297.
 28. Hampson, M.C., History, biology and control of potato wart disease in Canada. *Canadian Journal of Plant Pathology*, 1993. 15(4): p. 223-244.
 29. Baayen, R.P., *et al.*, History of potato wart disease in Europe - a proposal for harmonisation in defining pathotypes. *European Journal of Plant Pathology*, 2006. 116(1): p. 21-31.
 30. OEPP/EPPO, PM 7/28 (1) *Synchytrium endobioticum*. *EPPO Bulletin*, 2004. 34(2): p. 213-218.
 31. Prodhomme, C., *et al.*, Comparative Subreads Sets Analysis (CoSSA) is a robust approach to identify haplotype specific SNPs; Mapping and pedigree analysis of a potato wart disease resistance gene *Sen3*. *Plant Methods*, unpublished.
 32. OEPP/EPPO, PM 7/28 (2) *Synchytrium endobioticum*. *EPPO Bulletin*, 2017. 47(3): p. 420-440.
 33. Przetakiewicz, J., Assessment of the resistance of potato cultivars to *Synchytrium endobioticum* (Schilb.) Perc. in Poland. *EPPO Bulletin*, 2008. 38(2): p. 211-215.
 34. Anonymous, *Synchytrium endobioticum* found in Denmark. *EPPO Reporting Service*, 2014. 11(206).
 35. Przetakiewicz, J., The Viability of Winter Sporangia of *Synchytrium endobioticum* (Schilb.) Perc. from Poland. *American Journal of Potato Research*, 2015. 92(6): p. 704-708.
 36. Cotton, A.D.C., Host Plants of *Synchytrium endobioticum*. *Studies from the Pathological Laboratory: IV*. Royal Botanic Gardens, 1916. 1916(10): p. 272-275.
 37. McDonald, B.A. and C. Linde, Pathogen population genetics, evolutionary potential, and durable resistance. *Annu Rev Phytopathol*, 2002. 40: p. 349-79.
 38. Raffaele, S. and S. Kamoun, Genome evolution in filamentous plant pathogens: why bigger can be better. *Nat Rev Microbiol*, 2012. 10(6): p. 417-30.
 39. Fokkens, L., *et al.*, The multi-speed genome of *Fusarium oxysporum* reveals association of histone modifications with sequence divergence and footprints of past horizontal chromosome transfer events. *bioRxiv*, 2018: p. 465070.
 40. Kemen, A.C., M.T. Agler, and E. Kemen, Host-microbe and microbe-microbe interactions in the evolution of obligate plant parasitism. *New Phytol*, 2015. 206(4): p. 1207-28.
 41. Stajich, J.E., *Fungal Genomes and Insights into the Evolution of the Kingdom*. *Microbiol Spectr*, 2017. 5(4).
 42. Sánchez-Vallet, A., *et al.*, The Genome Biology of Effector Gene Evolution in Filamentous Plant Pathogens.

- Annual Review of Phytopathology, 2016.
43. Werren, J.H., Selfish genetic elements, genetic conflict, and evolutionary innovation. Proceedings of the National Academy of Sciences, 2011. 108(Supplement 2): p. 10863.
 44. Kemen, E. and J.D. Jones, Obligate biotroph parasitism: can we link genomes to lifestyles? Trends Plant Sci, 2012. 17(8): p. 448-57.
 45. Longcore, J.E., D.R. Simmons, and P.M. Letcher, *Synchytrium microbalum* sp. nov. is a saprobic species in a lineage of parasites. Fungal Biology, 2016. 120(9): p. 1156-1164.
 46. Black, W., et al., A proposal for an international nomenclature of races of *Phytophthora infestans* and of genes controlling immunity in *Solanum demissum* derivatives. Euphytica, 1953. 2(3): p. 173-179.
 47. Cooke, D.E.L. and A.K. Lees, Markers, old and new, for examining *Phytophthora infestans* diversity. Plant Pathology, 2004. 53(6): p. 692-704.
 48. Zhu, S., et al., An updated conventional- and a novel GM potato late blight *R* gene differential set for virulence monitoring of *Phytophthora infestans*. Euphytica, 2015. 202(2): p. 219-234.
 49. Malcolmson, J.F. and W. Black, New *R* genes in *Solanum demissum* Lindl. and their complementary races of *Phytophthora infestans* (Mont.) de Bary. Euphytica, 1966. 15(2): p. 199-203.
 50. Malcolmson, J.F., Races of *Phytophthora infestans* occurring in Great Britain. Transactions of the British Mycological Society, 1969. 53(3): p. 417-IN2.
 51. Hehl, R., et al., TMV resistance gene *N* homologues are linked to *Synchytrium endobioticum* resistance in potato. Theoretical and Applied Genetics, 1999. 98(3): p. 379-386.
 52. Plich, J., et al., Novel gene *Sen2* conferring broad-spectrum resistance to *Synchytrium endobioticum* mapped to potato chromosome XI. Theoretical and Applied Genetics, 2018. 131(11): p. 2321-2331.
 53. Ballvora, A., et al., Multiple alleles for resistance and susceptibility modulate the defense response in the interaction of tetraploid potato (*Solanum tuberosum*) with *Synchytrium endobioticum* pathotypes 1, 2, 6 and 18. Theor Appl Genet, 2011. 123(8): p. 1281-92.
 54. Groth, J., et al., Molecular characterisation of resistance against potato wart races 1, 2, 6 and 18 in a tetraploid population of potato (*Solanum tuberosum* subsp. *tuberosum*). J Appl Genet, 2013. 54(2): p. 169-78.
 55. Corwin, J.A. and D.J. Kliebenstein, Quantitative Resistance: More Than Just Perception of a Pathogen. The Plant Cell, 2017. 29(4): p. 655-665.
 56. Gebhardt, C. and J.P. Valkonen, Organization of genes controlling disease resistance in the potato genome. Annu Rev Phytopathol, 2001. 39: p. 79-102.
 57. Flor, H.H., Current Status of the Gene-For-Gene Concept. Annual Review of Phytopathology, 1971. 9(1): p. 275-296.
 58. Dutreux, F., et al., De novo assembly and annotation of three *Leptosphaeria* genomes using Oxford Nanopore MinION sequencing. Scientific Data, 2018. 5: p. 180235.
 59. Hulse-Kemp, A.M., et al., Reference quality assembly of the 3.5-Gb genome of *Capsicum annuum* from a single linked-read library. Horticulture Research, 2018. 5(1): p. 4.
 60. Su, X., et al., Chapter One - Heterologous Gene Expression in Filamentous Fungi, in Advances in Applied Microbiology. 2012, Academic Press. p. 1-61.
 61. Hyle, J.W., R.J. Shaw, and D. Reines, Functional distinctions between IMP dehydrogenase genes in providing mycophenolate resistance and guanine prototrophy to yeast. J Biol Chem, 2003. 278(31): p. 28470-8.
 62. Anonymous, 2017/625 (EU) on official controls and other official activities performed to ensure the application of food and feed law, rules on animal health and welfare, plant health and plant protection products.



Summary



Summary

Plant pathogens can have great social and economic impact, and are a continuous threat to food security. This is clearly the case for *Synchytrium endobioticum*, the species causing potato wart disease. *S. endobioticum* is an obligate biotrophic fungus of the phylum Chytridiomycota (chytrids), which is a basal lineage in the fungal kingdom. The lack of chemical control agents, the impact of the pathogen (complete loss of tuber yields from infected field have been reported), and the production of resting spores that remain viable and infectious in infested soils for decades, led to a quarantine status for *S. endobioticum* in most countries world-wide. Potato wart disease has been reported from all continents where potato is cultivated, and strict phytosanitary control measures are enforced to prevent the introduction and spread of the pathogen. The use of resistant potato varieties has proven successful in achieving these goals.

In **chapter 1**, over a century of potato wart research is reviewed to place our current knowledge in historic perspective. Observations from light microscopy and electron microscopy studies of *S. endobioticum* performed in the twentieth century are combined with recent molecular studies. Based on our current knowledge on molecular plant-pathogen interactions, a model is presented to describe the interaction between *S. endobioticum* and its host in particular with respect to plant resistance.

Chapter 2 describes the independent sequencing, assembly and functionally annotation of two *S. endobioticum* genomes. A comparative genomics approach, in which knowledge acquired from other fungal taxa is exploited, is used to gain insights into genomic features underlying the obligate biotrophic and pathogenic lifestyle of the pathogen. Our study underlines the high diversity in chytrids compared to the well-studied Ascomycota and Basidiomycota, and reflects biological differences between the phyla. Moreover, it highlights the surprising commonalities between plant pathogenic fungi that are so evolutionary distinct.

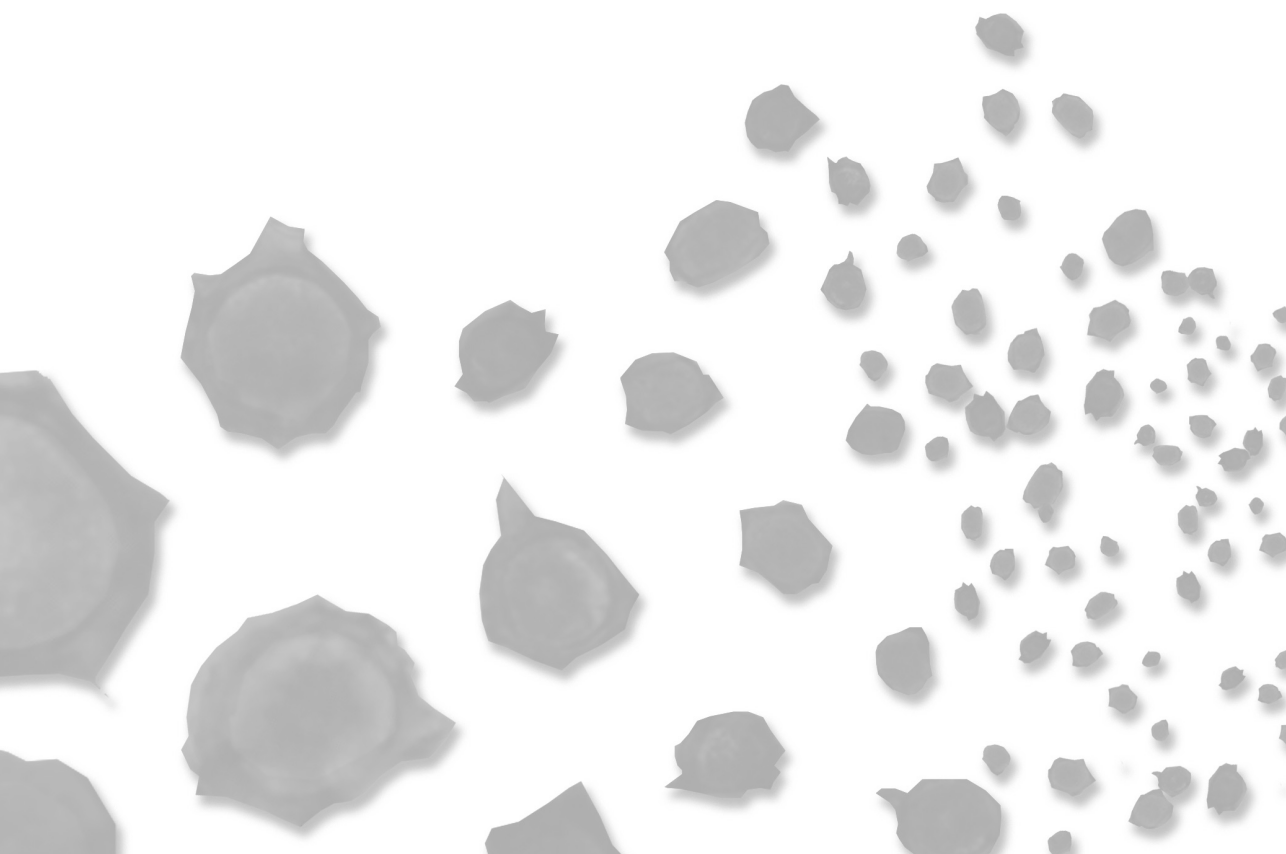
In **chapter 3**, the within-species diversity is determined using mitochondrial genomes, and four major mitochondrial lineages are identified. Furthermore, mitochondrial genomic variation shows that *S. endobioticum* has likely been introduced into Europe multiple times, and that several pathotypes emerged multiple times. We also demonstrate that isolates represent communities of different genotypes, and that the use of semi-resistant potato cultivars triggers a rapid shift in the mitochondrial haplotype. This shift is associated with

increased virulence and is likely the result of disruptive selection in the community. Our analysis reveals diversity of *S. endobioticum* isolates, which is undetected with the currently used bioassays.

In **chapter 4**, by means of an alternative “aboveground” bioassay, it is demonstrated that potato resistance genes are expressed in green aboveground plant organs similar to etiolated “belowground” sprouts. As potato wart resistance is active in both belowground and aboveground organs, the alternative bioassay can potentially speed up screening for *S. endobioticum* resistance in potato breeding programs as it omits the requirement for tuber formation. In addition, possibilities arise to express *S. endobioticum* effectors in potato leaves through agroinfiltration, thereby providing additional phenotyping tools for research and breeding.

Chapter 5 describes the identification of the *S. endobioticum* effector *AvrSen1*, which represents the first effector gene identified in Chytridiomycota. A single dominant gene (*Sen1*) governs pathotype 1(D1) resistance and we hypothesized that the underlying molecular model would involve a pathogen effector (*AvrSen1*) that is recognized by the host. Expression of *AvrSen1* in *Sen1* plants showed a hypersensitive response which co-segregated with *Sen1* in potato populations. In non-pathotypes 1(D1) isolates, five different variants resulting in the loss of recognition by the plant were observed suggesting that *AvrSen1* is under strong selective pressure.

Chapter 6 is a summarizing discussion in which the chapters are integrated, and several aspects linked to pathogenicity and pathotype diversity are further discussed. The impact of the research described in chapters 2 to 5 is specified. Finally, opportunities and challenges for future research on *S. endobioticum* are presented with a central role for an improved *Avr-R* gene based pathotyping system, which will contribute to the improvement of durable potato wart resistance breeding and phytosanitary control.



Acknowledgements



A PhD thesis can be expressed in many different ways. Be it the 290 pages that make up the book, the just under four years spent on the project, or the roughly 1,800 cups of coffee (totaling 360 liters, which equals 2 bathtubs filled to the brim) consumed in that time. Without a doubt the most important measure is the people that helped me in one way or another to reach this major milestone. Since my thesis is already rather thick (and I take it you read through the entire thing before ending up in this section), I have created a fun crossword puzzle to express my gratitude to the people involved instead of just adding another eight pages to the book.

Apart from the puzzle I would like to express my immense gratitude to promotor and co-promotors for their advice and support, and the trust bestowed upon me during the project. You have provided a creative environment in which I could explore the different aspects of potato wart disease and further develop my scientific and personal skills. I am sure that our collaboration will reach beyond the limits of this project, be it with potato wart or other plant pathogens and plant pests. I am looking forward to it!

Work is important, but family is even more important. Pascal, we both knew that starting this adventure would be challenging, even though we did not exactly know what was coming. When I was first asked if I wanted to be involved in the project, we just learned that Tijl (who was under a year old at that time) was to become a big brother to Loek and Noor. We had barely recovered from the first months after the birth of the twins, when the project at WUR started. Mieke and Henk played a big role by helping us to ease the (on occasion) heavy task of parenthood, and I am very thankful for that. Pascal, you were a steady rock during the project, which made it possible for me to complete this thesis. You are an excellent mother, a perfect girlfriend and an all-round superwoman. I love you!

Tijl, Loek and Noor, you are now 6, 4.5 and 4.5 years old. In that time you have learned how to sit, walk, talk, use cutlery at diner (more or less), open and close zippers, read and write your own names, and so on, and so on. These accomplishments require immense perseverance and the eagerness to constantly improve yourselves. Your curiosity and energy is very contagious and serves as a reminder to act in the same spirit. Now the project is finished, I am looking forward to spend more time with you catching frogs and playing football in the weekends instead of focussing on work. Tijl, Loek and Noor, it will take a while before you can read this section, but you are the greatest kids a dad could wish for.

And for the rest, thank you all!

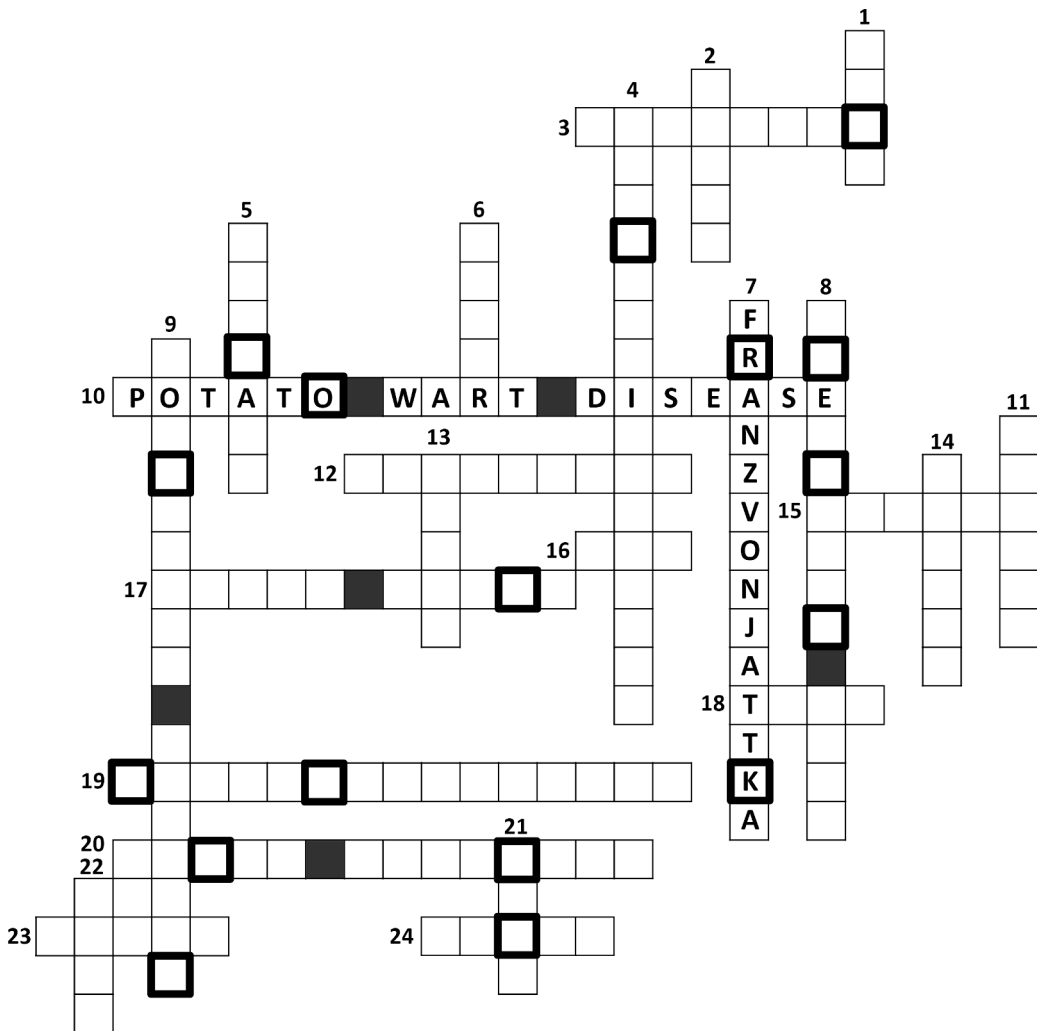
Horizontal

3. My former head of the National Reference Centre at NVWA who enthusiastically supported my decision to work on the PhD thesis.
10. My favorite plant disease.
12. The PhD student who worked on the potato-side of the plant-pest interaction. Good luck with finalizing your thesis!
15. My favorite dance partner at the Asilomar Fungal Genetics Conference. Hope to do it again real soon.
16. Without a doubt, you are my Canadian soulmate and I am looking forward to extend our collaboration beyond this project.
17. Carolien, Cees, Jan, Luc, Peter, Theo, Willem-Jan, Willem, apart from being very nice colleagues, you were also my regular [*complete the sentence*].
18. My co-promotor, mentor and excellent travel companion with a predilection towards kickstarter gadgets. We'll definitely keep bumping into each other.
19. The group at WUR that made me feel very much at home with Anne, Bram, Beatriz, Dirk-Jan, Els, Ilse, Marc, Odette, Stefan, Tanvi, Yvonne and many more!
20. The group where I did my promotion and where I learned many things that I'll be using for the rest of my life!
23. Without this key person in the lab and in the greenhouse this thesis would not exist.
24. I would very much like to visit the lobster festival once with this lovely lady from Prince Edward Island.

Vertical

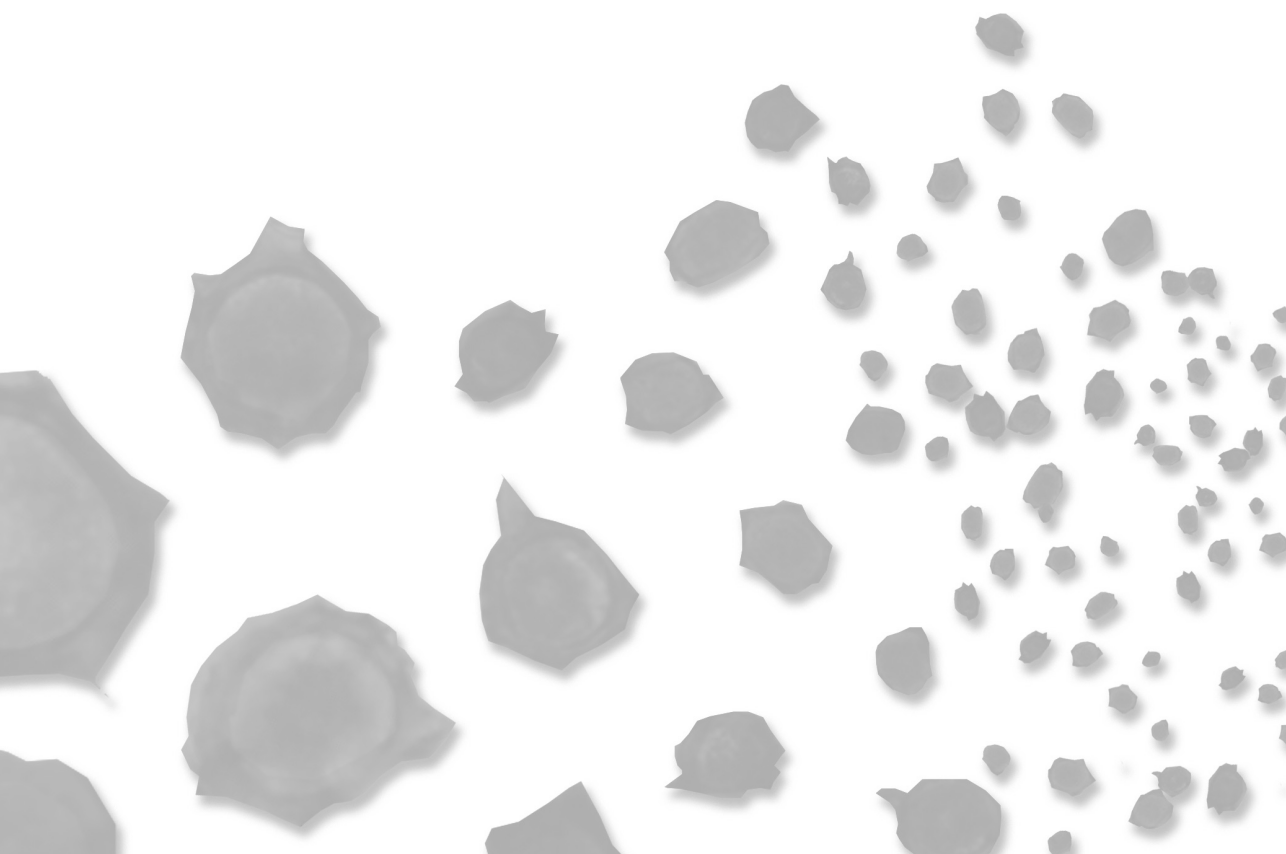
1. You had to wait a year for the baby birth gift of your oldest son. This time I'll be faster!
2. My former team leader at the NVWA who had been my mentor for many years and who enabled me working on this PhD project.
4. Gert and Marjan, thank you for maintaining the in vitro plant collection and performing all the [*complete the sentence*] experiments.
5. My promotor who provided a very productive and creative environment to study potato wart disease in all its aspects. Thank you for your advice and support!
6. Thank you for thinking about me as possible PhD student when the project proposal was accepted!
7. The first person to set the potato wart disease wheels in motion.
8. Those that deserve to be named, but who are not included in this puzzle.
9. I am proud to be part of this discipline at the NVWA with Esther, Eveline, Hans, Joris, Lucas, Marcel, Marleen, Marlies, Maureen, Naomi, Sonia and Tim.
11. The love of my life, excellent mother, and all-round superwoman.
13. A great source of inspiration, and the best canoe instructor I ever had.
14. The senior researcher who took over many of my tasks at the NVWA enabling me to spend time on the project.
21. The new PhD student who will carry the potato wart disease torch further into the future.
22. I am looking forward to chop some more wood with this co-promotor and great neighbour.

Answer the randomized questions, and collect the letters in the bold boxes. With these letters you can form a sentence and solve the puzzle. It is very likely that you will need the help of others to complete the puzzle, which shows that games are a great way to break down barriers between groups and to encourage collaboration.



Solution

 K O R



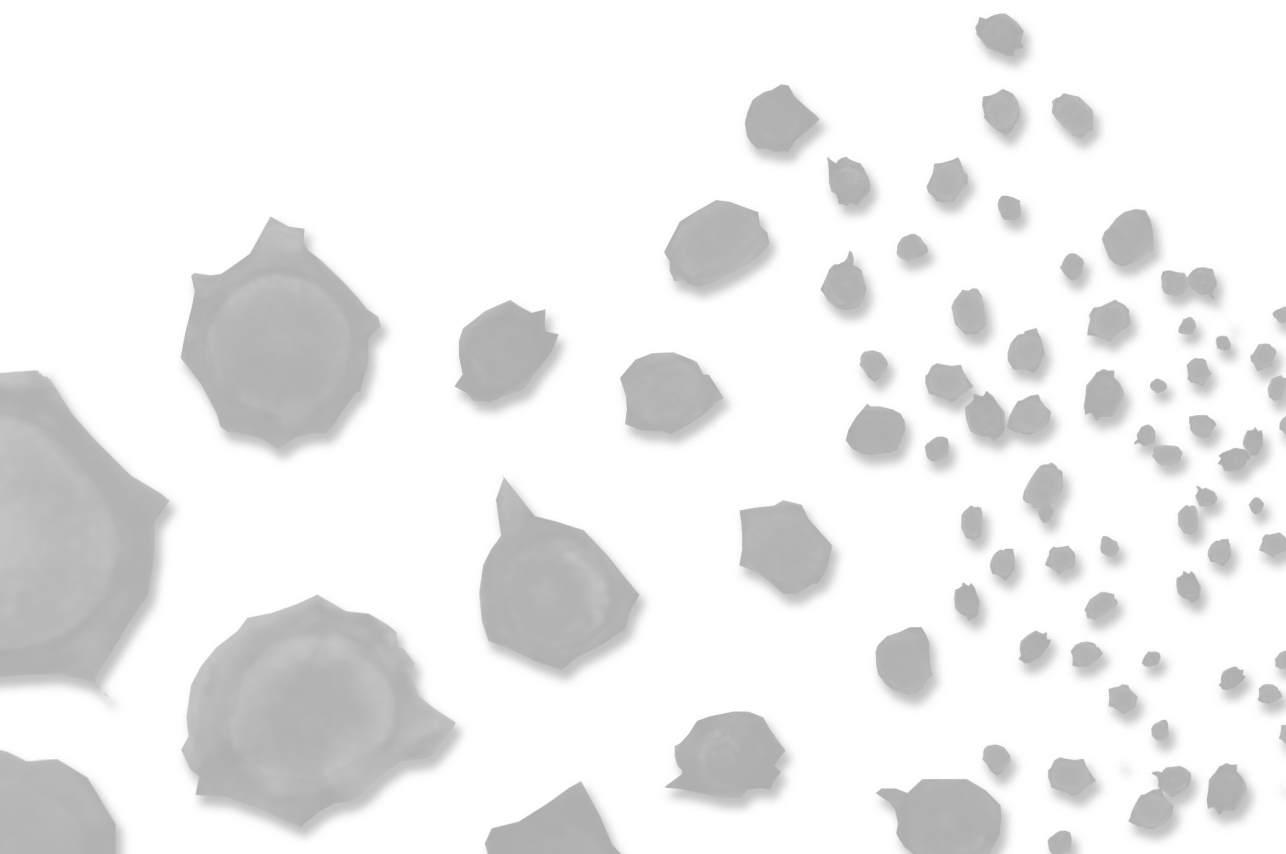
About the author





Bart van de Vossenberg was born on May 15th 1980 in Beek en Donk, the Netherlands. After obtaining his MAVO diploma in 1996, he graduated from the school for laboratory, environmental and engineering technologies (MLO) in 2000, and received a bachelor of science in Biology and Medical research at the Fontys University of Applied Sciences (HLO) in 2003. In the following year, he obtained a propaedeutic diploma in Arts at the Avans University of Applied Sciences. In 2004, Bart started working as a molecular

biological technician at the Plantenziektenkundige Dienst in Wageningen, which is nowadays part of National Reference Centre for plant health (NRC) of the Netherlands Food and Consumer Product Safety Authority (NVWA). In the molecular biological group of the NRC, molecular diagnostics and applied research projects are performed on a broad range of plant pathogens, plant pests, invasive plant species and insect vectors for human and veterinarian pathogens. In 2007, he was promoted to molecular biological senior specialist, and in 2012 Bart became lead molecular biologist. In this capacity he was responsible for the diagnostic and scientific output of the molecular biological group, and he coordinated or was involved in, numerous (inter)national research projects and training missions. As of 2010, Bart is a national representative in the “Diagnostics and quality assurance” panel of the European and Mediterranean plant protection organization (EPPO), and he is a member of the “Horizontal methods for molecular biomarker analysis” panel of the Dutch normalization institute (NEN). To further develop his scientific skills, Bart started working part-time as PhD student on the potato wart disease project in 2015, which led to this thesis. After finalizing his PdD thesis, Bart is now again fulltime employed as lead molecular biologist at the NRC. Besides his family and work, Bart is passionate about riding and customizing motorcycles, creating board games, and playing music.



Publications



van de Vossen B.T.L.H., Prodhomme C., van Arkel G., van Gent-Pelzer M.P.E., Bergervoet-van Deelen J.E.M., Brankovics B., Przetakiewicz J., Visser R.G.F, van der Lee T.A.J., and Vossen J.H. (2019) The *Synchytrium endobioticum* AvrSen1 triggers a Hypersensitive Response in Sen1 potatoes while natural variants evade detection, BioRxiv, doi.org/10.1101/646984

van de Vossen B.T.L.H., van Gent-Pelzer M.P.E., Boerma M., van der Gouw L.P., van der Lee T.A.J., and Vossen J.H. (2019) An alternative bioassay for *Synchytrium endobioticum* demonstrates the expression of potato wart resistance in aboveground plant parts, Phytopathology, doi: 10.1094/PHYTO-01-19-0024-R

van Valkenburg J.L.C.H., Schoenenberger N., **van de Vossen B.T.L.H.**, Man in 't Veld W.A., Westenberg M., Boer E. (2019) A natural hybrid of *Impatiens*, in the introduced range, demonstrated by sequence analysis of the nuclear ribosomal DNA-gene repeat, Botany Letters, doi: 10.1080/23818107.2019.1584863

van de Vossen B.T.L.H., Brankovics B., Nguyen H.D.T., van Gent-Pelzer M.P.E., Smith D., Dadej K., Przetakiewicz J., Kreuze J.F., Boerma M., van Leeuwen G.C.M., Lévesque C.A., and van der Lee T.A.J. (2018) The linear mitochondrial genome of the quarantine chytrid *Synchytrium endobioticum*; insights into the evolution and recent history of an obligate biotrophic plant pathogen, BMC Evol Biol. 18: 136.

van de Vossen B.T.L.H., Westenberg M., Adams I. *et al.* (2018). Euphresco Sendo: An international laboratory comparison study of molecular tests for *Synchytrium endobioticum* detection and identification. European Journal of Plant Pathology, 151 (3), 757-766

Vreeburg R.A.M., Bergsma-Vlami M., Bollema R.M., de Haan E.G., Kooman-Gersmann M., Smits-Mastebroek L., Tameling W.I.L., Tjou-Tam-Sin N.N.A., **van de Vossen B.T.L.H.**, and Janse J.D. (2016) Performance of real-time PCR and immunofluorescence for the detection of *Clavibacter michiganensis* subsp. *sepedonicus* and *Ralstonia solanacearum* in potato tubers in routine testing, EPPO bulletin, 46(1), 112-121

Hassani-Mehraban A., Westenberg M., Verhoeven J.T.J., **van de Vossen B.T.L.H.**, Kormelink R., Roenhorst A.J.W. (2016). Generic RT-PCR tests for detection and identification of tospoviruses. Journal of Virological Methods 233: 89-96.

Gilligan T.M., Tembrock L.R., Farris R.E., Barr N.B., van der Straten M.J., **van de Vossen B.T.L.H.**, and Metz-Verschure E. (2015). A Multiplex Real-Time PCR Assay to Diagnose and Separate *Helicoverpa armigera* and *H. zea* (Lepidoptera: Noctuidae) in the New World. PLoS One 10(11): e0142912. <https://doi.org/10.1371/journal.pone.0142912>

Bergsma-Vlami M., van de Bilt J.L.J., Tjou-Tam-Sin N.N.A., **van de Vossen B.T.L.H.**, Westenberg M., (2015). *Xylella fastidiosa* in *Coffea arabica* ornamental plants imported from Costa Rica in the Netherlands. Journal of Plant Pathology 97: 395

Braun-Kiewnick A., Viaene N., Folcher L., Ollivier F., Anthoine G., Niere B., Sapp M., **van de Vossen B.**, Toktay H. and Kiewnick S. (2015) Assessment of a new qPCR tool for the detection and identification of the root-knot nematode *Meloidogyne enterolobii* by an international test performance study, DOI 10.1007/s10658-015-0754-0

van de Vossen B.T.L.H., Ibáñez-Justicia A., Metz-Verschure E., van Veen E.J., Bruil-Dieters M.L., Scholte E.J. (2015) Real-time PCR tests in Dutch exotic mosquito surveys; implementation of *Aedes aegypti* and *Aedes albopictus* identification tests, and the development of tests for the identification of *Aedes atropalpus* and *Aedes japonicus japonicus* (Diptera: Culicidae), 3. Med. Entomol. 52(3): 336-350.

Soes D.M., Vierbergen G., **van de Vossen B.T.L.H.**, Breugelmanns K. and Backeljau T. (2015) A Banana slug *Ariolimax columbianus* (Gould, 1851) imported with Salal, Spirula, 402, 42-44 (in Dutch)

Man In 't Veld W.A., Rosendahl K., van Rijswijk P., Meffert J., Westenberg M., **van de Vossen B.T.L.H.**, Denton G., van Kuik F. (2014) *Phytophthora terminalis* sp. nov. and *Phytophthora occultans* sp. nov., two invasive pathogens of ornamental plants in Europe. Mycologia, doi: 10.3852/12-371.

van de Vossen B.T.L.H. and van der Straten M.J. (2014) Development and Validation of Real-Time PCR Tests for the Identification of Four Spodoptera Species: *Spodoptera eridania*, *Spodoptera frugiperda*, *Spodoptera littoralis*, and *Spodoptera litura* (Lepidoptera: Noctuidae), Journal of Economic Entomology 107(4): 1643-1654.

Ibáñez-Justicia A., Kampen H., Braks M., Schaffner F., Steeghs M., Werner D., Zielke D., den Hartog W., Brooks M., Dik M., **van de Vossen B.**, Scholte E-J. (2014) First report of established population of *Aedes japonicus japonicus* (Theobald, 1901) (Diptera, Culicidae) in the Netherlands, Jour. of Eur. Mos. Contr. Ass., published online 24th April 2014

van de Vossen B. T. L. H., Voogd J.G.B., Westenberg M. and Karssen G. (2014) Comparison of three conventional PCR test (Bulman & Marshall, 1997) versions for the molecular identification of *Globodera pallida* and *G. rostochiensis* cysts and juveniles, Bulletin OEPP/EPPO Bulletin 44 (1), 27-33

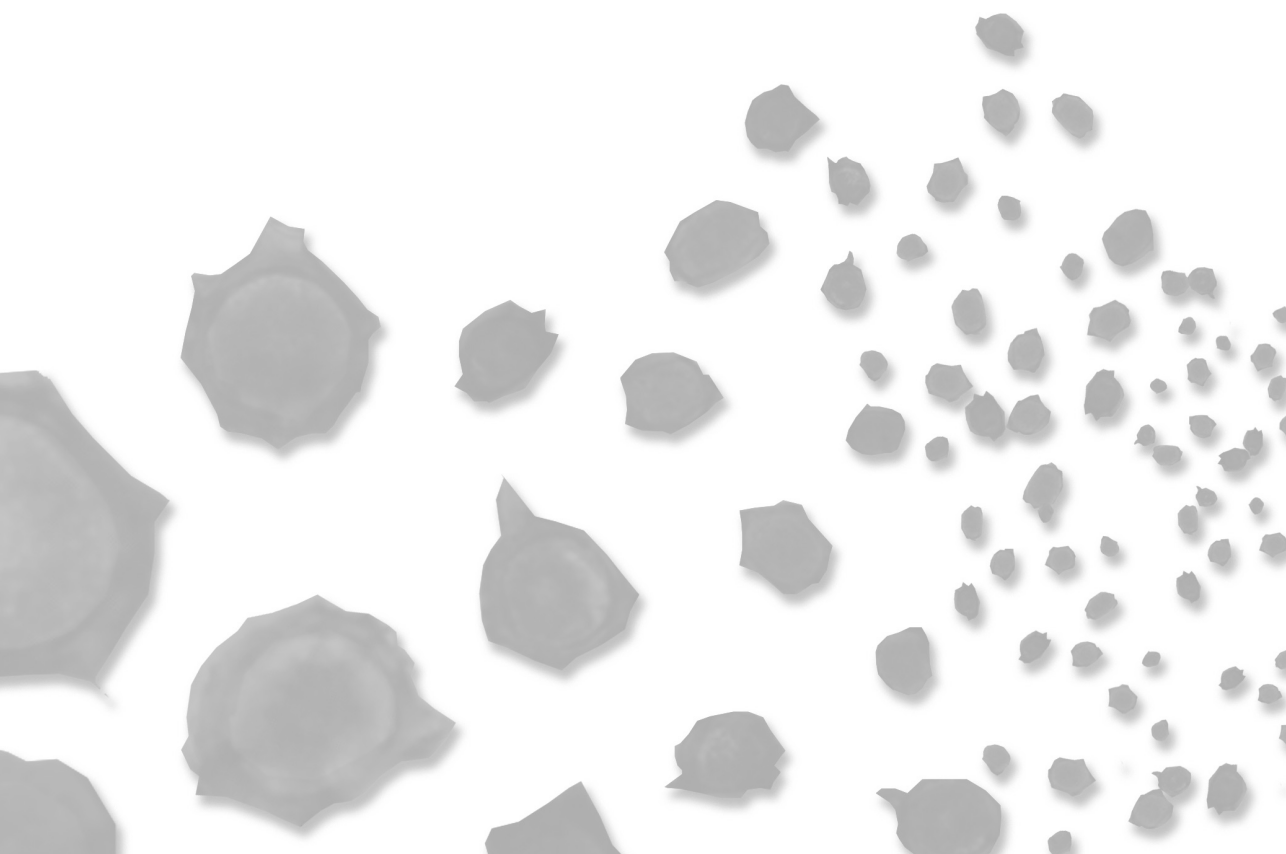
Ahmed M., **van de Vossen B.T.L.H.**, Cornelisse C., Karssen G. (2013) On the species status of the root-knot nematode *Meloidogyne ulmi* Palmisano & Ambrogioni, 2000 (Nematoda, Meloidogynidae), ZooKeys 362: 1-27 (2013)

Staats M., Erkens R. H. J., **van de Vossen B.T.L.H.**, Wieringa J.J., Kraaijeveld K., Stielow B., Gennl J., Richardson J.E., Bakker F.T. (2013) Genomic Treasure Troves: Complete Genome Sequencing of Herbarium and Insect Museum Specimens, PLOS ONE, 8(7): e69189. doi:10.1371/journal.pone.0069189

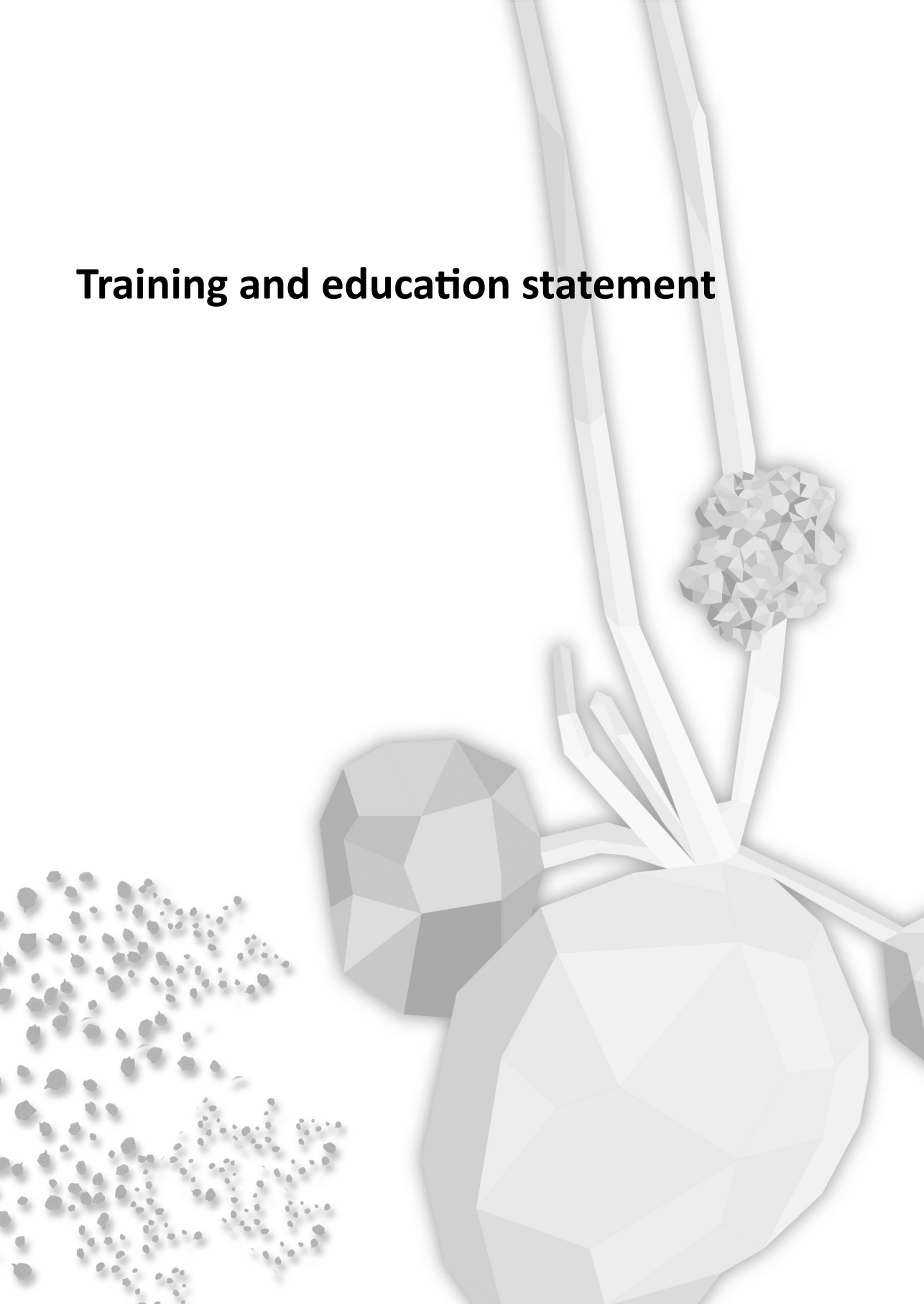
Botermans M., **van de Vossen B.T.L.H.**, Verhoeven J.T., Roenhorst J.W., Hooftman M., Dekter R., Meekes E.T. (2013) Development and validation of a real-time RT-PCR assay for generic detection of pospiviroids, J Virol Methods., 187(1):43-50

van de Vossen B.T.L.H., Westenberg M., Bonants P.J.M (2013) DNA barcoding as an identification tool for selected EU-regulated plant pests: an international collaborative test performance study among 14 laboratories, Bulletin OEPP/EPPO Bulletin (2013) 43 (2), 216-228

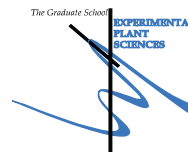
Kox L.F.F., van Brouwershaven I.R., **van de Vossen B.T.L.H.**, van den Beld H.E., Bonants P.J.M., de Gruyter J. (2007) Diagnostic Values and Utility of Immunological, Morphological, and Molecular Methods for *In Planta* Detection of *Phytophthora ramorum*, Phytopathology 97:9, 1119



Training and education statement



Education Statement of the Graduate School
Experimental Plant Sciences



Issued to: Bart T.L.H. van de Vossenbergh
Date: 1 July 2019
Group: Plant Breeding & Biointeractions and Plant Health
University: Wageningen University and Research

1) Start-Up Phase	<u>date</u>	<u>CP</u>
▶ First presentation of your project Unravelling potato wart disease; determination of the draft genome of the obligate biotrophic fungus	1 Jul 2015	1.5
▶ Writing or rewriting a project proposal		
▶ Writing a review or book chapter		
▶ MSc courses		

Subtotal Start-Up Phase

1.5

2) Scientific Exposure	<u>date</u>	<u>CP</u>
▶ EPS PhD student days		
EPS Get2Gether, Soest, the Netherlands	28-29 Jan 2016	0.6
EPS Get2Gether, Soest, the Netherlands	15-16 Feb 2018	0.6
▶ EPS theme symposia		
Theme 4 Symposium 'Genome Biology', Amsterdam, The Netherlands	15 Dec 2015	0.3
Theme 2 Symposium 'Interactions between Plants and Biotic Agents', Leiden, The Netherlands	22 Jan 2016	0.3
Theme 2 Symposium 'Interactions between Plants and Biotic Agents', Amsterdam, The Netherlands	24 Jan 2018	0.3
▶ Lunteren Days and other national platforms		
Annual meeting "Experimental Plant Sciences", Lunteren, the Netherlands	13 -14 Apr 2015	0.6
Annual meeting "Experimental Plant Sciences", Lunteren, the Netherlands	11-12 Apr 2016	0.6
Annual meeting "Experimental Plant Sciences", Lunteren, the Netherlands	9-10 Apr 2018	0.6
▶ Seminars (series), workshops and symposia		
Workshop: COST Evolutionary genomics of plant pathogens, Kiel, Germany	26-28 Aug 2015	0.9
6th Meeting of the EPPO Panel on Diagnostics and Quality Assurance, Paris, France	19-21 Jan 2016	0.6
Symposium: Plant Microbiome Network meeting, Wageningen, the Netherlands	29 Jun 2016	0.2
Seminar: Agri-food and agriculture canada (AAFC) seminar, Ottawa, Canada	19 Oct 2016	0.3
Workshop: Joint EPPO/EEC Workshop on Euphresco, Moscow, Russian Federation	27-28 Jul 2016	0.6
7th Meeting of the EPPO Panel on Diagnostics and Quality Assurance, Tbilisi, Georgia	17-19 Jan 2017	0.6
8th Meeting of the EPPO Panel on Diagnostics and Quality Assurance, Lisbon, Portugal	5-7 Mar 2018	0.6
▶ Seminar plus		
▶ International symposia and congresses		
Testa - EPPO Conference on diagnostics for plant pests, Angers, France	30 Nov - 4 Dec 2015	1.5
Congress: Fungal genetics conference, Assilomar, United States of America	14-19 Mar 2017	1.8
Congress: International Congress of Plant Pathology, Boston, United States of America	29 Jul - 3 Aug 2018	1.8
▶ Presentations		
Poster: COST workshop - Evolutionary genomics of plant pathogens, Kiel, Germany	27 Aug 2015	1.0
Talk: Testa-EPPO conference on diagnostics for plant pests, Angers, France	2 Dec 2015	1.0
Talk: Plant Microbiome Network meeting, Wageningen, the Netherlands	26 Jun 2016	1.0
Talk: Joint EPPO/EEC Workshop on Euphresco, Moscow, Russian Federation	27 Jul 2016	1.0
Talk: Agri-food and agriculture canada (AAFC) seminar, Ottawa, Canada	20 Oct 2016	1.0
Talk: Canadian Food Inspection Agency (CFIA) seminar, Ottawa, Canada	20 Oct 2016	1.0
Poster: Fungal genetics conference, Assilomar, United States of America	16 Mar 2017	1.0
Talk: EPS Get2Gether, Soest, the Netherlands	16 Feb 2018	1.0
Talk: Annual meeting "Experimental Plant Sciences", Lunteren, the Netherlands	10 Apr 2018	1.0
Poster: International Congress of Plant Pathology, Boston, United States of America	1 Aug 2018	1.0
▶ IAB interview		
▶ Excursions		

Subtotal Scientific Exposure

22.8

3) In-Depth Studies	<i>date</i>	<i>cp</i>
▶ Advanced scientific courses & workshops		
Toegepaste Statistiek, Wageningen, the Netherlands	Jan-Apr 2015	2.4
The power of RNA-seq, Wageningen, the Netherlands	10-12 Feb 2016	0.8
Comparative Genomics: from evolution to function, Nijmegen, the Netherlands	21-25 Nov 2016	1.5
▶ Journal club		
▶ Individual research training		
Guest researcher Agri-food and agriculture canada (AAFC), Ottawa, Canada	18-21 Oct 2016	1.2

Subtotal In-Depth Studies

5.9

4) Personal Development	<i>date</i>	<i>cp</i>
▶ General skill training courses		
PhD Competence Assessment, Wageningen, the Netherlands	2-3 Jun 2015	0.3
EPS introduction course, Wageningen, the Netherlands	22 Sep 2015	0.2
Information Literacy for PhD including EndNote Introduction, Wageningen, the Netherlands	14-15 Jun 2016	0.6
▶ Organisation of meetings, PhD courses or outreach activities		
Development educational board game: Quarantine laboratory	19 Apr 2018	1.0
Development educational board game: Potato Wart Disease!	8 Dec 2018	1.0
Organisation of the International Potato Wart Disease Workshop, Wageningen, the Netherlands	26-28 Jun 2019	1.0
▶ Membership of EPS PhD Council		

Subtotal Personal Development

4.1

TOTAL NUMBER OF CREDIT POINTS*	34.3
Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS with a minimum total of 30 ECTS credits.	
* A credit represents a normative study load of 28 hours of study.	

The research described in this thesis was financially supported by the Dutch Topsector Horticulture & Starting Materials (grant 1406-056). Within the Topsector, private industry, knowledge institutes and the government are working together on innovations for sustainable production of safe and healthy food and the development of a healthy green environment.

Financial support from Wageningen University & Research and the Dutch National Plant Protection Organization for printing this thesis is gratefully acknowledged.

