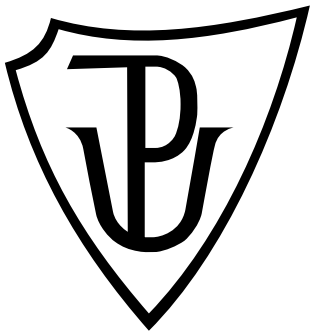
PALACKÝ UNIVERSITY IN OLOMOUC

Faculty of Science

Master′s Thesis

Olomouc 2016 **Eva Malečková**

PALACKÝ UNIVERSITY IN OLOMOUC
Faculty of Science
Department of Cell Biology and Genetics



Molecular Characterization of Fungal Pathogen *Blumeria graminis* and Development of Suitable Markers

Master's Thesis

Eva Malečková

Study program: Biology
Field of Study: Molecular and Cell Biology
Form: Full-time

Supervisor: Mgr. Miroslav Valárik, Ph.D.

Olomouc 2016

Declaration

I declare that this thesis has been composed by me and is entirely my own work, except where explicitly stated otherwise in the text, and that used sources are completely listed in the References chapter.

The work was done under the guidance of Miroslav Valárik, Ph.D. at the Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research in Olomouc, Czech Republic.

Olomouc, April 30, 2016

…….

Eva Malečková

Summary

The barley powdery mildew-causing fungus *Blumeria graminis* f. sp. *hordei* is considered a high-risk pathogen and has been studies intensively both for its agricultural importance and as a model for obligatory biotrophy. In the Czech Republic, the fungus has been receiving attention for decades but relying solely on phenotyping, e.g. assessments of virulence frequencies. Recently the whole genome sequence of *Blumeria graminis* f. sp. *hordei* was released, costs of next-generation sequencing have dropped down and a pioneering study suggested the power of molecular markers to study spatial and time dynamics of *Blumeria graminis* f. sp. *hordei* populations. Building on the recent progress, additional DNA markers were designed to further exploit their capability to differentiate isolates captured in a geographically limited area. In whole-genome sequences (Illumina MiSeq) of ten Czech isolates captured in 2014, there were identified 84,471 transposable element-based polymorphisms, 65,535 single nucleotide polymorphisms and 9,068 simple sequence repeats. Primer pairs were designed for 10 repeat-junction markers, 21 simple sequence repeats and for 20 amplicons carrying at least one single nucleotide polymorphisms. Repeat-junction markers yielded only 10 % of polymorphic markers but nearly 67 % of simple sequence repeat markers resulted into polymorphic and reproducible patterns. Out of 20 primer pairs for Sanger sequencing, only 4 were excluded after initial testing, i.e. 80 % have potential to be used in the future. In total, 232 isolates were screened for 158 polymorphisms: 141 simple sequence repeat alleles and 17 single nucleotide polymorphisms. Neighbor-joining algorithm was used to visualize power of the markers to discriminate individual isolates. In the cladogram, 97.4 % of all isolates could be unambiguously identified. Screened isolates originated from 17 and 15 defined locations in the Czech Republic collected in seasons 2014 and 2015, respectively but no correlation between genotype and geographical origin was revealed. Surprisingly, no isolates with identical set of polymorphisms were identified between two subsequent seasons. It was thus suggested that sexual reproduction, i.e. recombination, and/or migration of spores may play a more important role than expected, especially in a region in the heart of Europe.

Key words: *Blumeria graminis*, DNA markers, genetic diversity, population structure, powdery mildew, simple sequence repeats, single nucleotide polymorphism, transposable elements

Souhrn

Blumeria graminis f.sp. *hordei* je patogenní houba způsobující onemocnění ječmene nazývané padlí travní. Vzhledem k rozsahu způsobovaných ztrát je tento patogen intenzivně studován. Pozornost je mu věnována také proto, že se stal modelem pro studium obligátní biotrofie. V České republice je *B. graminis* intenzivně studována už desítky let, avšak výhradně s využitím fenotypování, tj. na základě hodnocení frekvence známých virulencí. Nedávno však byla zveřejněna referenční genomová sekvence tohoto organizmu. Pokles nákladů na celogenomové sekvenování umožnil předběžné studium využitelnosti molekulárních markerů za účelem studia diverzity populací *B. graminis* f.sp. *hordei* se slibnými výsledky. Na základě těchto poznatků byly proto navrženy další DNA markery s cílem umožnit rozlišení izolátů pocházejících z geograficky omezené oblasti. Z celogenomových sekvencí (Illumina MiSeq) deseti vybraných izolátů pocházejících z České republiky (rok 2014) bylo detekováno 84 471 potenciálních markerů založených na transponovatelných elementech, 65 535 jednonukleotidových polymorfizmů a 9 068 mikrosatelitů. Primery byly navrženy pro 10 markerů pro místa inzerce transponovatelných elementů, 21 mikrosatelitů a 20 amplikonů obsahujících jednonukleotidové polymorfizmy. Pouze 10 % markerů odvozených z inzerčních míst transponovatelných elementů bylo polymorfních. Na druhou stranu 67 % mikrosatelitů bylo polymorfních. Z 20 párů primerů navržených za účelem detekce jednonukleotidových polymorfizmů bylo až 80 % použitelných a může být v budoucnu zdrojem markerů. Celkem 232 izolátů bylo hodnoceno na přítomnost 158 polymorfizmů: 141 různých alel mikrosatelitů a 17 jednonukleotidových polymorfizmů. K vizualizaci rozlišovací schopnosti tohoto panelu markerů (tj. variability detekovaných genotypů) byl použit *neigbor-joining* algoritmus a v takto získaném kladogramu bylo 97,4 % studovaných izolátů charakterizováno jedinečnou sadou polymorfizmů. Studované izoláty pocházely ze 17 (rok 2014) či 15 (rok 2015) přesně vymezených lokalit v rámci České republiky, avšak korelace mezi genotypy izolátů a jejich geografickým původem nebyla patrná. Kromě toho nebyl identifikován jediný izolát, který by se mezi oběma sezonami vyznačoval shodnou sadou polymorfizmů. Je proto pravděpodobné, že sexuální reprodukce (tedy rekombinace) nebo disperze spor přispívá k diverzitě patogenu více, než se dosud předpokládalo, zejména pak v relativně malé oblasti ve středu evropského kontinentu.

Klíčová slova: Blumeria graminis, DNA markery, genetická diverzita, jednonukleotidový polymorfizmus, mikrosatelit, padlí, struktura populace, transponovatelný element

Acknowledgments

This work has been supported by grant LD14105 from Ministry of Education, Youth and Sports of the Czech Republic and is part of COST action FA1208 „Pathogen-informed strategies for sustainable broad-spectrum crop resistance".

First, I would like to thank my supervisor Mgr. Miroslav Valárik, Ph.D. Institute of Experimental Botany AS CR. The door to his office was always open whenever I ran into a trouble in the laboratory or when I had questions regarding bioinformatics or thesis writing.

This work could not be done without the help of Doc. Ing. Antonín Dreiseitl, CSc. (Agrotest Fyto Ltd., Kroměříž) who kindly provided collections of *Bgh* isolates from 2014 and 2015 seasons and seeds of susceptible barley, Mgr. Kateřina Holušová who constructed NGS sequencing libraries and Michael Abrouk, Ph.D. who performed computational analysis of obtained sequences and preselected candidate markers.

I would like to thank Mgr. Eva Komínková for introducing me to the topic and for willingly sharing her experience and gained knowledge with me, thus contributing to smooth progress of my work and avoidance of unnecessary mistakes. Most importantly, Eva was of a great help when she took care of all *Bgh* isolates in the 2015 season. My sincere thanks go to Mgr. Hana Vanžurová who also was very helpful whenever I needed to consult any issues related to laboratory work. Both ladies also helped me very much with initial inoculations.

Last but not least, I would like to acknowledge workers and students of the Centre of Plant Structural and Functional Genomics for creating supportive and friendly working environment, particularly Ing. Marie Seifertová for her assistance with Sanger sequencing procedures.

Finally, I must express my profound gratitude to my parents for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.

Table of contents

1 Introduction

Powdery mildew disease is a common cause of reduced yields as well as deterioration of end-product quality. The disease causing fungus *Blumeria graminis* infects four major *Poaceae* crops: wheat *Triticum aestivum* L., barley *Hordeum vulgare* L., oat *Avena sativa* L. and rye *Secale cereale* L. *B. graminis* f.sp. *hordei (Bgh)*, the *forma specialis* attacking barley, has been accepted as a model to study both host-parasite interactions and mechanisms underlying obligatory biotrophy. In agriculture, the significance of *Bgh* is illustrated by its inclusion on the list of Top Ten fungal pathogens (Dean *et al.*, 2012; Wyand and Brown, 2003). Barley powdery mildew disease causes significant losses worldwide and the Czech Republic, for which malting barley is an important economic commodity, is no exception (Dreiseitl, 2010).

Due to its agricultural importance, *Bgh* has been studied from numerous points of view: spore germination (Kunoh *et al.*, 1979; Wright *et al.*, 2000), leaf surface recognition by *Bgh* conidia (Nielsen *et al.*, 2000) and *Bgh* adaptation to barley resistance genes (e.g. Dreiseitl, 2003). Finally, reference genome sequence was obtained only a few year ago (Spanu *et al.*, 2010), triggering numerous follow-up studies, such as avirulence gene clustering (Skamnioti *et al.*, 2008), estimation of TE content (Amselem *et al.*, 2015), or coevolution of virulence effectors with retroelements (Amselem *et al.*, 2015; Sacristán and García-Arenal, 2008).

However, studies of the pathogen phylogeny have been primarily focused on ff. spp. classification (e.g. Troch *et al.*, 2014) or on a position of *Bgh* among other powdery mildews (Takamatsu, 2013a) because the resolution of used markers was not sufficient for characterization at a sub-species level. First DNA markers employed in *Bgh* population genetics were RFLP markers used to describe *Bgh* population in Denmark (Damgaard and Giese, 1996). Subsequently, a combination of RAPD and SCAR markers was employed in a study carried out in France (Caffier *et al.*, 1999). Despite increasing utilization of DNA markers, *Bgh* phenotyping even nowadays plays the most important role in *Bgh* isolate characterization as illustrated by recent studies done for example in the Czech Republic (Dreiseitl, 2015), Latvia (Kokina and Rashal, 2012) or Poland (Tratwal and Bocianowski, 2014). It is assumed that a combination of phenotyping and genotyping approaches might further deepen our understanding of *Bgh* diversity, spatial structure and dispersal patterns.

Because of its parasitic lifestyle, *Bgh* has closely co-evolved with its host, the major components of the continuous host-pathogen arms race being arsenal of *Bgh* effector proteins

encoded by avirulence genes and corresponding resistance genes in barley (Schulze-Lefert and Panstruga, 2011). Except for fungicide use, control of powdery mildew disease relies on breeding of resistant barley varieties but *Bgh* can acquire new avirulence genes rapidly, thus breaking the host resistance. In order to make the control of the disease more effective, better understanding of molecular mechanisms underlying host-pathogen interactions is essential (Oberhaensli *et al.*, 2011). Major outcomes of intensive *Bgh* research include cloning and mapping of barley resistance genes as well as fungal avirulence genes (Pedersen *et al.*, 2002; Seeholzer *et al.*, 2010). The *Bgh* genome has been sequenced and annotated (Spanu *et al.*, 2010) and components of *Bgh* proteome described (Bindschedler *et al.*, 2009; Kusch *et al.*, 2014), enabling subsequent molecular studies.

Bgh genetic diversity has remained largely unclear, partially due to labor intensive research of an obligatory biotroph. Increased employment of techniques of molecular biology promises faster progress in deciphering *Bgh* population structure and dynamics. DNA markers have been shown as an effective tool for population studies (Sunnucks, 2000) and the present work thus follows the pioneering study by Komínková *et al.* (2016, submitted), which revealed high resolution power of molecular markers to distinguish *Bgh* isolates unambiguously. Besides enriching the previously developed genotyping panel, this work aimed to assess *Bgh* diversity in the Czech Republic in two subsequent seasons.

When rapid and high-throughput methodology of marker development is established, molecular studies can be combined with traditional approaches based on differential tests to develop molecular markers tightly associated with *Bgh* genes essential for a successful infection, e.g. avirulence genes. Molecular markers also have the potential to further our understanding of seasonal changes in *Bgh* populations and decipher relative contributions of recombination, gene flow, genetic drift and migration to *Bgh* diversity. Taken together, knowledge of structure of local populations and mechanisms giving rise to new strains can help ensure longevity of resistant barley varieties produced in the future, thus contributing to more efficient control of *Bgh* epidemics. Last but not least, if efficient strategy to study *Bgh* population genetics is developed, it can be applied to other *formae speciales* of *B. graminis* or other fungi whose *in vitro* cultivation is demanding or impossible.

2 Background

2.1 *Blumeria graminis* f.sp. *hordei*

2.1.1 Powdery mildew pathogen in brief

Taxonomy of *B.graminis* f.sp. *hordei* and disease symptoms

The obligate biotroph *Blumeria (*syn. *Erysiphe) graminis* (DC.) Speer f.sp. *hordei* Marschal *(Bgh),* the causing agent of barley powdery mildew, was recognized as the sixth most important plant-pathogenic fungus because of its risk potential (Dean *et al.,* 2012). The *B. graminis* is a filamentous fungus of the phylum *Ascomycota* (order *Erysiphales*) (Schweizer, 2007; Thomas *et al.,* 2001). Barley *Hordeum vulgare* L. is an exclusive host of *Bgh* and since similarly specialized host-parasite interactions have been described, the *B. graminis* species has been classified into host-species-associated forms *(formae speciales,* ff.spp.*)*. These include three more special forms infesting cereal crops (f.sp. *avenae* on oat *Avena sativa* L., f.sp. *tritici* on wheat *Triticum aestivum* L. and f.sp. *secalis* on rye *Secale cereale* L.) and four *formae speciales* causing powdery mildew of wild grasses (f.sp. *poae* on genus *Poa,* f.sp. *bromi* on genus *Bromus,* f.sp. *agropyri* on genus *Agropyron* and f.sp. *dactylidis* on genus *Dactylis*). *B. graminis* ff. spp. from cultivated monocots indeed show high level of host specialization but the specificity to a single host, and thus the f. sp. classification, of *formae speciales* from wild grasses has been questioned (Inuma *et al.,* 2007; Troch *et al.*, 2014; Wyand and Brown, 2003). While *Bgh* is a useful model organism to study both host-parasite interactions and population genetics of pathogens causing powdery mildew, the disease is not restricted to monocotyledous plants only (McDonald and Linde, 2002; Wyand and Brown, 2003). Species attacked by other *Erysiphales* include legumes *(Fabaceae* Lindl.*)* (e.g. Sillero *et al.,* 2006), cultivated fruit trees from the *Rosaceae* family (e.g Pessina *et al.,* 2014) as well as genera of maple *(Acer* L.*)* and oak *(Quercus* L.*),* fruit-bearing plants like strawberry *(Fragaria* L.*)* (e.g. Mori *et al.*, 2015) and grape *(Vitis* L.*)* (e.g. Feechan *et al.*, 2008) and also ornamentals, e.g. lilacs *(Syringa* L.*)* (Seko *et al.*, 2008). Overall, powdery mildews affect around 10,000 angiosperm species with dicot powdery mildews showing significantly broader host range (Dean *et al.*, 2012; Glawe, 2008).

The most apparent symptom of powdery mildew disease is presence of white, powdery pustules on the surface of host tissues, typically leaves. If ascocarps (sexual fruiting

bodies) develop, they form dark-colored colonies within white asexual colonies (for *Bgh* reproduction, see page 12). Although *Bgh* usually does not kill their hosts, the infection nevertheless affects underlying tissues and numerous biochemical processes occurring in them. To support further *Bgh* development, nutrients rerouting from host to pathogen cells occurs. In response to infection, plants show increased rates of respiration and transpiration as well as reduced photosynthesis and chlorosis in the place of older *Bgh* colonies (Both *et al.*, 2005; Bushnell and Allen, 1962).

Geographical distribution of *B. graminis* f.sp. *hordei* and its importance for Czech agriculture

Bgh is a common and worldwide-spread pathogen but naturally, it is more abundantly present in areas with high density of its host, i.e. in colder latitudes of the Northern Hemisphere (Leff *et al.*, 2004). The first infection can occur early in the year and persist throughout the whole growing season. The infection causes major reductions in yield as well as in quality of harvested grains which result in considerable economical losses. The importance of powdery mildew for Czech agriculture was described by Dreiseitl (2011), who observed that between 1989 and 2000, 40 % of winter barley epidemic and 50 % of spring barley epidemic was caused by *Bgh*. The extent of losses caused by powdery mildew is largely depended on time of infection. If the disease occurs early enough, number of barley offshoots, and thus harvestable grains, can be markedly reduced. Even though yield reduction by 20 % or more is rather rare, losses as high as 40 % have been reported. The value of barley production in the Czech Republic depends on a number of factors, however, Dreiseitl (2010) estimated the annual value at 4 billion CZK and the losses caused by *Bgh* were estimated at 200 million CZK, i.e. 5 % of annual production.

Barley, the exclusive host of *Blumeria graminis* f.sp. hordei

Barley, however, is not of regional importance only. Being cultivated on 1,580,000 km^2 (9 % of world arable land), barley was the top fourth crop worldwide in early 2000s (Leff *et al.*, 2004). The global production for 2014/2015 season was estimated at 138.3 million tons and about 40 % (58.4 million tons) was harvested in the European Union. Barley produced in the Czech Republic represents nearly 2 million tons annually. There, barley is grown on about 25 % (1.4 million ha) of cultivated land and it is the second most commonly grown cereal after wheat. About 70 % of Czech barley is spring varieties but nevertheless, winter barley

plays an important role in powdery mildew epidemic as it serves as *Bgh* inoculum reservoir (Dreiseitl, 2005; Kůst and Potměšilová, 2014). Barley produced in the Czech Republic is mainly used for malting and as stock feed, besides that, it is country's important export commodity (Russell and Mikulasova, 2015).

2.1.2 The pathogen's life cycle

The reproductive strategies of *Fungi* are largely variable, exhibiting sexual (i.e. recombinant) or asexual (i.e. clonal) modes of reproduction as well as combinations thereof. Considering reproductive strategies as a classification criterion, Taylor et al. (1999) described three distinct meiosporic groups and one mitosporic group of *Fungi*. *Bgh* was classified as a *heterothallic type with a definite asexual phase*, reflecting the existence of both sexual and asexual phase in its life cycle. Indeed, both of the stages have been described in detail and can be distinguished with naked eye. Mature sexual fruiting bodies are visible as dark-colored dots, while asexually formed colonies are white to yellowish (Glawe, 2008).

For its entire nutritional dependence on a living host, the timing of *Bgh* reproduction needs to be synchronized with its host's life cycle, which explains why the annual sexual stage (teleomorph) does not occur until late in the barley growth season, i.e. in the fall (Glawe, 2008; Wolfe and McDermott, 1994). Thus, the ultimate aim of sexual reproduction is the formation of diploid chasmothecia (formerly cleistothecia), specialized structures capable of overwintering and producing haploid ascospores once the environmental conditions become favorable again. Then, the chasmothecia absorb water and break open, releasing the ascospores. On average, the contribution of ascospores to *Bgh* epidemic is relatively small, however, their importance becomes apparent from the adaptation and increasing variability point of view. In other words, the sexual cycle has the potential to give rise to many new allelic combinations during ascospores formation. Moreover, ascospores may play a bigger role in seasons with weather unfavorable for asexual development or when the transmission from spring- to autumn-sown barley populations is limited (Pariaud *et al.*, 2009; Ridout *et al.*, 2006; Ridout, 2009; Zhang *et al.*, 2005).

On the other hand, the asexual state (anamorph) is the predominant phase that is known for its very rapid progress and massive occurrence of conidiophores on leaves but also other aerial plant organs. Conidiophores are desiccation-sensitive and short-lived fruiting bodies but nevertheless capable of prolific production of oval-shaped asexual spores, conidiospores, that are organized in chains and dispersed by wind easily. The repetitive

asexual generations of annual life cycle thus make powdery mildew a polycyclic disease and enable rapid increase in the frequencies of advantageous allele combinations, ultimately contributing to epidemic spread of the disease (McDonald and Linde, 2002; Ridout, 2009; Zhang *et al.*, 2005).

As mentioned earlier, *Bgh* can survive either in forms of asexual colonies or as sexual ascospores inside chasmothecia in the absence of host. Another factor that contributes to spread of barley powdery mildew disease is growing of winter and spring barley cultivars. Such continuity of host populations creates so called *green bridge* that limits seasonal bottlenecks and makes it easier for the pathogen to transfer from one host plant to another. This makes *Bgh* a high-risk pathogen with a high potential for gene and genotype flow (McDonald and Linde, 2002; Wolfe and McDermott, 1994).

2.1.3 Host cell infection: cytological changes

Being a biotrophic ectoparasite, the pathogen's survival and reproduction is limited to a living host's epidermal cells and the infection process consists of a number of consequential morphological differentiation steps. The infection begins with a unicellular ascospore or conidium landing on a susceptible host. Physical contact with the host surface needs to be established and a germ tube forms. However, unlike other *Erysiphales*, *Bgh* germinates with primary and secondary germ tube. Shorter and rapidly emerging (0.5 to 2 hours after inoculation) primary germ tube secures attachment to the leaf surface, absorbs water from the infected cell and eventually penetrates through the epidermal cell wall. The secondary germ tube forms after 3 to 4 hours and in the presence of a functional primary germ tube becomes a foundation for the development of a haustorium. Once formed, the secondary germ tube swells, elongates and finally (8 to 10 hours later) differentiates into a hooked digitate appressorium, a specialized infection structure. By 15 to 18 hours after inoculation, a penetration peg emerges from beneath the mature appressorium and penetrates the leaf cuticle and epidermal cell wall, which is achieved through a combined effect of increased turgor pressure in the appressorium and enzymatic degradation of the host cell wall. Once in the inner environment of the host cell, the penetration peg starts differentiating into haustorium, a specialized intracellular hypha providing nutrients required for growth of hyphae forming on the leaf surface. Other roles of haustorium include host perception and control of host defense mechanisms and metabolism through effector delivery (Panstruga and Dodds, 2009). To secure formation of additional haustoria, the mycelium proliferates and gives rise to

secondary, elongating appressoria. Growth of circular colonies observable on the leaf surface is dependent on ambient temperature. Under optimal conditions, the newly formed colonies are visible to the naked eye in 3 days post-inoculation and the sporulation begins 4 to 5 days post-inoculation. Aerial reproductive structures develop from somatic hyphae and a single colony can release 300,000 spores (Catanzariti *et al.*, 2007; Giese *et al.*, 1997; Jørgensen, 2013; Kunoh *et al.*, 1979; Panstruga and Dodds, 2009; Pryce-Jones *et al.*, 1999; Wright *et al.*, 2000).

As a result of a successful host cell penetration, both invaded hyphae and host cell undergo structural changes. Haustoria are sealed in epidermal cell wall with a neckband, whose molecular composition has not been characterized yet. A separate compartment, extrahaustorial matrix, is formed between host cell cytosol and fungal extrahaustorial membrane, which is derived from plant plasma membrane. Although derived from plant plasma membrane, the extrahaustorial membrane shows some unique characteristics. It lacks several components of plasma membrane, while it also contains transmembrane protein RPW8. Interestingly, Arabidopsis RPW8 was identified as a resistance protein conferring non-race specific resistance to the host. Presence of a host protein on the host-pathogen interface proves established vesicular trafficking between the two organisms, which is an important part of plant-pathogen interaction (Catanzariti *et al.*, 2007; Giraldo and Valent, 2013; Opalski *et al.*, 2005; Wang *et al.*, 2009).

2.1.4 Host cell infection: gene expression changes and *Bgh* proteins

The development of *Bgh* during infection has been studied at the level of gene expression, too. Both *et al.* (2005) analyzed gene expression throughout asexual cycle of *Bgh* race IM82 and observed a global switch in gene expression between the pre-germination stage and post-penetration stage. The authors suggested that the observed changes reflected alternation of the fungus primary metabolism and occurrence of cellular processes specific to penetration and hyphae and haustorium development. Furthermore, number of up-regulated genes was homologous to described or putative proteins associated with pathogenicity and virulence of other fungi.

Over 800 proteins expressed in hyphae, haustoria and conidia of *Bgh* DH14 isolate (a source for reference genome sequence) were identified *in vivo* and their function was predicted (PANTHER classification). Haustorial proteome was rich in proteins of monosaccharide metabolism which is consistent with the feeding function of haustoria.

Proteins of nucleic acid metabolism were relatively more abundant in hyphae. This reflects active proliferation and nuclear division during conidia production. Finally, proteins identified in conidia included enzymes of lipid, carbohydrate and phosphate metabolism that are required for storage compound breakdown during spore germination. In contrast to tissue-specific proteins were heat-shock proteins (HSPs) whose presence was confirmed in haustoria, hyphae as well as conidia (Bindschedler *et al.*, 2009). More recently, the abundance of HSPs was confirmed in an *in silico* analysis of *Bgh* DH14 isolate genome. However, protein families with the largest number of predicted members are EKA (= effectors with similarity to avirulence proteins AVR_{k1} and AVR_{a10}) and CSEP (= candidate secreted effector proteins; proteins involved in host colonization), both protein families playing a role in plant-pathogen interaction (Kusch *et al.*, 2014).

Proteins of plant-pathogen interaction

Upon successful invasion of a host cell, its inner environment has to be changed so that it is favorable for further colonization by the pathogen and its reproduction. To modify cellular processes of the host, the pathogen secretes numerous effector molecules inhibiting plant defense response and enhancing pathogen virulence. Effectors can be either proteins or secondary metabolites, however, this text will focus on the effector proteins only. Based on reference genome sequence, existence of 1,300 effector protein paralogues has been proposed (Catanzariti *et al.*, 2007; Pedersen *et al.*, 2012; Schmidt *et al.*, 2014). Possible explanation for their abundance as well as their organization within the *Bgh* genome are discussed later. Here, proposed functions of selected effector proteins are described.

In the plethora of effector proteins, particular attention has been paid to avirulence factors. Encoded by *AVR* genes, they are usually small in size (less than 300 amino acids) and common among other eukaryotic plant-infecting pathogens as well as bacteria (Catanzariti *et al.*, 2007; Schmidt *et al.*, 2014). Although they too act as effectors enhancing infection, according to the *gene-for-gene* model, they are recognized by compatible resistance (R) proteins of the host upon entering the host cell (Flor, 1971). Thus, if a given pathogen strain lacks effector(s) capable of action in the host, the infection fails and the host is said to possess nonhost resistance (Schweizer, 2007). Both, avirulence proteins in *Bgh* and R proteins in barley have been identified.

There are at least 25 avirulence proteins in the EKA family, the most intensively studied proteins being AVR_{a10} and AVR_{k1}, both with paralogues in other powdery mildew

causing fungi. Contributing to successful infection, they are expected to be expressed in the fungal haustorium as it is the only tissue in direct contact with the intracellular environment of the host. However, unlike rust and oomycete avirulence proteins, AVR_{a10} and AVR_{k1} lack N-terminal secretory signal sequence, and thus an existence of a secretory pathway independent of endoplasmic reticulum was proposed (Amselem *et al.*, 2015; Catanzariti *et al.*, 2007; Ridout *et al.*, 2006). In barley, more than 85 R proteins have been discovered, including MLA_{10} and MLK_1, recognizing AVR_{a10} and AVR_{k1}, respectively (Schmidt *et al.*, 2014). Being receptor-like in nature, R proteins recognize compatible avirulence proteins secreted into the cytoplasm of the host cell and trigger a radical cellular response - necrosis of the infected cell. This hypersensitive response is typical of gene-for-gene interactions and restricts pathogen invasion to neighboring cells (Catanzariti *et al.*, 2007; Ridout *et al.*, 2006). Expression of R proteins by plant hosts thus provides an alternative explanation for nonhost resistance (Schweizer, 2007).

Another family of abundant proteins comprise of CSEPs. Initially, nearly 250 *Bgh* CSEPs were compared with proteins of *Erysiphe pisi* DC. and *Golovinomyces orontii* (Castagne) V.P. Heluta, infecting pea *(Pisum sativum* L.*)* and *Arabidopsis thaliana* (L.) Heynh., respectively. Only a fraction (10 CSEPs) were shared among all three powdery mildews, indicating that CSEPs might be highly species-specific (Spanu *et al.*, 2010). Since then, a total of 491 CSEPs was identified in the *Bgh* genome. Their function remains unknown, however, haustorial expression was confirmed for vast majority of them. Specifically, short (100-150 amino acids) CSEPs are believed to be involved in haustoria formation (Bindschedler *et al.*, 2011; Pedersen *et al.*, 2012; Spanu *et al.*, 2010). An example of well-studied *Bgh* (isolate DH14) candidate secreted effector protein is CSEP0055, which is thought to support fungal growth at the site of penetration (Zhang *et al.*, 2012). Involvement of CSEPs in the infection process is also supported by the fact that majority of haustorially-expressed CSEPs are down-regulated during an incompatible inoculation (Hacquard *et al.*, 2013).

Independently, five *Blumeria* effector candidates *(BEC1* to *BEC5)* showing differential expression throughout *Bgh* infection process were identified. However, only *BEC3* and *BEC4* were not identified previously as CSEP. Based on their expression profiles and amino acid composition, possible roles during the infection process were assigned to BEC2 to BEC4. BEC2 is thought to act in the apoplastic space, preparing the host cell for fungus entry. BEC3 and BEC4 were suggested to be involved in pathogen penetration and onset of haustoria formation. Taken together, fungal effector proteins in general might be

released very early in the infection process, possibly already during primary germ tube penetration. Hence, effectors do not probably play a role in haustorium formation only but also inhibit very early defense response, preventing cell death of invaded plant cells (Nowara *et al.*, 2010; Schmidt *et al.*, 2014).

2.1.5 *Blumeria graminis* f.sp. *hordei* genome

The *Bgh* genome size was estimated to be approximately 120 Mb (Spanu *et al.*, 2010). The closely related wheat powdery mildew (*B. graminis* f.sp. *tritici*, *Bgt*) is predicted to have a genome larger than 170 Mb (Parlange *et al.*, 2011). Other powdery mildew species also have genomes up to four times larger than non-parasitic *Ascomycetes*, suggesting that genome expansion is a common phenomenon among obligatory biotrophs (Spanu *et al.*, 2010).

Such extreme genome expansion can be explained by intensive proliferation of transposable elements (TEs). The most recent determination of TE content in *Bgh* DH14 isolate is 67 % (Amselem *et al.*, 2015) and the proportion of repetitive DNA in the *Bgt* genome might even exceed 90 % (Wicker *et al.*, 2013). The most abundant TEs in the *Bgh* genome are non-long terminal repeat (non-LTR) retrotransposons: LINEs and SINEs (Parlange *et al.*, 2011; Spanu *et al.*, 2010). LINEs and SINEs were estimated to comprise 24.5 % and 11 % of the *Bgh* genome. Nonetheless, accounting for over 20 % of the *Bgh* genome, LTR TEs also represent a significant proportion of the *Bgh* genome. In contrast to Class I (RNA transposons), the contribution of DNA transposons (Class II) to the size of the *Bgh* genome is less than 2 % (Amselem *et al.*, 2015).

With the *Bgh* DH14 genome assembly available, questions about the distribution of TEs and coding sequences throughout the *Bgh* genome arose. While evidence for clustering of coding sequences exists, there is some discrepancy regarding distribution of TEs. Two known clusters of avirulence genes were described, one of them containing both AVR_{a10} and AVR_{k1} (Ridout *et al.*, 2006; Skamnioti *et al.*, 2008). *CSEP* genes belonging to the same family tend to be clustered as well (Pedersen *et al.*, 2012). Based on an analysis of a representative portion of the *Bgh* genome (accounting for 9 % of the genome), clustering of TEs was excluded (Spanu *et al.*, 2010; Spanu, 2012). However, another group of authors identified two loci rich in TE sequences, suggesting clustering of TEs at least in some genome regions (Oberhaensli *et al.*, 2011).

Despite the contradiction about TE clustering, a contribution of TEs to reshuffling, amplification and diversification of genes interspaced between the repetitive sequences has

been accepted. Such genome flexibility is of particular importance for genes encoding proteins involved in host-pathogen interactions as changes of effector proteins provide a mechanism to escape recognition by the host. Multiplication of both *EKA* and *CSEP* genes probably occurred through action of TEs. For example, AVR_{k1} effector family coevolved closely with TE1a class of LINE retrotransposons (Pedersen *et al.*, 2012; Sacristán *et al.*, 2009). In fact, the relationship between retrotransposons and effector genes might be even tighter than that. An analysis of *Bgh* CC148 isolate genome revealed that AVR_{k1} and AVR_{a10} have evolved from the truncated ORF1 of class *I* LINE retrotransposons with two ORFs *(*elements *Satine* and *Kryze).* A striking similarity - 99 % nucleotide identity - was found between ORF1 of *Satine* and AVR_{a10}. The authors thus suggested that recycling and neofunctionalization of degenerate TE products may be a mechanism to generate the variability of avirulence and effector genes, thus increasing the change that fungal proteins will not be recognized by the host R proteins. Although the reference DH14 isolate lacks both AVR_{k1} and AVR_{a10}, it is intriguing to think that at least some of many *AVR* paralogues evolved through the described mechanism as well (Amselem *et al.*, 2015). It is also worth noting that sequence divergence of AVR_{k1} genes has been observed among the *formae speciales* of *B. graminis*, suggesting a possible mechanism underlying the strict host specialization (Raffaele and Kamoun, 2012; Sacristán *et al.*, 2009; Sacristán and García-Arenal, 2008).

Given the large size of the *Bgh* genome, a relatively small number of protein-coding sequences (6,470 annotated genes according to (BluGen - the *Blumeria* Sequencing Project; 27 April 2016) might be surprising. It turns out, however, that the gene loss largely reflects parasitic lifestyle of *Bgh*. Among genes that are believed to be reduced or entirely absent in the *Bgh* genome are genes of anaerobic fermentation, glycerol biosynthesis, inorganic nitrate and sulfur assimilation, and thiamine biosynthesis, which is in agreement with the fungus dependence on a host. Another feature of biotrophy seems to be a convergent loss of primary and secondary metabolism enzymes. Unlike in other fungal pathogens, plant cell wall depolymerizing enzymes are significantly reduced, too (Spanu *et al.*, 2010).

On the other hand, genes important for virulence and host invasion are amply represented in the *Bgh* genome. There are at least 1,350 *EKA* genes, paralogues of AVR_{k1} and AVR_{a10}, and nearly 500 *CSEP* genes were identified in *Bgh* haustoria and their expression has been understood to some extent. Both protein families seem to be required for successful host infection and their function in this process was discussed earlier (Kusch *et al.*, 2014; Pedersen *et al.*, 2012). Another unusually expanded protein family comprise fungus-specific, pathogenesis-related kinases, many of them possibly *forma speciale*.-specific as the *Bgt*

genome contains about a third of pathogenesis-related kinase genes identified in the *Bgh* genome.

Taken together, the *Bgh* genome illustrates evolutionary trade-offs of obligatory biotrophic lifestyle. First, irreversible gene loss disables release from host-dependent life and reproduction but on the other hand, TEs contribute to genome flexibility that can generate genetic variation, and ultimately adaptation to host resistance genes, without sexual reproduction. Second, functional and evolutional gains come at the costs of excess DNA maintenance (Parlange *et al.*, 2011; Raffaele and Kamoun, 2012; Spanu *et al.*, 2010). Another conclusion drawn on content of the *Bgh* genome says that *Bgh* probably reached its maximal adaptation and further evolutionary innovation seems unlikely (Kemen and Jones, 2012).

Comparison of *Blumeria graminis* f.sp. *hordei* genome to other powdery mildews

Sequencing of selected orthologous regions of *Bgh* and *Bgt* genomes revealed that protein-coding sequence are collinear, while the order of intergenic regions composed of TEs is far less preserved. Such observation suggests that most of the TE-mediated genome reorganization happened after the divergence of *Bgh* and *Bgt* (Oberhaensli *et al.*, 2011; Wicker *et al.*, 2013). This is in agreement with a large difference in TE content between the two *formae speciales* [67 % in *Bgh* genome (Amselem *et al.*, 2015) and 85 % in *Bgt* (Parlange *et al.*, 2011)] and a high degree of repetitive DNA similarity throughout the *Bgt* genome, suggesting only recent reshuffling of the *Bgt* genome (Parlange *et al.*, 2011).

Slightly over 90 % of predicted *Bgt* genes have homologous sequences in the *Bgh* genome, suggesting very similar gene content between the two *formae speciales*. Moreover, about a third of the homologs belong to *CSEP* gene family, while *CSEP* genes were not identified in powdery mildews of other genera. These findings imply that *CSEPs* are *Blumeria* specific genes, with a couple of *CSEP* families significantly expanded in *Bgh* in comparison to *Bgt* (Hacquard *et al.*, 2013; Kusch *et al.*, 2014; Wicker *et al.*, 2013). While a large proportion of both *Bgh* and *Bgt* genome is comprised of transposable elements with non-LTR being predominant, the TE composition seem to vary between the *formae speciales*. LINEs are the most abundant repeats in both genomes but the estimation of SINEs is significantly different: 3 % of the *Bgt* genome in comparison to 11 % in *Bgh* (Parlange *et al.*, 2011).

Comparison of genomes of *Bgh* and more distant powdery mildews *Erysiphe pisi* DC. (pea powdery mildew) and *Golovinomyces orontii* (Castagne) V. P. Heluta (*Arabidopsis thaliana* powdery mildew) revealed that genomes of these powdery mildews are

up to four times larger than genomes of most ascomycetes. Besides that, systematic analysis of core ascomycete genes was performed, revealing 99 homologous genes missing in all three powdery mildews. Many of these genes are not present in genomes of other obligatory biotrophs, suggesting that gene loss is common among obligatory biotrophs. Also, consistent with expanded diversity of *CSEPs* in *Blumeria* genomes is the finding that only a tiny fraction of *Blumeria CSEPs* were found in genomes of pea and Arabidopsis powdery mildews (Spanu, 2012).

2.1.6 Evolution of *Blumeria graminis* f.sp. *hordei*

Evolution of *Bgh* was studied with regard to other *formae speciales* as well as other pathogens causing powdery mildew disease. There is an agreement that powdery mildews form a monophyletic group (Takamatsu, 2013a). Originally, phylogenetic studies within the group were based on morphology of the fungi and host-pathogen relationships. However, the development of molecular methods resulted in reorganization of traditional classification with respect to the newest findings. rDNA internal transcribed spacer (rDNA-ITS) regions are common sequences to use in phylogenic enquires and they were also suggested to serve as a DNA barcode for *Fungi*, however their use in phylogeny of *Erysiphales* was refused by some authors (Álvarez and Wendel, 2003; Takamatsu, 2004; Wyand and Brown, 2003). Thus, coding sequences, such as 18S, 28S, and 5.8S rDNA, have been used instead. (Mori *et al.*, 2000).

Most recently, *EKA* and *CSEP* genes received attention in phylogenetic studies. Analysis of candidate *EKA* genes revealed they have arisen multiple times in different lineages, supporting their putative role in the adaptation of *Blumeria graminis* to various hosts (Amselem *et al.*, 2015). In another study, the ratio of nonsynonymous to synonymous substitutions was estimated for *CSEP* genes. This ratio was revealed to be unusually high, meaning that these sequences are under selection pressure to evolve rapidly (Wicker *et al.*, 2013). Furthermore, *CSEPs* were suggested to be *Blumeria*-specific proteins which supports the hypothesis of *Blumeria* species being only distantly related to other powdery mildews (Takamatsu, 2013b)

Based on molecular phylogenetic analyses and observed reproductive isolation of the four *formae speciales* affecting cultivated monocots, *Bgh* was proposed to be a sister clade to *Bgt* (Inuma *et al.*, 2007) and this finding was later supported by additional molecular analyses as well as by phylogenic trees based on phenotype characteristics (Takamatsu, 2013a). This is

also in agreement with conclusions based on TE-based studies (Wicker *et al.*, 2013). Work of Wyand and Brown (2003) identified close relation between *Bgt* and *B. graminis* f.sp. *secalis*, which is reflective of the phylogeny of their hosts, wheat and rye (Inuma *et al.*, 2007). However, the remaining two host-specific forms were not proven to reflect the phylogeny of their respective hosts. Overall, the evidence of phylogeny shared between *B. graminis* and host plant species remains ambiguous, probably due to host-jumping that occurred multiple times during *B. graminis* evolution (Inuma *et al.*, 2007; Wyand and Brown, 2003).

Due to their agronomical importance, there were several attempts to estimate the divergence time of *Bgh* and *Bgt.* Initially, the calculation was based on conserved intergenic sequences and estimated to occur approximately 10 million years ago (Oberhaensli *et al.*, 2011). However, Wicker *et al.* (2013) based their estimation on a number of substitutions in synonymous sites of more than 5,200 homologs and concluded that divergence of wheat and barley powdery mildews occurred later than originally believed, approximately 6.3 million years ago. The later estimation seems to be more accurate as the intergenic sequence-based molecular clock might be "faster" due to extensive TE activity. This new finding also suggests that *Bgh* and *Bgt* diverged after the divergence of their hosts 10 to 12 million years ago (Chalupska *et al.*, 2008).

2.1.7 Diversity studies of *Blumeria graminis*

Studies based on pathotype classification

Before molecular markers became a widespread tool for diversity estimations, structure of *Bgh* populations - as well as those of other pathogens - could be described in terms of pathotypes only. A classification into pathotypes is based on pathogenicity of studied isolates to a set of plant hosts carrying different resistance genes. Based on the severity of symptoms, *Bgh* pathotypes are classified on a nine-point scale (Caffier *et al.*, 1999; Dreiseitl and Pickering, 1999; Usami *et al.*, 2007). Testing *Bgh* virulence to known resistance genes enables comparison of populations from distant geographic regions, e.g. Europe *versus* China (Dreiseitl and Wang, 2007) or the Czech Republic *versus* Israel (Dreiseitl *et al.*, 2006), as well as estimations of population complexity and identification of prevailing virulence frequencies within a given population. Furthermore, this approach enables selection of *Bgh* pathotypes present in subsequent seasons, e.g. Dreiseitl *et al.* (2013) or Dreiseitl (2011).

The structure of *Bgh* populations has been intensively studied in the Czech Republic at least since mid-1970s but primarily by the means of *Bgh* phenotyping (Dreiseitl, 2005). Apart

from Czech *Bgh* populations, the pathogen pathotypes were studied at numerous other locations, including sites of only sporadic *Bgh* occurrence, e.g. New Zealand (Dreiseitl and Pickering, 1999) or North Dakota, USA (Dreiseitl and Steffenson, 2000).

Although studying pathotypes based on their virulence provides valuable information about population structure and dynamics, unfortunately it is not possible to link described virulence phenotypes directly to genotypes of respective *Bgh* isolates. For this reason, it is necessary to focus on diversity at the level of DNA as well. Thankfully, recent develop in the field of molecular biology makes it possible.

Molecular marker-based studies

The first molecular markers used to assess population diversity were based on enzyme variants, isoenzymes. Isozymes catalyze the same biochemical reaction but vary in their amino acid composition. Of particular importance are allozymes, allelic versions of the same gene. The changes in amino acid sequence translates into changes in protein net charge and ultimately in electrophoretic mobility. Thus, the analysis of isoenzymes/allozymes is based on separation by native gel electrophoresis and subsequent staining. Although isozyme/allozyme analysis is relatively inexpensive and fast, number of polymorphic alleles is usually limited. Furthermore, fresh samples and protein isolation within short time from sampling are required (Nybom *et al.*, 2014; Schlötterer, 2004; Weising *et al.*, 2005).

An example of isozyme-based study of *Bgh* is the work of Koch and Köhler (1990). Isozymes of ten enzymes together with analysis of unspecific proteins were used to study variation among *Bgh* isolates as well as among the four ff. spp. found on cultivated monocots. Although 131 distinct isozymes were identified, the observed polymorphism in banding patterns was very low (only 8 bands differed among the samples). Nevertheless, assessment of homology was possible, leading to conclusion that the most homologous ff. spp. were *B. graminis* f.sp. *secalis* and *B. graminis* f.sp. *tritici*.

Another drawback of isozymes/allozymes is that these markers do not allow identification of changes in DNA sequence directly. Thanks to rapid progress in molecular biology, this was soon compensated for with development of first DNA-based molecular markers. Each type of DNA markers has its advantages over other DNA markers as well as over isozymes/allozymes. For example, microsatellites are more likely to be neutral than allozymes and mutate at a higher rate (Delmotte *et al.*, 1999; Schlötterer, 2004).

DNA markers that first replaced allozyme analysis relied on use of restriction enzymes, e.g. restriction fragment length polymorphism (RFLP), but the true boom of DNA markers was triggered with the invention of PCR, enabling analysis of a specific genomic region in a large number of individuals simultaneously. Since then, DNA markers have been employed in a variety of applications, ranging from population studies to map construction and integration. Many types of DNA markers have been developed, utilizing polymorphism either in DNA sequence or DNA repeats and detecting variation in single loci or multiple genomic locations simultaneously (Nybom *et al.*, 2014; Schlötterer, 2004).

Since early 1990s, DNA markers have been employed to study *Bgh* diversity. Diversity of European *Bgh* populations have been studies using RFLP and RAPD markers, and first SCAR and VNTR (variable number of tandem repeats) markers were developed, too (Caffier *et al.*, 1999; Damgaard and Giese, 1996; McDermott *et al.*, 1994). Except for assessing *Bgh* diversity, AFLPs were used in combination with RFLPs and SNPs to construct first genetic map of *Bgh*. In the case of AFLPs and RFLPs, high degree of homology between target sequences and LTR retroelements was revealed (Giese *et al.*, 1999; Pedersen *et al.*, 2002).

In the last decade, simple sequence repeats (SSRs) were used to study cereal-infecting fungi such as *Parastagonospora nodorum* (Stukenbrock *et al.*, 2005) and *Zymoseptoria tritici* (Goodwin *et al.*, 2007). Only few years ago, SSRs were finally used for population studies of *Bgh*, too. Tucker *et al.* (2015) and Wang *et al.* (2014) used SSRs only to study diversity of *Bgh* and closely related *Bgt*, respectively. Komínková *et al.* (2016, submitted) employed SSRs in combination with ISBPs and SNPs to study diversity of *Bgh* isolates in the Czech Republic and compare them with selected isolates from a world *Bgh* collection. Although several TE-based marker systems have been developed and the *Bgh* genome is extremely rich in repetitive sequences, to the best of my knowledge, this study is also the only example of use of TE-based markers to study *Bgh* diversity.

In the last decade, single nucleotide polymorphism (SNP) has become an important and widespread type of polymorphism. With *Bgh* reference genome sequence and NGS sequencing platforms available, the popularity of SNPs is reflected in emerging studies of this type of polymorphism in the genome of *B. graminis*. In the work of Hacquard *et al.* (2013), genome sequence of the reference isolate and two additional isolates A6 and K1 from Sweden and Germany, respectively were compared. A mosaic structure of alternating monomorphic and polymorphic DNA blocks was revealed. Moreover, isolate-specific monomorphic blocks showed SNP densities as low as 0.05 SNP/kb and as high as 1.68 SNPs/kb. Such striking

contrast of SNP occurrence within the same genome was observed in the *Bgt* genome, too (Wicker *et al.*, 2013).

Taken together, so far published studies based both on pathotype and molecular characterization of *Bgh* have revealed that large genetic variation is maintained through frequent asexual reproduction and readily available for local selection exists in the pathogen populations (Hacquard *et al.*, 2013; Wolfe and McDermott, 1994). This thesis aims to describe diversity and dynamics of two *Bgh* population from the Czech Republic taking advantage of DNA molecular markers.

2.2 DNA molecular markers

Three types of DNA markers - repeat junction markers (RJMs), simple sequence repeats (SSRs) and single nucleotide polymorphisms (SNPs) - were employed in the experimental part of the thesis. Development of these marker types, polymorphism detection, applications and their advantages and limitations are described in detail.

2.2.1 Transposable elements and repeat junction markers

Transposable elements constitute a significant proportion of eukaryotic genomes and are thus a suitable genome component for development of molecular markers. TEs are dynamic elements of genomes that frequently insert into new genomic locations through copy-and-paste (Class I - retrotransposons) or cut-and-paste (Class II - DNA transposons) mechanism. Further classification of TEs distinguishes between long terminal repeat (LTR) elements and non-LTR transposable elements (LINE and SINE elements). TE classification system based on this division was proposed by Wicker *et al.* (2007) and will be used throughout the rest of the text. It is not rare for TE insertions to be nested one into another, which oftentimes give rise to unique TE junctions, i.e. boundaries between TE sequence and sequence of target locus. Many insertions are neutral to selection, resulting in their stable maintenance in genomes. In summary, activity of TE is considered to largely contribute to genomic diversity and evolution of the species (Bennetzen, 2000; Feschotte *et al.*, 2002; Kalendar *et al.*, 2011; Mazaheri *et al.*, 2014; Wanjugi *et al.*, 2009; You *et al.*, 2010).

TE-based markers rely on discovered or predicted junctions and five classes of TE markers based on PCR have been developed up to date: repeat junction markers (RJMs), repeat junction-junction markers (RJJMs), insertion site-based polymorphisms (ISBPs), inter-

retrotransposon amplified polymorphisms (IRAPs) and retrotransposon-based insertion polymorphisms (RBIPs). Of particular importance for population studies are RJMs because they were shown to generate highest number of primer pairs by available computational tools and, unlike other TE-based markers, they are also able to capture genes. RJM primer pairs consist of a primer specific for the junction and another primer complementary to any type of sequence (unique sequence or another TE) in neighborhood of the junction. The resulting amplicons are scored for presence/absence, using gel electrophoresis and PCR product visualization (Devos *et al.*, 2005; Kalendar *et al.*, 2011; Mazaheri *et al.*, 2014; You *et al.*, 2010).

RJMs are suitable for high-throughput genotyping and some RJM applications include studying functional genome diversity (Yadav *et al.*, 2015) and genetic diversity assays (You *et al.*, 2010). RJMs were also successfully used for phylogeny reconstruction, to construct physical and genetic linkage maps in wheat as well as in radiation mapping, e.g. of the D genome of bread wheat *Triticum aestivum* L. (Wanjugi *et al.*, 2009). Major drawbacks of RJMs are i) the need for junction sequence information in order to design specific primers and ii) the fact that transposon arrangements can be too variable to use them for comparisons of distant lineages (Kalendar *et al.*, 2011; Wanjugi *et al.*, 2009).

2.2.2 Simple sequence repeats

Simple sequence repeats a.k.a microsatellites consist of tandem repeated units one to six nucleotides in length and are highly abundant in eukaryotic genomes, both in genic (developed as EST-SSRs) and non-coding regions (Varshney *et al.*, 2005; Wang *et al.*, 2014). Although SSRs are prone to point mutations, the main driving force for polymorphism generation is slipped-strand mispairing during DNA replication. This mechanism explains existence of a number of alleles varying in the length of the repetitive motif at a single SSR locus. Since DNA sequences flanking the repeats are conserved, primer pairs can be designed to these regions. Thus, the variation is size of PCR amplicons reflects the variation in number of repetitive units (Akkaya *et al.*, 1992; Goodwin *et al.*, 2007). PCR products are separated by gel electrophoresis to size individual alleles. Alternatively, using fluorescently labeled primer, size of amplicons can be determined by capillary electrophoresis. Fluorescent labeling also enables multiplex analysis and scoring with available software (Tucker *et al.*, 2015).

There are several features of SSRs that make them a very powerful genetic marker for a variety of studies: relatively easy and cheap development (definitely for sequenced species),

their relative abundance in genomes and good genome coverage, relatively rapid mutation rate, high polymorphism, co-dominant inheritance, ease and reproducibility of assays (PCR). However, a significant drawback of SSR application in studies of fungi is the need for *de novo* development of SSR markers for every studied species as the cross-amplification even between closely related fungi, e.g. *Bgh* and *Bgt*, is very low. Possible solution might be development of genic SSRs, in which case primers are designed for more conserved sequences. Unfortunately, SSRs are also hard to automate (Fazekas *et al.*, 2010; Powell *et al.*, 1996; Robinson *et al.*, 2004; Schlötterer, 2004; Wang *et al.*, 2014).

Simple sequence repeats were used to study diversity of domesticated crops such as sunflower *Helianthus annuus* L. (Filippi *et al.*, 2015), rice *Oryza sativa* L. (Chakravarthi and Naravaneni, 2006) and *Sorghum* accessions (Agrama and Tuinstra, 2003). The variety of SSR applications include fingerprinting (including forensic purposes), population genetic and taxonomy studies, construction of genetic maps, e.g. of *Hordeum vulgare* L. (Varshney *et al.*, 2007), as well as genetic and physical map integration (Varshney *et al.*, 2005). For association studies, however, it should be kept in mind that alleles of the same size may have different evolutionary origin (Estoup *et al.*, 1995).

2.2.3 Single nucleotide polymorphisms

SNP is a polymorphism in the smallest building unit of DNA, i.e. a single nucleotide. Besides transitions and transversions, small indels are considered SNPs, too (Riju and Arunachalam, 2010). SNPs represent the most abundant type of allelic variation and are thus an increasingly popular type of molecular marker. Based on comparison of three *Bgh* isolates, a frequency of isolate-specific nucleotide variants was estimated at ~1 SNP/kb. SNPs can be found both in coding and non-coding sequences. Since a genic (i.e. in coding sequences) mutation can have deleterious effect on an individual's fitness, SNPs in coding regions are rather rare - less than 4 % in the *Bgt* genome (Wicker *et al.*, 2013). Nevertheless, both genic and genomic SNPs are used. Discovery of SNPs in coding regions take advantage of numerous EST databases. They can also be directly linked to gene function and are thus called 'perfect markers.' However, genomic SNPs have some advantages, too. They can be identified in any region of the studied genome. Furthermore, genomic SNPs are usually neutral to selection, making them more suitable for diversity estimations (Batley *et al.*, 2003; Hayward *et al.*, 2012).

The very first approaches to SNP discovery were relatively low-throughput but on the other hand, sequence knowledge was not required. Those days, SNP discovery was based on recognition of changes in restriction sites, DNA conformational changes or Targeting Induced Local Lesions in Genomes (TILLING) relying on use of a mismatch-specific enzyme. Nowadays, SNP discovery is possible through sequencing of PCR products, in a process called local SNP discovery (Shattuck-Eidens *et al.*, 1990). Recently, SNP discovery has largely benefited from next-generation sequencing (NGS), which enabled global *in silico* SNP discovery. However, for SNP discovery based on NGS data, a reference genomic sequence is required and is thus easier to achieve in model species (Kumar *et al.*, 2012; Li *et al.*, 2009; Morin *et al.*, 2004). An effort to simplify SNP discovery in non-model species resulted in introduction of restriction site-associated DNA sequencing (RAD-seq) (Etter *et al.*, 2011).

The most demanding approach to determine sequence variation is by gel electrophoresis performed under conditions that enable distinguish differences in electrophoretic mobility of amplified fragments caused by differences in their DNA sequence (Schlötterer, 2004). More advanced methodologies - such as allele-specific probe hybridization, primer extension assays and allele-specific PCR - rely on detection of fluorescence signal. In comparison to previously discussed molecular markers, SNP genotyping can be easily performed in a high-throughput mode, i.e. via Sanger sequencing or newly developed genotyping-by-sequencing and microarray approaches (reviewed by Chen and Sullivan, 2003; Kwok and Chen, 2003). Technologies currently available thus make SNPs suitable for autoimmunization and ultimately for high-throughput analysis. Furthermore, SNPs have a great potential for association studies between allelic forms of genes and phenotypes (Goldstein, 2001; Wang *et al.*, 2015). Biallelic nature of SNPs is a two-edged characteristics - while SNP scoring is less error-prone, they are also of smaller informational value. Thus, to obtain more information, multiple SNPs from a given region of genome, so called haplotypes, should be examined (Rafalski, 2002; Wang *et al.*, 2015).

SNP markers have been successfully used in construction of high density genetic maps necessary for quantitative trait loci mapping, genetic diversity studies, association mapping as well as in marker-assisted selection (Kumar *et al.*, 2012; Wang *et al.*, 2015). Sequence variants can be aligned to one another and based on homology of examined sequences, the extent of homology can be deduced from the alignment (Weising *et al.*, 2005).

3 Aims and objectives

Aims

This thesis builds on previous findings (Komínková *et al.*, 2016 submitted) and aims to enrich a panel of available DNA markers specific for *Blumeria graminis* f.sp. *hordei*, to use them in screening of *Blumeria graminis* f.sp. *hordei* isolates captured in 2014 and 2015 in the Czech Republic and to assess the pathogen diversity in the country.

Research objectives

1. Genomic DNA extraction from *Blumeria graminis* f.sp. *hordei* isolates
2. Next-generation sequencing and sequence analysis
3. Development and optimization of suitable molecular markers
4. Molecular characterization of selected *Blumeria graminis* f.sp. *hordei* isolates
5. Literature review, data evaluation and interpretation of results

4 Materials and Methods

Fungi and plants

Molecular analyses were performed on 232 *Blumeria graminis* f. sp. *hordei* isolates. The conidia of 119 and 113 *Bgh* isolates were captured in summer 2014 and 2015, respectively (Tabs. II and III), using phytopathological machine for collecting spores from the air mounted on the roof a car, as described previously (Dreiseitl, 2015; Schwarzbach, 1979). All capture sites were located along the routes in the main barley growing areas of the Czech Republic (Figs. 2 and 3). In 2015, 11 *Bgh* isolates from location HE and 21 *Bgh* isolates from location KM (Tab. III) were collected manually. The capture dates were chosen in accordance with expected intensive production of *Bgh* conidia in the month of June. The *Bgh* isolates were kindly provided by Doc. Ing. Antonín Dreiseitl, CSc. (Agrotest Fyto Ltd., Kroměříž).

Seeds of *Bgh*-susceptible barley *(Hordeum vulgare* L.*)* cv. Stirling used for conidia multiplication were kindly provided by Doc. Ing. Antonín Dreiseitl, CSc. Seedlings of cv. Stirling were grown in soil pots in a growth chamber for approximately 14 days. The growth conditions for plants and all *B. graminis* isolates were as follows: 19 °C/12 h and 17 °C/12 h, with fluorescent light intensity 35 $\mu mol \cdot m^{-2} \cdot s^{-1}$ and 50% humidity.

The isolate of *Blumeria graminis* f.sp. *tritici (Bgtm258)* captured in Olomouc, Czech Republic in 2010 was used as an outgroup. It was multiplied on leave segments of susceptible *Triticum monococcum* L. plants, under the same conditions as *Bgh* isolates. The growth conditions of *Bgtm258* and *T. monococcum* plants were the same as for *Bgh* isolates and barley plants.

Multiplication of inoculum

The *Bgh* conidia were cultivated on detached leaf segments as previously described by Dreiseitl and Wang (2007). In brief, about 25 mm long primary leaf segments were cut from 14 day old barley seedlings and placed in Petri dishes (140 mm in diameter) on 0.8% agar with 35 ppm benzimidazole, adaxial surface facing upwards. Petri dishes were placed at the bottom of a metal inoculation tower and spores harvested approximately 14 days after the previous inoculation were shaken on a clean ~ 5 × 5cm square of clean black paper to control the amount of inoculum visually. The paper square was then rolled and the inoculum was blown through a hole (15 mm in diameter) in the upper part of 30 cm tall metal inoculation

tower (Fig. 1). To ensure proper inoculation, the spores were allowed to settle down for 30 seconds before placing the closed Petri dish into a growth chamber. Conidia from fully grown mycelia were harvested into 2mL micro-centrifugation tubes after approximately 14 days and the samples were stored at -80 °C until used.

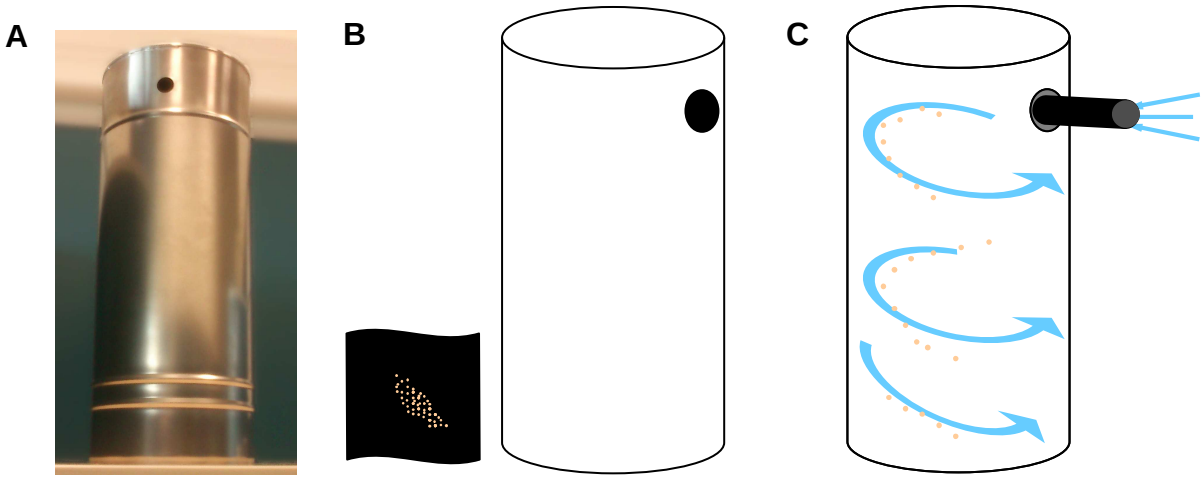


Figure 1: Inocula were multiplied using an inoculation tower. **(A)** A photograph of the inoculation tower used for inoculum multiplication. **(B)** First, spores were shaken on a clean 5 × 5cm square of clean black paper. **(C)** Second, the paper square was rolled and the inoculum was blown into the inoculation tower through a hole in the upper part.

DNA extraction

Genomic DNA was extracted according to the protocol developed previously by Komínková *et al.* (2016, submitted) with minor modifications. 10-15 mg of spores were treated with 300 µL solution consisting of 1 % (w/v) Lysing Enzymes from *Trichoderma harzianum* a.k.a. Glucanex® (Sigma-Aldrich, USA), 1 % (v/v) Triton™ X-100 (Sigma-Aldrich, USA) and 4 % (v/v) β-mercaptoethanol (Sigma-Aldrich, USA) and 285 µL 50mM EDTA pH 5.6 (Sigma-Aldrich, USA) per sample. The spores were gently but completely suspended and incubated at 37 °C for 2 hours. In the half of the incubation time, the mixture was vortexed briefly. After incubation, 250 µL lysis buffer [0.83 % (w/v) L-ascorbic acid (Sigma-Aldrich, USA), 0.05 % (w/v) proteinase K (Roche Diagnostics, Switzerland), 1.67 % (v/v) SDS, 6.67 % (v/v) β-mercaptoethanol, 500mM NaCl, 100 mM Tris-HCl pH 8.0 and 50mM EDTA pH 8.0 (Sigma-Aldrich, USA)] were added. The tubes were inverted few times during the 70 minutes long incubation at 65 °C. The lysate was cooled on ice and DNA was extracted with 250 µL phenol (pH 7.0), hand shaken and centrifuged for 5 minutes. The aqueous phase was pipetted into a fresh tube and phenol was removed using extraction to 250 µL of chloroform:isoamyl alcohol 24:1 mixture and

centrifuged for 3 min. The obtained aqueous phase was transferred to fresh tube and freed from residual phenol with additional chloroform:isoamyl alcohol extraction (500 µL) as described above. Extracted DNA was precipitated with 50 µL 3M NcAc (pH 5.2) and 440 µL pre-chilled isopropanol. After inverting gently and 5 minute centrifugation, the supernatant was discarded. Precipitated DNA was washed with 500 µL ethanol twice. First, 70% ethanol was used and after 5 minutes of centrifugation and discarding the supernatant, 96% ethanol was used the same way. The washed DNA was air dried for about 15 minutes at room temperature and dissolved in 20 µL sterile re-distilled water. Finally, the samples were treated with 10 µg/ml RNase A (Sigma-Aldrich, UK). All centrifugation steps were carried out in a pre-cooled Jouan GR 20-22 centrifuge (Thermo, USA) at 10,000 g and 4°C.

Concentration of DNA in each sample was measured with PicoGreen® dsDNA Quantitation Reagent (Molecular Probes, Inc., USA) and Turner BioSystem fluorometer (USA) according to the manufacturer's protocol. Quality of extracted genomic DNA was checked using electrophoresis in randomly chosen samples. 1 µL DNA was mixed with 4 µL re-distilled water and 1 µL 6× STOP C loading buffer (100 mM EDTA, 1 % SDS, 0.05 % bromphenol blue, 0.05 % xylene cyanol and 42.5 % glycerol). Samples were allowed to migrate in 0.8% agarose (Serva, Germany) and 0.5× TBE buffer (45 mM Tris-HCl pH 8.0, 45 mM boric acid and 1 mM EDTA) for approximately 60 minutes (5 V/cm). Owl B2 (Thermo Scientific, USA) electrophoresis chamber and MP-500 V (Major Science, USA) power supply were used. The gel was stained in 0.05% ethidium bromide (Sigma-Aldrich, USA) and visualized with Syngene UV transilluminator and GeneSnap software (Syngene, UK).

NGS Illumina library construction and sequencing

DNA sequencing libraries of 10 isolates were prepared from 50 ng of extracted DNA per isolate using Nextera XT DNA Library Preparation Kit (Illumina, USA) according to the manufacturer's instructions with the following modifications. The tagmentation reaction contained 6 µL ATM (amplicon tagment mix) and the subsequent incubation step lasted for 6 minutes. Libraries were amplified in C 1000™ Thermal Cycler (Bio-Rad, USA) starting with 72 °C for 3 minutes and 95°C for 30 seconds, followed by 14 cycles of 95 °C for 10 seconds, 55 °C for 30 seconds and 72°C for 30 seconds, and finished by final extension at 72°C for 5 minutes. Then, the reactions were finished with a continuous 10°C step. The library normalization step was omitted and instead, all libraries were quantified by qPCR

using the KAPA Library Quantification Kit for Illumina Genome Analyzer Platform (KAPA Biosystems, Inc., USA). To assist with insert size estimation, agarose gel electrophoresis was employed. Following equimolar library pooling, the final library concentration was checked again. The libraries were paired end sequences using MiSeq Reagent Kit v3 (600 cycles) (Illumina, Inc., USA) according to the manufacturer's instructions at the Institute of Molecular and Translational Medicine (Olomouc, Czech Republic).

Sequence analysis and primer design

Obtained whole-genome sequences were used to select candidate molecular markers of three types: repeat junction marker (RJM), simple sequence repeat (SSR) and single nucleotide polymorphism (SNP). Selection of candidate SNPs was performed as follows. The SNPs were called from reads of the ten isolates aligned against the *Blumeria graminis* f.sp. *hordei* DH14 version 3 reference sequence generated by the BluGen consortium (Spanu *et al.*, 2010). BWA software (Li and Durbin, 2010) was used to map the paired-end reads from each isolate to the *Bgh* reference genome sequence and variant detection was performed using GATK software (DePristo *et al.*, 2011). As a last step, home-made script was used to sort and select ~ 65,000 stringent SNPs per isolate. To ensure amplification of unique sequences, the candidate sequences were also masked for transposable elements using the *Blumeria graminis* f.sp. *tritici* transposable element database (TREP) developed by Wicker *et al.* (2013). Only, SNPs polymorphic in four to six isolates were considered informative.

To detect RJMs and SSRs, MaSuRCA assembler (Zimin *et al.*, 2013) was used to obtain an assembly of each *Bgh* isolate. However, for candidate RJM and SSR marker selection, only the assembly of *Bgh* isolate O4 was used for it was sequenced with the highest coverage (23 ×). The RJMs and SSRs with 2bp to 4bp repeat motifs were detected directly on contig assemblies using isbpFinder.pl (Paux *et al.*, 2010) and ssrFinder.pl scripts, respectively. To search for TE junctions the *Bgt* TREP was used as source of TE motifs.

Primer pairs for chosen candidate sequences were designed using Primer3web version 4.0.0 (Untergrass *et al.*, 2012) and its default settings except that the primer length was 18 to 23 nucleotides with an optimum of 20 nucleotides and T_m optimum was 60 °C. The optimal PCR product length was marker-dependent: ~ 500 bp for RJMs, 90 bp to < 200 bp for SSRs (depending on the length and number of repeat units), and ~ 300 bp to 450 bp for SNPs. SSRs with the highest number of repeat units were prioritized for primer design. RJM primers were designed so that at least one primer was spanning the TE junction.

PCR amplification and PCR amplicon separation

Besides newly designed primer pairs, five primer pairs originally developed by Komínková *et al.* (2016, submited) were used for PCR amplifications. Each PCR reaction (total volume of 15 μL) contained 1 × PCR buffer (10mM Tris-HCl, 50mM KCl, 1.5mM $MgCl_2$), 1 × Cresol Red (0,01 % o-cresolsulfonephthalein; 1.5 % sacharose), 200 μM of each dNTP, 1 μM of each primer, 0.6 U of *Taq* polymerase and 50 pg genomic DNA. To rule out the option of amplification failure, a primer pair (0.75 to 1 μM of each primer) specific for glyceraldehyde-3-phosphate dehydrogenase gene (GPD) developed previously (Komínková, 2011) was used as an internal standard. The combination of GPD primers (Table I) was chosen so that the internal standard product could be easily differentiated from RJM- or SSR-specific amplicons. The primer annealing temperature of these reactions was set according to the optimal temperature for RJM and SSR primer pairs.

Table I: Primer pair combinations amplifying fragments of glyceraldehyde-3-phosphate dehydrogenase gene, an internal standard used in selected PCR reactions

primer	sequence (5' - 3')	*in silico* product size [bp]
Bgh GPD F2	ACGCACCCATGTTTGTCAT	102
Bgh GPD R2	CCAATGGGGCAAGACAGTTA	
Bgh GPD F2	ACGCACCCATGTTTGTCAT	259
Bgh GPD R3	GCACCAGTGCTGCTAGGAAT	

PCR reactions were performed in C 1000™ or C 1000 Touch™ Thermal Cycler (Bio-Rad, USA). The PCR amplification steps included an initial denaturation of 95 °C for 5 min followed by 40 cycles of 30 s at 95 °C, 30 s at the primer-specific annealing temperature (see Tabs. IV, V and X), elongation at 72 °C for 30 s, and a final 10 min long extension step at 72 °C. Then, the reaction was finished with a continuous 4°C step.

The PCR amplicons were separated and visualized on 4% and 6% non-denaturing polyacrylamide gels [acrylamide-bisacrylamide 19:1 according to the required concentration, Bio-Rad Laboratories, Inc., USA; 45 mM Trizma® base (Sigma-Aldrich, USA), 45 mM boric acid (Lach-Ner, Czech Republic), 1 mM EDTA (Sigma-Aldrich, USA), 0.00073 % TEMED (Bio-Rad Laboratories, Inc., USA) and 0.067 % APS (Sigma-Aldrich, USA) were run in 0.5× TBE buffer [45 mM Trizma® base (Sigma-Aldrich, USA), 45 mM boric acid (Lach-Ner, Czech Republic) and 1 mM EDTA (Sigma-Aldrich, USA)] containing 0.005 mg ethidium bromide (Sigma-Aldrich, USA). 4% gels were used to separate RJM and SNP products and SSR alleles were separated on 6% gels. Before loading the samples, the gel was run at 300 V for 60 to 90 minutes to allow ethidium bromide to run up the gel. The PCR products were

separated at 350 V, time of separation being dependent on size of amplicons and in the case of SSR markers also on size differences among individual alleles. The power supply MP-500 V (Major Science, USA) and Standard Power Pack P25 (Biometra®, Germany) were used. The gels were visualized using INGENIUOS documentation system and GeneSnap software (Syngene, UK).

Sanger Sequencing

Prior to running sequencing reactions, 20 to 40 ng of PCR product (assessed from intensity of bands) were Exo-SAP purified. Besides PCR products, each purification reaction (total volume of 7 μL) contained 1 U of Exonuclease I (Thermo Scientific™, USA), 0.5 U of FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific™, USA) and 1 × PCR buffer (10mM Tris-HCl, 50mM KCl and 1.5mM $MgCl_2$). The mixture was incubated for 30 minutes at 37 °C and enzymes were inactivated with 95 °C for 5 minutes. The incubations were performed in C 1000™ or C 1000 Touch™ Thermal Cycler (Bio-Rad, USA).

Purified PCR products were cycle sequenced using BigDye® Terminator v3.1 Cycle Sequencing Kit **(**Applied Biosystems, USA). Each sequencing reaction (volume of 10 μL) contained 1.5 μL of sequencing buffer, 0.875 μL of DBX64, 0.125 μL of BigDye®, 10 μM primer and 10-20 ng of template DNA. Sequencing reaction were performed on C 1000™ or C 1000 Touch™ Thermal Cycler (Bio-Rad, USA). The reaction began with initial denaturation at 98 °C for 5 min, followed by 60 cycles of 10 s at 96 °C, 5 s at the 50 °C and an elongation at 60 °C for 4 min. Then, the reaction was finished with a continuous 4°C step. Prior to subsequent analysis, sequencing reactions were purified on Biomek® NXP (Beckman Coulter, USA) automation workstation using magnetic beads Agencourt® CleanSeq® (Beckman Coulter, USA) using manufacturer instructions. Finally, capillary electrophoresis was performed on ABI™ 3730xl DNA analyzer (Applied Biosystems, USA).

Data analysis and cladogram construction

Variability of *Bgh* genotypes was visualized in a cladogram. Scoring of identified polymorphisms was adjusted to the presence/absence nature of RJMs, i.e. all polymorphisms were converted to a binary code. In the case of SSRs, each polymorphic allele was scored separately. For SNP analysis, sequences of PCR amplicons were aligned using MEGA6 (Koichiro *et al.*, 2013) and manually edited if needed. Each position with nucleotide variation was scored separately for presence/absence of observed nucleotide variants. This strategy

made it possible to obtain a set of binary data in a format compatible with tools of PHYLIP version 3.695 (Felsenstein, 2005). First, input files were analyzed with restdist.exe in the fragment mode and in the next step, neighbor.exe tools was employed to obtain an unrooted tree. FigTree version 1.4.2 was used to visualize the cladogram.

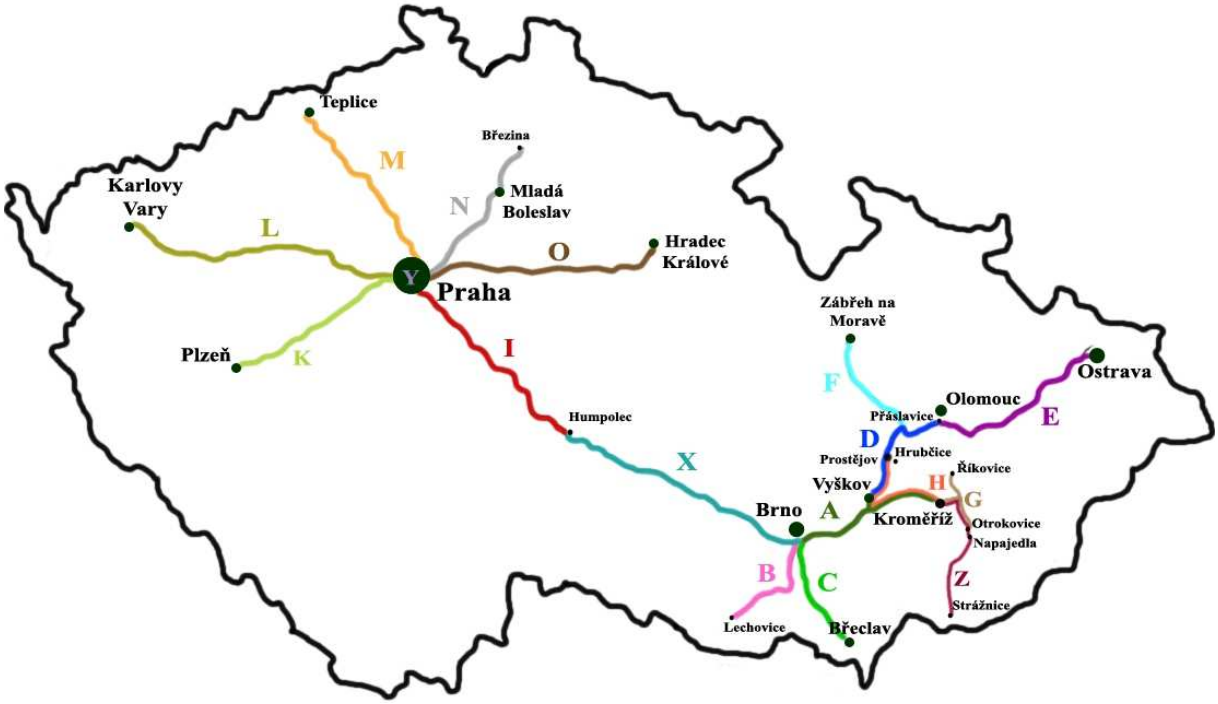


Figure 2: Geographical origin of *Bgh* isolates captured in the Czech Republic in June 2014. The color code of capture routes corresponds to color code for genotype visualization (Fig. 4).

Table II: *Bgh* isolates captured in June 2014 and their geographical origin

Capture route	Distance [km]	Isolates (total number of isolates)
Kroměříž - Vyškov - Brno	70	A1-2, A1-3, A1-4, A1-5, A1-6, A1-7, A1-9, A1-10, A1-11, A1-12, A1-13, A2-1, A2-3, A2-4, A2-5, A2-8, A2-9, A2-10, A2-11, A2-12, A2-13, A2-16, A2-17, A2-18, A2-19, A2-20, A2-22, A2-22, A2-23, A2-25, A2-27 (31)
Brno - Lechovice	55	B1-1, **B1-2**, B1-3, B1-4, **B2-1**, **B2-2**, B2-3, B2-4, B2-5 (9)
Brno - Břeclav	60	C1, **C2**
Vyškov - Olomouc - Přáslavice	55	D1, D2
Přáslavice - Ostrava Vítkovice - Mankovice - Velká Bystřice	165	E1, E3
Olomouc - Zábřeh na Moravě	45	F1
Kroměříž - Otrokovice - Říkovice - Kroměříž	70	G1
Prostějov - Kroměříž	50	H1, H2, H3
freway D1 km 90 - Praha Stodůlky	115	I1, I2, I3, I4, I5, **I6**, I7, I8, I9, **I10** (10)
Praha freeway D5 km 4 - Plzeň - Praha freeway D5 km 4	145	K1
Praha Stodůlky - Karlovy Vary	120	L1, L2, L3, L4, L5, L6, L7, L8, L9, L10, L11, L12 (12)
Praha - Teplice - Lovosice	115	M1, M2, **M3**, M4, M5, M6, M7 (7)
Praha - Mladá Boleslav - Březina	100	N1, N2, N3, N4, N5, N6, **N7**, N8 (8)
Praha - Hradec Králové - Chlumec nad Cidlinou	155	O1, O2, O3, **O4**, O5, O6, O7, O8, O9, O10, O11, O12, O13 (13)
Brno - freeway D1 km 90-199	220	X1, X2, X3, X4, X5, X6, X7, X8, X9 (9)
Pražský okruh (freeway D0)	40	Y1
Kroměříž - Napajedla - Strážnice - Kroměříž	345	Z1, Z2, Z3, Z4, Z5, **Z6**, Z7 (7)

isolates in **bold** were sequenced

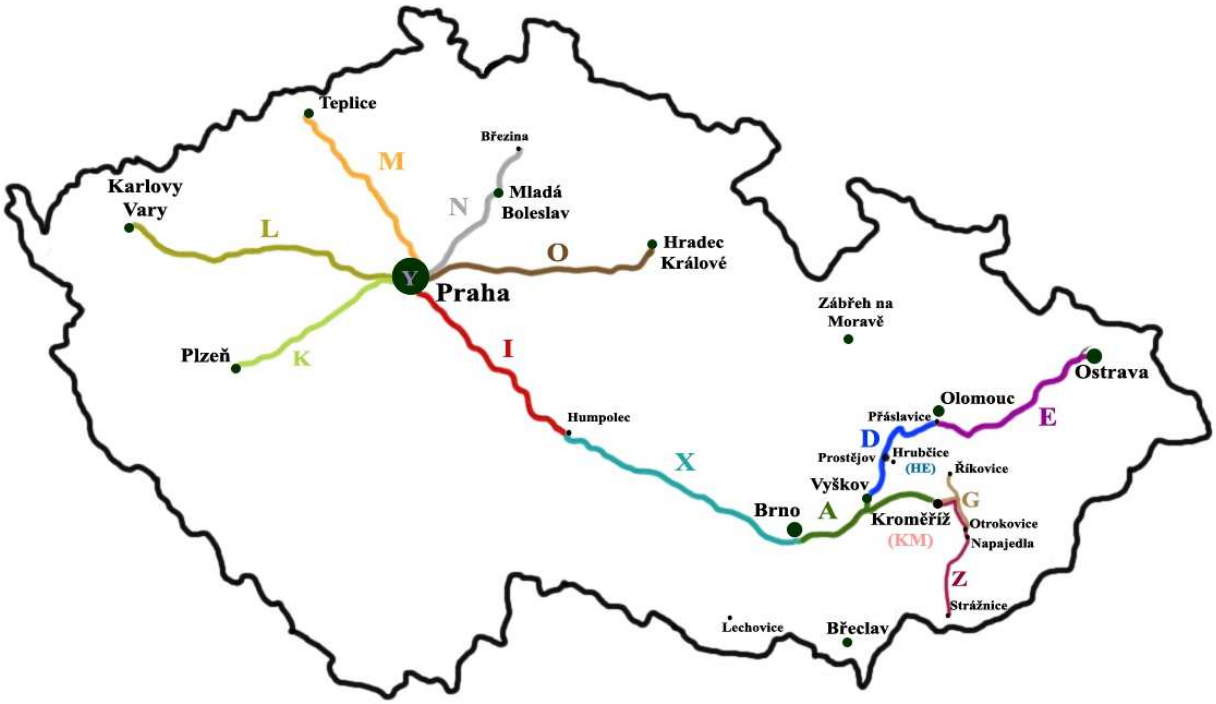


Figure 3: Geographical origin of *Bgh* isolates captured in the Czech Republic in June 2015. The color code of capture routes corresponds to color code for genotype visualization (Fig. 4).

Table III: *Bgh* isolates captured in June 2015 and their geographical origin

Capture route	Distance [km]	Isolates (total number of isolates)
Vyškov - Olomouc - Přáslavice	55	D1, D2, D3
Přáslavice - Ostrava Vítkovice - Mankovice - Velká Bystřice	165	E1, E2, E3, E4, E5, E6, E7, E8, E9, E10, E11, E12, E13 (13)
Kroměříž - Otrokovice - Říkovice - Kroměříž	70	G1, G2
Hrubčice - *collected manually*	-	HE2, HE3, HE4, HE5, HE6, HE7, HE8, HE9, HE10, HE11, HE12 (11)
freway D1 km 90 - Praha Stodůlky	115	I1, I2
Praha freeway D5 km 4 - Plzeň - Praha freeway D5 km 4	145	K1, K2, K3, K4, K5, K6 (6)
Kroměříž - *collected manually*	-	KM1, KM2, KM3, KM4, KM5, KM6, KM7, KM8, KM9, KM10, KM11, KM12, KM13, KM14, KM15, KM16, KM17, KM18, KM19, KM20, KM21 (21)
Praha Stodůlky - Karlovy Vary	120	L1, L2, L3, L4, L5, L6, L7, L8, L9, L10, L11L, L12, L13, L14, L15, L16, L17, L18 (18)
Praha - Teplice - Lovosice	115	M1, M2, M3, M4, M5, M6, M8, M9, M10, M11, M12, M13, M14, M15, M16, M17, M18, M19, M20, M21, M22, M23 (21)
Praha - Mladá Boleslav - Březina	100	N1
Praha - Hradec Králové - Chlumec nad Cidlinou	155	O1
Brno - freeway D1 km 90-199	220	X1
Pražský okruh (freeway D0)	40	Y1, Y2
Kroměříž - Napajedla - Strážnice - Kroměříž	345	Z1, Z2

Table IV: SSR markers tested for the presence of length polymorphism

	DH14 reference contig	Marker	Repeat motif	Primer sequence (5'-3')	Ta [°C]	Number of polymorphic alleles	Internal standard
●	000610	*obm31*	TGAA(28)	F CCAGACAATGTGTAAAGACTCCCTA R AGGGCTTCAGCCTGATACG	55	17	no
●	005527	*obm32*	TGAA(25)	F GATACAATCTAAAGAAGAACGTGTCA R CTAGGCTAGTGCCCGTCA	52	16	no
●	003454	*obm33*	CACT(24)	F TGACTCAAGCAGTACAATAAACAAA R TTTATAGTATCAGCTGCACTCTCG	52	11	no
	N/A	*obm34*	AAAT(15)	F GGACGTTAAACTAGATCAATGAATG R GGAGACTCCTATTGAGTTCTGTAAA	52	N/A	GPD F2/R3
●	003977	*obm35*	CACT(14)	F CCCAGGTTTTTAGAGGACACATA R ACCTGTACGCGGTGTCCTA	55	15	no
●	007505	*obm36*	TATT(13)	F TCTTTTACTAAAGCTGCTGGAACA R CGGGCACGACGTTAAGTATT	55	8 alleles could be scored reproducibly; possibly up to 13 alleles	none
●	003645	*obm37*	TTGA(11)	F GGCTAGCACGAGACCTTACG R CGGGCTAGACCTAGGACCTAAA	55	11	no
	N/A	*obm38*	ATG(30)	F CAGGTATGAAATCGATAATGGTAATG R TGTGCCATATCAGAGAAAGCAT	55	N/A	GPD F2/R2
	N/A	*obm39*	AAT(22)	F ATAGGGCGCCAAATCGGG R GCTTTGGGCCCTTGTCAGT	55	at least 12	no
	N/A	*obm40*	ATT(20)	F GCGCAGATTTCTGGCAAA R GGCAGATTTTTCAAAATGATGATG	55	N/A	no

Table IV - continued: SSR markers tested for the presence of length polymorphism

	DH14 reference contig	Marker	Repeat motif	Primer sequence (5'-3')	Ta [°C]	Number of polymorphic alleles	Internal standard
	N/A	*obm41*	ATT(20)	F CGCCAAATCAGGATAATGAA R GGACGTTAAACTAGATCAATGAATG	52	N/A	no
●	005896	*obm42*	TGA(19)	F TGAAAATTGAGGCTTGTCG R CTTTCAGACTTGCTTGGGTAA	52	15	no
●	005961	*obm43*	TAA(19)	F AGACTCAGACAGGACAATAAAATGA R CTTTGGGGTCGCCATATC	52	6	GPD F2/R3
●	001472	*obm44*	(CT)19	F AATTTTCGGCACCTCTTCC R CCGCCATACTTGGATCATTT	55	8	GPD F2/R3
●	000433	*obm45*	(TC)18	F TGTTTACTGTACCCTGTGAAACCT R CCTGATGAGGCTTTGGTATGA	55	8	no
●	004373	*obm46*	TG(16)	F CGACCAATACACTGGATTCTG R TTTCGACAAAAGTGCTACCAA	52	9	no
●	007133	*obm47*	AG(16)	F TCTCAGCTGTAGGGTGACGA R GACTCAAGCAGTACAATAACATACCC	55	8	no
	005585	*obm48*	TA(15)	F TCTCAGCTGTAGGGTGACGA R GACTCAAGCAGTACAATAACATACCC	55	N/A	no
	002693	*obm49*	AT(15)	F TCCATATCTTATTTACCCCGTATT R TGCTGATGCTGATGTTATGTG	52	N/A	no
●	003717	*obm50*	AC(14)	F GCCCTGTCTACCCTCACCT R TCACATCCATTTACGGCTGA	55	7	no
●	005538	*obm67*	(GAT)13	F TAGATGGACGAGGGAGCAGG R CATCGTTGAAGTCGCTGTGC	55	7	no

● indicates SSR markers used for genotyping

Table V: RJM markers tested for presence/absence variation

	DH14 reference contig	Marker	Primer sequence (5'-3')	Junction		*in silico* product size [bp]	Ta [°C]
				left element (class, order, superfamily) *	right element (class, order, superfamily) *		
●	003577	*obm51*	F TTTTATGTCCCATTTTTTTAACCA R TGAGCTTTCGAATTGTGGAA	Bast (retrotransposon, LTR, unknown)	Yho (retrotransposon, SINE, unknown)	486	59
	004339	*obm52*	F GGTTTTTATGGTCGGTCAGA R AGTATTCCGACAAAAAGCTATTCT	Sparta (retrotransposon, LTR, Gypsy)	Athena (retrotransposon, LTR, Copia)	485	50
	000086	*obm53*	F ACCTTGGCCTTGTATCTGAA R GCTTTGATGTTCCAAAATT	Sven (DNA-transposon, TIR, Mariner)	Lono (retrotransposon, SINE, unknown)	518	50
	003016	*obm54*	F GTCTGACTGTATTTCTCTGCATC R GGGCGACATGAATCTACCTA	Nova (DNA-transposon, unknown, unknown)	low copy DNA	484	50
	002565	*obm55*	F CGTAAGGAAGTGTTGAAATCGG R AGCCTCTGGGGATTATGAGG	Fuji (retrotransposon, SINE, unknown)	Crius (retrotransposon, LTR, unknown)	515	55
	002350	*obm56*	F AGTGTTACAGTTTTACCGTTTTAC R ACAGGCGTTGAAGTTTCG	low copy DNA	Tethys (retrotransposon, LTR, Gypsy)	504	50
	000296	*obm57*	F GCACTCAGTCAGAGGCCTG R TCGCTGCATCTCACAAAGTT	Yho (retrotransposon, SINE, unknown)	Nova (DNA-transposon, unknown, unknown)	505	55
	003639	*obm58*	F TTCCGTGACACTCCCAGATGC R CGAACTTCACGCGGTTCTTT	Sparta (retrotransposon, LTR, Gypsy)	Agemem (retrotransposon, SINE, unknown)	518	55
	008599	*obm59*	F CTCGTTTTAGCAGCCCTTC R TGACAGGAGTTGAGAATAGG	Hestia (retrotransposon, LTR, Gypsy)	Lie (retrotransposon, SINE, unknown)	524	50
	004219	*obm60*	F TCTTGCTGTAGCTTCAACACC R TGTCCACATCGTTTCATCTTC	Nova (DNA-transposon, unknown, unknown)	Crius (retrotransposon, LTR, unknown)	503	55

● indicates RJM marker used in this study

5 Results

The main aims of the thesis were i) development of molecular markers which could unambiguously discriminate isolates of *Blumeria graminis* f.sp. *hordei* from the Czech Republic and ii) their employment in studying *Bgh* diversity in the Czech Republic. In order to do so, the essential steps were first, extraction of genomic DNA of sufficient quality and quantity, second, whole-genome sequencing of ten selected *Bgh* isolates and lastly, identification of candidate polymorphisms (RJMs, SSRs and SNPs) and their validation.

5.1 DNA extraction and NGS Illumina sequencing

Genomic DNA extraction yields ranged from 62 ng to 640 ng (mean 230 ± 102.6 ng) using 10-15 mg of *Bgh* spores. About 50 pg of DNA was sufficient for PCR reactions.

Ten *Bgh* isolates exhibiting the largest virulence variability (unpublished data) were sequenced with Illumina MiSeq. The RAW sequences with coverage from 10 to 23 fold were processed and assembled using MaSuRCA assembler (Tab. VI). The obtained reads (SNP markers) and resulting assemblies (RJM and SSR markers) were used for marker development.

Table VI: Sequencing and assembly statistics for ten selected *Bgh* isolates

Isolate	B1-2	B2-1	B2-2	C2	I6	I10	M3	N7	O4	Z6	Mean
Coverage	18	13	10	10	15	14	14	14	23	12	14
No. of scaffolds	218,469	25,859	21,302	15,566	27,217	41,390	42,542	13,685	77,903	6,136	49,007
Min. scaffold length [bp]	185	300	300	374	300	300	303	380	300	334	307.6
Max. scaffold length [bp]	38,890	9,128	10,332	8,110	9,070	23,099	28,137	8,416	48,206	6,400	18978.8
N_{50} [bp]	1,088	1,455	1,374	1,400	1,545	2,097	2,454	1,217	2,402	847	1587.9

5.2 Marker development

The analysis of NGS sequences of ten *Bgh* isolates identified 84,471 ISBP/RJM polymorphisms, 65,535 SNPs and 9,068 SSR polymorphisms. Out of all identified SNPs, only those polymorphic in four to five *Bgh* isolates were considered as suitable candidates for marker development. Based on polymorphisms identified in the whole-genome sequence of *Bgh* isolate O4 (isolate with highest sequencing coverage), primer pairs for 10 candidate RJMs (Tab. V) and 21 candidate SSRs (Tab. IV) were designed. To assess SNP, variability in all sequenced isolates was considered and primer pairs for 20 amplicons containing at least one SNP (Tab. X) were designed.

Seven primer pairs were designed to target SSRs with tetra-nucleotide repeat units, seven for tri-nucleotide repeat units and seven for di-nucleotide repeat units. However, five of seven markers developed for tri-nucleotide units were not used for genotyping, and thus markers newly added to the current genotyping panel targeted primarily motifs with tetra- and di-nucleotide repeat units (~ 43 % of the panel each). Primer pairs were preferentially designed for SSRs with highest number of repeat units detected in the whole-genome sequence of *Bgh* isolate O4. The relationship between number of repeat units and number of identified alleles could thus be examined. SSR sequences with the highest number of *in silico* identified subunits were those consisting of tri- and tetra-nucleotide repeat units. The two reproducible SSR markers with tri-nucleotide repeat units both contained 19 subunits in the SSR motif. Six reproducible SSR markers with tetra-nucleotide repeat units, on average contained 19 repeat units in a SSR motif (minimum of 11 and maximum of 28 repeat units per a SSR motif). In contrast, six reproducible SSR markers with di-nucleotide repeat units consisted of 16 repeat units on average (minimum of 13 and minimum of 19 repeat units). Depending on the length of a repeat unit, an average of 14, 11.5 and 7.8 alleles were identified in SSR motifs with repeat units 4, 3 and 2 nucleotides in length, respectively (Tab. VII).

SSR markers were first tested on a subset of eight *Bgh* isolates to confirm presence of polymorphic patterns. In the initial screening, all primer pairs showed polymorphisms and were thus used on *Bgh* isolates from at least one season. However, seven SSR markers were not suitable for genotyping because of i) non-reproducible patterns when an internal standard was used at the same time *(obm34* and *obm38),* ii) excess satellite band production *(obm39, obm40* and *obm41)* or iii) smeared PCR products *(obm48* and *obm49)* under the used PCR conditions. Taken together, 14 SSR markers with reproducible polymorphic patterns enabled scoring of 141 polymorphic alleles (Tab. IX).

Table VII: Number of *in silico* identified repeat units and number of experimentally confirmed alleles for SSR markers used in this study.

	Marker	Number of subunits in O4 isolate	Number of experimentally identified alleles
tetra-nucleotide basic repeat units	*obm31*	28	19
	obm32	25	18
	obm33	24	11
	obm35	14	17
	obm36	13	8
	obm37	11	11
average		**19.1**	**14**
tri- nucleotide basic repeat units	*obm42*	19	17
	obm43	19	6
average		**19**	**11.5**
di- nucleotide basic repeat units	*obm44*	19	8
	obm45	18	8
	obm46	16	9
	obm47	16	8
	obm50	14	7
	obm67	13	7
average		**16**	**7.8**

Besides development of new primer pairs, four SSR markers designed previously by Komínková *et al.* (2016, submitted) were used on the *Bgh* collections from both seasons. Except for alleles already identified in the collection of Czech *Bgh* isolates captured in 2012, new alleles were identified for both primer pairs targeting a SSR sequence with tetra-nucleotide repeat unit. In the case of *obm24*, there were four alleles described originally and another allele (140 bp in length) was identified in the *Bgh* isolates collected in 2014 and 2015. Marker *obm29* was originally characterized with five alleles only (Komínková *et al.* 2016 submitted) but additional four alleles (104 bp, 100 bp, 96 bp and 86 bp in length) were revealed in season 2014 and confirmed in season 2015 (Tab. VII).

Table VIII: Nomenclature and PCR product sizes of *obm24* and *obm29* alleles newly described in *Bgh* isolates captured in seasons 2014 and 2015

Marker	Alleles	Allele size estimation [bp]
obm24	*obm24.6*	140
obm29	*obm29.6*	104
	obm29.7	100
	obm29.8	96
	obm29.9	86

RJMs were also tested on a set of eight *Bgh* isolates first. Although only *obm51* marker revealed a polymorphic presence/absence pattern in the initial testing, all RJMs were tested on the *Bgh* collection from season 2014 to confirm this finding. Indeed, out of ten candidate RJMs designed to capture the variability of identified TE junctions, nine primer

pairs were excluded i) for the lack of presence/absence variation *(obm52, bom53, obm55)*, ii) because multiple target sequences were amplified *(obm56, obm57, obm58, obm59)* or iii) the amplification was not reproducible *(obm54, obm60)*. The only polymorphic RJM *(obm51)* was designed for a LTR-SINE junction and the primer pair yielded a PCR product of expected length in 70 % and 90 % of screened *Bgh* isolates collected in 2014 and 2015, respectively.

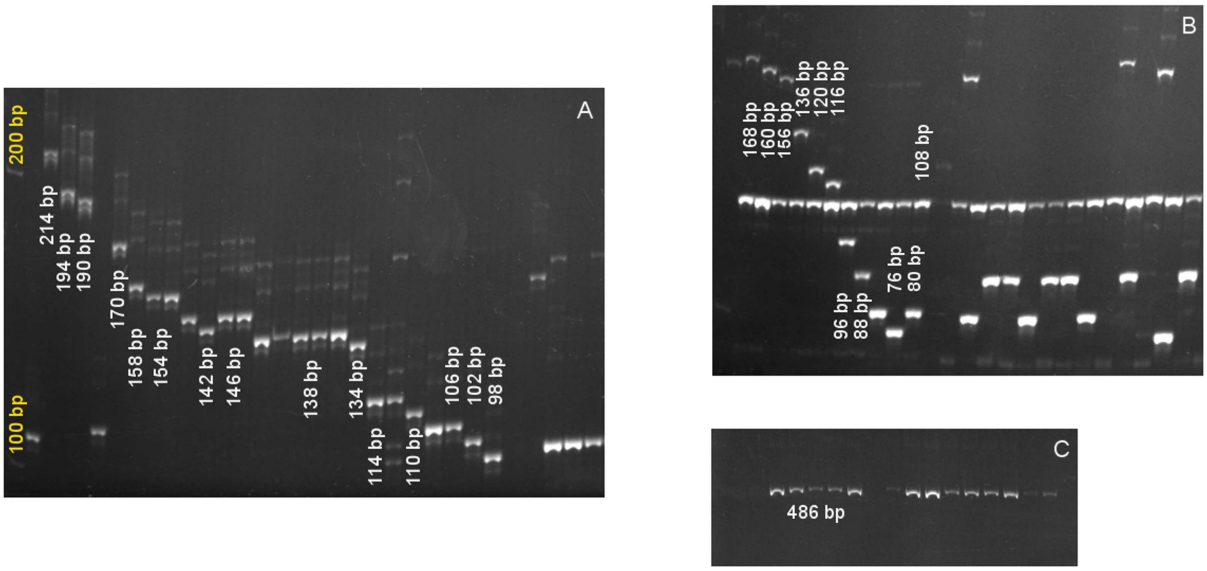


Figure 3: Selected polymorphic markers. (**A)** *obm32* - 15 out of 18 identified alleles, first lane: GeneRuler 100 bp Plus DNA Ladder. (**B)** *obm33*. (**C)** *obm51*. Alleles of SSR markers *obm32* and *obm33* were resolved on 6% polyacrilamide gels. PCR products obtained with RJM marker *obm51* were separated on 4% polyacrilamide gels. PCR product size estimations were based on migration of GeneRuler 100 bp Plus DNA Ladder (Thermo Fisher Scientific, USA).

Primer pairs designed to assess variability in SNPs (Tab. X), were tested on four *Bgh* isolates to i) confirm presence of a single PCR product of an expected size and ii) exclude amplifications of heterogeneous sequences. Only one primer pair *(obm80)* yielded several bands and primer pair *obm77* was excluded from further use due to presence of several amplicons of the same size but varying in sequence. Moreover, two primer pairs *(obm72* and *obm75)* failed to amplify the target DNA under the standard PCR conditions. Out of the remaining 16 primer pairs, 3 randomly chosen primer pairs *(obm73, obm81* and *obm86)* were used on the *Bgh* collection. The amplicon 358 bp in length contained 3 SNPs *(obm73)*, the 400bp amplicon contained 2 SNPs *(obm81)* and the amplicon 376 bp long contained 3 SNPs *(obm86)*. Out of the eight nucleotide variants, seven nucleotide variations were present in both seasons but one *(obm81.1)* emerged in the season 2015 for the first time.

Furthermore, SNPs identified in a collection of Czech *Bgh* isolates from 2012 (Komínková *et al.*, 2016, submitted) were scored as well. Three primer pairs described previously *(obm6, obm9* and *obm20)* were used to assess SNPs in the obtained amplicons. In sequences amplified with *obm6* and *obm20* primer pairs, new SNPs were discovered. In the case of *obm6*, three new SNPs were revealed alongside the original four SNPs. In amplicons produced with *obm20* primer pair, there were seven variable positions in 2012 and additional two variants were revealed in 2014 and confirmed to present in isolates captured in 2015.

5.3 Diversity of *B. graminis* f.sp. *hordei* in the Czech Republic in seasons 2014 and 2015

The genotyping panel applied on 119 *Bgh* isolates collected in 2014 and 113 *Bgh* isolates from season 2015 comprised 18 most polymorphic and most reliable markers: 14 SSR markers and 4 primer pairs targeting sequences with SNPs. In total, 158 polymorphisms could be scored with this panel, 141 of polymorphisms being SSR alleles (~ 89 % of the panel) and 17 polymorphisms being SNPs (~ 11 %). RJM markers were not used because of their lower information value and lack of time. Out of 232 *Bgh* isolates, 97.4 % were identified as unique genotypes. Two isolates (A1-5 and A1-12) could not be unambiguously identified in the collection from 2014 and two more undistinguishable pairs (HE4 and HE5, HE11 and HE12) occurred in 2015 (Fig. 4). The unresolved pairs of *Bgh* isolates originated in the same geographical region but showed different virulences in differential tests (unpublished data).

While there are 12 sister groups formed by *Bgh* isolates (Fig. 4), i.e. pairs of *Bgh* isolates with similar genotypes, that were captured in the same region (unresolved pair A1-12_2014/A1-5_2014 and A1_2015, A2-16_2014 and A4_2015, E1_2015 and E11_2015, E10_2015 and E13_2015, I3_2014 and I10_2014, KM5_2015 and KM6_2015, KM13_2015 and KM19_2015, L4_2015 and L8_2015, L5_2014 and L13_2015, L8_2014 and L9_2014, M1_2014 and M13_2015, N2_2014 and N7_2014), the overall topology of a cladogram obtained with neighbor-joining algorithm does not reveal any correlation between genotypes of *Bgh* isolates and their geographical origin (Fig. 4). Sister groups formed by *Bgh* isolates captured in the same location but in different seasons are of an interests: unresolved pair A1-12_2014/A1-5_2014 and A1_2015, L5_2014 and L13_2015, M1_2014 and M13_2015. However, any generalizations about inter-seasonal similarity of predominant genotypes originating in the same region cannot be drawn. Isolates with a similar set of polymorphisms A1-10_2014 and M7_2015 were captured at locations possibly as far away as 300 km.

Similarly, isolates L1_2014 and E12_2015 originated from locations up to 390 km apart. Furthermore, 4 *Bgh* isolates captured in 2015 (A3, KM11, KM12, M3) form 4 individual branches separated from the remaining 228 *Bgh* isolates (next to the root of the cladogram).

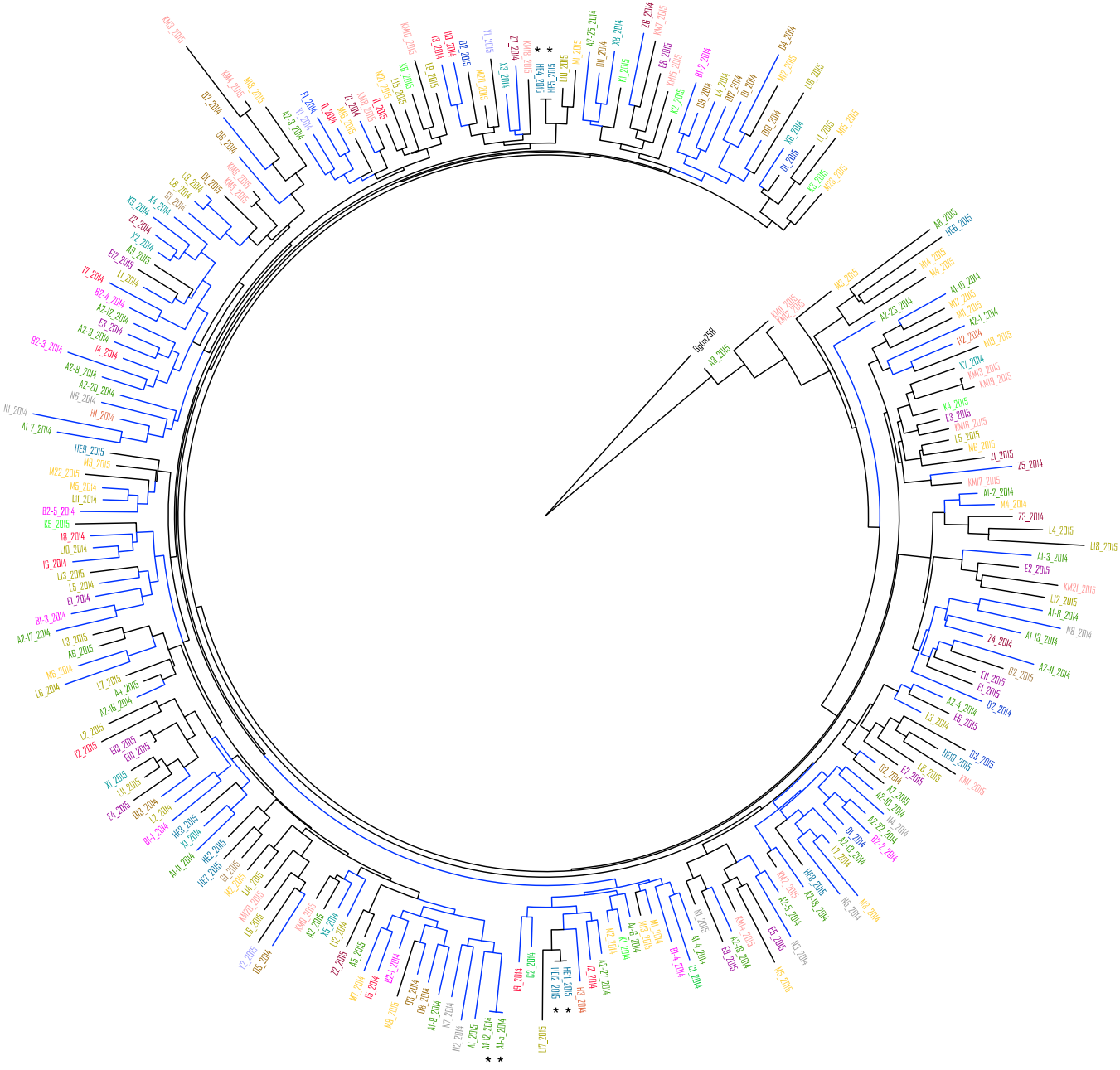


Figure 4: A cladogram constructed with neighbor-joining algorithm (PHYLIP) to visualize variability of genotypes identified with a set of 18 primer pairs. In the amplicons, 158 polymorphisms were present: 141 SSR alleles and 17 SNPs. Out of 232 *Bgh* isolates, 97.4 % were characterized with unique set of polymorphisms and thus unambiguously identified. The six *Bgh* isolates that could not be resolved with used marker panel (asterisk), however, possessed different virulences (data not shown), revealing limited resolution power of used marker panel to distinguish *Bgh* isolates originating from rather limited geographical locations. Taxon color code corresponds to the geographical origin of *Bgh* isolates (see Figs. 2 and 3). Branches in blue represent *Bgh* isolates captured in 2014, black branches represent *Bgh* isolates capture in 2015. *Bgtm258* isolate was used as an outgroup.

Table IX: Alleles and estimated PCR product sizes of markers added to the genotyping panel of *Blumeria graminis* f.sp. *hordei*

Marker	Alleles	Allele size estimation [bp]
obm31	*obm31.1*	340
	obm31.2	260
	obm31.3	236
	obm31.4	202
	obm31.5	194
	obm31.6	190
	obm31.7	186
	obm31.8	178
	obm31.9	154
	obm31.10	150
	obm31.11	142
	obm31.12	138
	obm31.13	114
	obm31.14	78
	obm31.15	66
	obm31.16	54
	obm31.17	214
	obm31.18	122
	obm31.19	86
obm36	*obm36.1*	96
	obm36.2	92
	obm36.3	84
	obm36.4	80
	obm36.5	76
	obm36.6	68
	obm36.7	120
	obm36.8	88

Marker	Alleles	Allele size estimation [bp]
obm32 (Fig. 3)	*obm32.1*	214
	obm32.2	194
	obm32.3	190
	obm32.4	170
	obm32.5	158
	obm32.6	154
	obm32.7	146
	obm32.8	142
	obm32.9	138
	obm32.10	134
	obm32.11	114
	obm32.12	110
	obm32.13	106
	obm32.14	102
	obm32.15	98
	obm32.16	90
	obm32.17	86
	obm32.18	78
obm37	*obm37.1*	132
	obm37.2	124
	obm37.3	116
	obm37.4	112
	obm37.5	108
	obm37.6	104
	obm37.7	100
	obm37.8	96
	obm37.9	84
	obm37.10	80
	obm37.11	76

Marker	Alleles	Allele size estimation [bp]
obm33 (Fig 3)	*obm33.1*	168
	obm33.2	160
	obm33.3	156
	obm33.4	136
	obm33.5	120
	obm33.6	116
	obm33.7	108
	obm33.8	96
	obm33.9	88
	obm33.10	80
	obm33.11	76
obm42	*obm42.1*	127
	obm42.2	121
	obm42.3	118
	obm42.4	106
	obm42.5	103
	obm42.6	88
	obm42.7	79
	obm42.8	76
	obm42.9	70
	obm42.10	67
	obm42.11	109
	obm42.12	100
	obm42.13	94
	obm42.14	91
	obm42.15	85
	obm42.16	130
	obm42.17	184

Marker	Alleles	Allele size estimation [bp]
obm35	*obm35.1*	160
	obm35.2	152
	obm35.3	148
	obm35.4	144
	obm35.5	140
	obm35.6	128
	obm35.7	124
	obm35.8	120
	obm35.9	108
	obm35.10	104
	obm35.11	100
	obm35.12	92
	obm35.13	84
	obm35.14	132
	obm35.15	156
	obm35.16	128
	obm35.17	116
obm43	*obm43.1*	116
	obm43.2	113
	obm43.3	110
	obm43.4	107
	obm43.5	101
	obm43.6	89

Table IX - continued: Alleles and estimated PCR product sizes of markers added to the genotyping panel for *Blumeria graminis* f.sp. *hordei*

Marker	Alleles	Allele size estimation [bp]	Marker	Alleles	Allele size estimation [bp]	Marker	Alleles	Allele size estimation [bp]
obm44	*obm44.1*	110	***obm45***	*obm45.1*	100	***obm46***	*obm46.1*	98
	obm44.2	104		*obm45.2*	96		*obm46.2*	96
	obm44.3	101		*obm45.3*	94		*obm46.3*	94
	obm44.4	99		*obm45.4*	91		*obm46.4*	92
	obm44.5	97		*obm45.5*	89		*obm46.5*	90
	obm44.6	95		*obm45.6*	83		*obm46.6*	86
	obm44.7	91		*obm45.7*	79		*obm46.7*	84
	obm44.8	85		*obm45.8*	81		*obm46.8*	82
							obm46.9	76
obm47	*obm47.1*	103	***obm50***	*obm50.1*	92	***obm67***	*obm67.1*	112
	obm47.2	99		*obm50.2*	86		*obm67.2*	85
	obm47.3	97		*obm50.3*	80		*obm67.3*	76
	obm47.4	93		*obm50.4*	72		*obm67.4*	70
	obm47.5	91		*obm50.5*	84		*obm67.5*	64
	obm47.6	85		*obm50.6*	82		*obm67.6*	55
	obm47.7	95		*obm50.7*	88		*obm67.7*	88
	obm47.8	105						

Table X: Primer pairs designed for amplicons containing SNPs. SNPs in sequences amplified with primer pairs *obm73, obm81* and *obm86* were confirmed experimentally, for the rest of primer pairs, SNP variants are expected variations based on sequence analysis.

	DH14 reference supercontig	marker		primer sequence (5'-3')	Ta [°C]	SNP variants		SNP position*	*in silico* product size [bp]	notes
●	contig 00781	*obm20*	F	AGTCCGTGACACCTACTCAGG	55	obm20.8	A/T	469	524	new polymorphisms in a primer pair developed by Komínková *et al.* (2016, submitted)
			R	TGAGGCTTGGGTAGAAGTCAA		obm20.9	single A insertion	471		
	000195	*obm71*	F	TCCATACACGGCTAAGAACA	55		A/G	84	332	to be used in the future
			R	CACACAGCACAGCACACATG						
†	001608	*obm72*	F	ACCGAGTCCCTTTCTATTTAAGCT	55		C/T	84	307	does not amplify
			R	TCCGTTCACAGTAACCCAGC						
●	003540	*obm73*	F	GCTAGCAGCGAGTGTCCTTA	55	obm73.3	A/C	104	358	included in the genotyping panel
			R	CCGCTTCGATCTTTCAAAACGA		obm73.2	A/C	108		
						obm73.1	C/T	270		
	005427	*obm74*	F	CTTCTCTAACCACCGCCGTA	54		C/T	52	302	to be used in the future
			R	CAGCCAGTACGCCAAAGAAA						
†	005442	*obm75*	F	TTTCTTCGCACTCACCCTCC	54		C/T	230	388	does not amplify
			R	AGTAATGGTCAAATGGTAGTGGG						
	005453	*obm76*	F	CCATAACCTTTCCCAAGCACC	55		A/C	221	433	to be used in the future
			R	TAGAGTGGCAGCTGTGTTGG						
†	005459	*obm77*	F	GTGGATGGCTGTGGTTTTGG	55		C/T	194	400	yields heterogenous sequences
			R	CCAGAGACCACCACCAAGAG						
	005464	*obm78*	F	ATTCGCCAAAGAATTGCATC	55		C/T	211	437	to be used in the future
			R	TAAGGCATGGATGCTGTCTG						
	005465	*obm79*	F	TTGTGTGCAGGTTATGGAGT	52		A/G	232	440	to be used in the future
			R	TCATGCACGGACCTTTGTAT						
†	005466	*obm80*	F	CGGTCGGACAATGGCCTT	54		A/G	247	496	several PCR products
			R	ACGAAATCTCTGCTTGTAACTCT						

Table X - continued: Primer pairs designed for amplicons containing SNPs. SNPs in sequences amplified with primer pairs *obm73, obm81* and *obm86* were confirmed experimentally, for the rest of primer pairs, SNP variants are expected variations based on sequence analysis.

	DH14 reference supercontig	marker		primer sequence (5'-3')	Ta [°C]	SNP variants		SNP position*	*in silico* product size [bp]	notes
●	005469	*obm81*	F	TGGCAAATTTTCAGAGGCAGG	55	obm81.2	A/G	167	400	included in the genotyping panel
			R	CGAAACAGCGGCAAACTTGA		obm81.1	A/T	187		
	005470	*obm82*	F	GCCGCGATTTCTTTCTTGCA	55		A/G	127	303	to be used in the future
			R	GCTGAATTCCGCCGGTAAAC						
	005472	*obm83*	F	ACCATTTGTCTCTAGGGCCG	54		A/G	197	443	to be used in the future
			R	TCTGTGGGTACGAGCATTGT						
	005474	*obm84*	F	AGAACCACTCAGACTACCCA	54		A/G	239	400	to be used in the future
			R	AGCTAAACAGTAAACCTCCCTTGA						
	005478	*obm85*	F	TCTATGTGTTGTTTGAGTCACTCCA	55		C/T	232	400	to be used in the future
			R	GTGTGGCCTGGTTGAGAGAA						
●	005483	*obm86*	F	CCCACCTCAGACTCCACCTA	55	obm86.1	A/G	214	376	included in the genotyping panel
			R	AGTCAGGAAAACCATGGGCA		obm86.2	A/T	223		
						obm86.3	C/T	226		
	005483	*obm87*	F	GGTCAAGGCCTAGTCCACTT	55		A/G	161	343	to be used in the future
			R	TCCTCGCAACGTGAATGACA						
	005486	*obm88*	F	ATGGTCCCTCAGCGATTTGG	56		A/T	63	310	to be used in the future
			R	AATCAAGGGTGCGTGATGGG						
	005496	*obm89*	F	TCTCGTCACTAAGGATCCCGA	55		A/G	186	422	to be used in the future
			R	TGCGAGTCATGTGTTGTGGA						
	005502	*obm90*	F	CCAACACGTCCAGATTCCCA	55		A/G	218	413	to be used in the future
			R	GCCCTACCATCCATCGCTAC						

† indicates primer pairs excluded from the genotyping panel; ● indicates primer pairs used for genotyping; * SNP position determined from the first nucleotide of the forward primer

6 Discussion

Blumeria graminis f.sp. *hordei* is an agriculturally important fungal pathogen. Although studied intensively, the pathogen phylogeny have been primarily focused on ff. spp. classification (e.g. Troch *et al.*, 2014) or on a position of *Bgh* among other powdery mildews (Takamatsu, 2013b) because the resolution of DNA markers was not sufficient to assess diversity within a single *forma specialis*. Diversity studies of *Bgh* populations have therefore relied primarily on phenotyping, i.e. assessment of virulences or fungicide resistance (Mcdonald and Mcdermott, 1993). The earliest non-phenotyping-based diversity studies of cereal powdery mildews were based on biochemical markers. Isoenzyme variants were sufficient to identify predominant genotypes in each of the cereal ff. spp. but isozyme analysis of *Bgh* isolates from a number of European countries revealed only low polymorphism level (Koch and Köhler, 1990; Koch and Köhler, 1991). Later on, DNA markers started to be used in population studies: RFLPs (Damgaard and Giese, 1996; O'Dell *et al.*, 1989), RAPDs (Mcdonald and Mcdermott, 1993; Wolfe and McDermott, 1994) or SCARs combined with RAPDs (Caffier *et al.*, 1999). In a *Bgh* genetic map construction, including mapping of seven avirulence genes, AFLPs and RFLPs were employed (Pedersen *et al.*, 2002).

Although the earliest population studies utilizing DNA markers illustrate the need for more molecular studies, development of additional markers has remained limited. Previous analysis of 15 *B. graminis* isolates (12 *Bgh* isolates of Czech origin, 2 *Bgh* isolates from the South African Republic and a *Bgt* isolate) has revealed an extremely low sequence diversity in coding sequences (Komínková, 2011). Thus, use of non-coding sequences to assess diversity has been thus proposed: TEs constitute for a large portion of the *Bgh* genome (Amselem *et al.*, 2015), SSRs have been widely employed in population genetics (Cavalli-Sforza, 1998; Sunnucks, 2000) and SNPs have become increasingly popular DNA marker (Morin *et al.*, 2004), yet none of the markers has been exploited systematically in *Bgh*. In the pioneering work by Komínková (2013), random sequence scaffolds from reference sequence of *Bgh* isolate DH14 (Spanu 2010) were selected. The scaffolds were annotated and putative ISBP and SSR features were tested as potential sources of polymorphisms within isolates from the Czech Republic. In contrast to gene- and ITS-based markers, SSR, ISBP, and SNP markers from non-coding regions showed fairly good level of polymorphism for marker development and enabled unambiguous discrimination of 73.2 % of 97 Czech isolates. To further increase the reliability and efficiency of marker development, the present study aimed

to develop new markers from genomic sequences of ten *Bgh* isolates from the Czech Republic. The strategy for marker discovery employed here yielded 84,471 candidate ISBP, 65,535 candidate SNP and 9,068 candidate SSR markers. Primer pairs were designed for 10 RJMs, 20 amplicons containing at least one SNP and 21 SSR motifs.

Bgh TEs have remained largely unexploited as a source of DNA markers, except for mapping of 42 LTR-retrotransposon loci (Pedersen *et al.*, 2002). In the above mentioned work of Komínková *et al.* (2016, submitted), 20 ISBP/RJM markers were designed but only two primer pairs enabled detection of reproducible presence/absence variation (PAV). Ten RJM primer pairs designed here targeted a variety of TE junctions but only one RJM yielded reproducible PAV pattern (Tab. V). In other words, even though ISBP/RJM polymorphisms were the most abundant, the success rate remained the lowest (10 %). The low reliability of the ISBP/RJM polymorphic sites could be due to a combination of insufficient sequencing coverage and presence absence nature of the polymorphism or due to high conservation of the insertion sites. However, Oberhaensli *et al.* (2011) compared TE content and distribution between two selected genomic regions of *Bgh* and closely related *Bgt* and they revealed very limited conservation of TEs between the two ff. spp.. This is indicative of intensive TE reshuffling since the divergence of *Bgh* and *Bgt*, which is estimated to occur 10 million years ago. Given the high activity of TEs, it is thus surprising that RJMs developed in this study failed to capture more variability in TE junctions in a collection of over 200 isolates. Komínková *et al.* (2016, submitted) applied their set of ISBPs/RJMs on a worldwide *Bgh* collection but neither this approach resulted into a higher degree of polymorphism.

Although Czech isolates are very diverse in their genotype (Fig. 4), it is possible that transposition does not occur continuously to create significant variability between isolates of a single *forma specialis*. Alternatively, the distance of TE insertion site from genes might play a role as TEs orthologous between *Bgh* and *Bgt* were less likely to be identified with increasing distance from coding regions (Oberhaensli *et al.*, 2011). If this is the case, future development of ISBPs and RJMs should consider DNA adjacent to TE junction. Nevertheless, RJM primer pairs yielding a single PCR amplicon *(obm52, bom53, obm55)* in all isolates might be a source of SNPs as described before (Komínková *et al.*, 2016 submitted) and their amplicons will be examined for SNP presence in the future.

High abundance of TEs in the *Bgh* genome is promising, but if their multiplication occurred within a limited time span, it may explain the results observed here. On the other hand, regions flanking the insertion sites may have accumulated mutations and REMAP markers (Kalendar and Schulman, 2006) may be more efficient than RJMs. However, this

hypothesis has to be tested. Additionally, a procedure to employ MITEs as polymorphic markers in rice have been described but unfortunately in the *Bgh* genome, MITEs constitute for less than 1 % (Amselem *et al.*, 2015; Casa *et al.*, 2000). In the ascomycete *Fusarium oxysporum* f.sp. *dianthi* race-specific primer pairs targeting *impala* and *Fot1* DNA transposons were successfully employed to differentiate several races. However, it failed to distinguish very closely related pathotypes undistinguishable by other approaches, including RFLP and RAPD profiles (Chiocchetti *et al.*, 1999). This observation is thus another example of limited variability in TEs between closely related samples. Because of TE abundance in the *Bgh* genome, as described e.g. by Amselem *et al.* (2015) and confirmed here, and expected association of TEs with genes involved in virulence, e.g. effector proteins, TEs and their utilization as molecular markers should continue to receive attention in the future.

20 primer pairs to score SNPs were designed. Two of them did not yield any amplification products and two primer pairs targeted multiple genomic regions (Tab. X). In total, 4 primer pairs were thus excluded from further use, while the remaining 16 primer pairs (80 %) are a promising source of SNPs. Out of candidate 16 primer pairs, 3 randomly chosen primer pairs were tested for reliability. Amplicons of an average length of 378 bp contained 2.67 SNPs per amplicon on average. The observed presence of several SNPs within a relatively short genomic region is in agreement with distribution of SNPs discovered previously (Komínková *et al.*, 2016 submitted). Hacquard *et al.* (2013) also suggested an existence of DNA blocks rich in isolate specific SNPs. Together with high reproducibility, genome-wide distribution and automated analysis, SNPs are powerful type of polymorphism for future high-throughput studies.

Out of three types of polymorphisms employed in the present study, SSRs were the least frequent. Low number of SSR loci has been identified in other fungi, too (Dutech *et al.*, 2007). However, the multi-allelic nature of SSRs makes them highly informative. Unlike in the case of TE-based markers, SSR markers for *B. graminis* have been developed previously (Tucker *et al.*, 2015; Wang *et al.*, 2014). Komínková *et al.* (2016 submitted) used the reference genome sequence to design 10 primer pairs for SSR motifs with di-, tri-, tetra- and hexa-nucleotide repeat units. When applied on a collection of Czech isolates captured in 2012, four of the primer pairs resulted into reproducible polymorphic patterns and the average number of alleles per primer pair was 4.5. The same primer pairs were also used to screen *Bgh* isolates from 2014 and 2015 and additional three alleles were identified, thus increasing the average allele number to 5.5 (data not shown). Interestingly, the new alleles emerged in SSR motifs consisting of tetra-nucleotide repeat units. Although the given primer set is way too

small to draw any general conclusions, increased variability in SSRs with longer repeat units agrees with observation made by Dutech *et al.* (2007) and Wang *et al.* (2014) and confirmed in this work.

SSR markers developed in other laboratories were used to study diversity of Australian *Bgh* and Chinese *Bgt* populations. In development of SSR markers for *Bgh*, Tucker *et al.* (2015) also used the reference genome sequence. Out of 30 primer pairs, 8 were polymorphic and on average yielded 7 alleles per primer pair. Wang *et al.* (2014) introduced a protocol to discover SSRs in enriched *Bgt* libraries. The enrichment-based strategy resulted into 31 primer pairs but only 5 yielded polymorphic products, with the mean number of alleles being 5.8. Here, reproducible SSR markers yielded on average 11 alleles per primer pair. DNA replication slippage is more likely to occur in sequences consisting of high number of basic repeats and as a result, SSR motifs consisting of high number of repeat units produce higher number of alleles (Dutech *et al.*, 2007; Wang *et al.*, 2014). Such correlation was observed in our set of SSR markers, too (Tab. VII). SSR markers developed here included a significant proportion (~ 43 %) of SSRs with tetra-nucleotide repeat units, while the primer pairs of Tucker *et. al* (2015) flanked SSRs with di- and tri-nucleotide repeat units only. This might have contributed to significantly increased number of polymorphic alleles, farther supporting the observations of Dutech *et al.*, 2007 and Wang *et al.*, 2014. Taken together, SSR discovery strategies by both Tucker *et al.* (2015) and Wang *et al.* (2014) were successfully employed to assess diversity of 90 or more *B. graminis* isolates but a strategy based on NGS seems to enable more rapid SSR discovery with higher success rate. It was also possible to mine specifically for SSR with longer repeat units, thus increasing chances for a high number of alleles per marker. Finally, SSR loci longer than eight repeats were reported to be hard to isolate (Dutech *et al.*, 2007) and NGS-based SSR discovery could overcome that.

Another limitation of enrichment-based strategy is that SSR motif to be enriched for has to be known in advance. In contrast, NGS-based SSR discovery proposed here enabled primer design for any motifs and moreover, showed significantly higher success rate. 21 SSR markers were designed and reproducible polymorphic patterns were obtained with 14 primer pairs (~ 67 %). It is hypothesized that the relatively high proportion of polymorphic markers was achieved thanks to utilization of whole genome sequences of multiple *Bgh* isolates. The fact that isolates originating in the same geographical region like all isolates screened might have contributed as well.

In an enriched library prepared by Dutech *et al.* (2007), SSRs with tri-nucleotide repeat units were scarce. In SSR-enriched libraries of peanut *Arachis hypogaea* L., the efficiency of tri-nucleotide SSR isolations was significantly lower in comparison to di-nucleotide SSR library (Yuan *et al.*, 2010). In the present study, primer pairs for tri-nucleotide repeat motifs showed significantly reduced reliability compared to primer pairs for SSRs with di- and tetra-nucleotide repeat units (each accounting for ~ 43 % of newly developed SSR markers). In contrast, five out of eight polymorphic SSRs developed by Tucker *et al.* (2015) consisted of tri-nucleotide repeat units. It is well known that bacteria do not tolerate long SSR motifs and this may be the reason for difficult preparation of libraries enriched for long tri-nucleotide SSRs as observed by Wang *et al.* (2014). In contrast to this finding, there is no evidence that NGS sequencing libraries may be affected by size of SSR motifs (Tucker *et al.*, 2015).

Screening with primer pairs developed for polymorphisms identified in *Bgh* isolates from 2012 (Komínková *et al.*, 2016 submitted) revealed that they can be applied on a set of isolates captured up to three years later. Moreover, new SSR alleles and SNPs were identified in recent *Bgh* isolates. It will be thus interesting to examine the longevity of markers designed previously as well as occurrence of new polymorphisms in the target sequences. Also, it would be interesting to examine the resolution power of markers designed specifically to capture variability in the Czech *Bgh* population when applied to isolates from other countries. For effective mapping of *Bgh* spore migration and epidemic predictions, this pathogen has to be studied internationally and development of a universal genotyping panel is thus necessary.

The previously developed genotyping panel was comprised of 9 primer pairs targeting 32 polymorphisms. Marker discovery strategy utilizing whole-genome sequences of ten Czech *Bgh* isolates enabled its enlargement with additional 15 more primer pairs: 1 RJM, 14 SSRs and 3 primer pairs for amplicons containing SNPs, contributing 174 more polymorphisms. However, in this work, 232 *Bgh* isolates and an outgroup isolate *Bgtm258* were screened for the most reproducible 158 polymorphisms: 141 SSR alleles and 17 SNPs. In the collection of 232 *Bgh* isolates, 226 unique genotypes were identified. High degree of *Bgh* diversity is in agreement with findings of Dreiseitl (2015) who described 495 unique pathotypes in a collection of 521 *Bgh* isolates captured between 2011 and 2014.

A panel capable of unambiguous identification of 97.4 % of *Bgh* isolates (Fig. 4) is a significant improvement as compared to 73.2 % of unique *Bgh* isolates with the original panel of molecular markers. The six unresolved isolates present in the examined *Bgh* collection formed three pairs identical in their composition of polymorphisms: A1-5_2014 and A1-

12_2014, HE4_2015 and HE5_2015, HE11_2015 and HE12_2015 (Fig. 4). The phenotypic analysis however revealed that each isolate possessed a unique set of virulences (unpublished data). Since irresolvable pairs of *Bgh* isolates originated in the same geographical locations, it is hypothesized that the resolution power of DNA markers might be limited in some cases, when isolates originate in a rather limited geographical region.

The *Bgh* isolates originated from 17 and 15 geographical locations in 2014 and 2015, respectively. Nonetheless, the obtained cladogram rejects any correlation between genotypes and geographical origin. O'Hara and Brown (1997 and 1998) hypothesized that powdery mildew epidemics start at a large number of infection sites but they soon overlap and merge, leading to aggregation of clones and in a large scale to random pathotype patterns in field conditions. This agrees with population structure revealed in *Bgh* population in the Czech Republic in 2012 (Komínková *et al.*, 2016 submitted) as well as in Denmark (Damgaard and Giese, 1996), China (Zhu *et al.*, 2010), Morocco (Jensen *et al.*, 2013) and most recently in Australia (Tucker *et al.*, 2015). Taken together, lack of clear spatial distribution within individual countries fits into the context of previous studies that concluded that in Europe, there is one large and widespread pool of variation with existing potential for interactions of *Bgh* field populations across the whole continent (reviewed by Wolfe and McDermott, 1994).

It is believed that *Bgh* survives in a form of sexually produced chasmothecia and one would thus expect to be able to identify (a) *Bgh* isolate(s) that overwinter between two subsequent season. However, no two *Bgh* isolates were alike between the two seasons. Sampling at the beginning of the epidemic does not capture genetic variability that may arise later in the season. Possible sources of changes in population structure might be recombination (sexual reproduction in midsummer) or new selection pressure introduced by winter barley varieties grown towards the end of the season (Brown, 1994). Overwintering *Bgh* isolates would then differ in their genotype from isolates widespread at the beginning of the summer epidemic.

Migration of *Bgh* spores also contributes to population diversity. Once introduced to a new location, mutations or recombination with local genotypes may cause further diversification (Brown and Hovmøller, 2002). This factor might be of particular importance for regions in the heart of Europe, exposed to airborne spores from neighboring as well as more distant countries. In fact, pathotype immigration from Germany and Poland has been described (Dreiseitl, 2003). Another documented *Bgh* migration route connects the Czech Republic and Great Britain (Brown and Hovmøller, 2002). Finally, the possibility that *Bgh* isolates identical between the two seasons were simply not included in both collections must

be kept in mind because the employed sampling method preferably captures predominating pathotypes/genotypes.

Given the obligatory biotrophic life style of *Bgh*, DNA markers represent a powerful tool for future inquiries about *Bgh* diversity in the Czech Republic, Europe and possibly worldwide. Diversity studies carried out for longer periods of time are required for detailed understanding of population structure, dynamics and factors contributing to it. Host population structure, environmental conditions and human interventions, such as fungicides used, introduce selection pressure on *Bgh* populations (Wolfe and McDermott, 1994; Zhu *et al.*, 2010). Furthermore, migration from Europe to North Africa as well as general west to east migration over the European continent was documented (Jensen *et al.*, 2013). Taken together, complex studies considering factors contributing to *Bgh* diversity and large scale migration monitoring are essential for epidemic prediction and ultimately effective control of powdery mildew disease. Since *Bgh* population is affected by population of its hosts, environmental conditions and studies taking all these factors into consideration are desirable.

7 Conclusions

The present thesis aimed to develop a panel of DNA markers capable of unambiguous identification of *Blumeria graminis* f.sp. *hordei* isolates. The markers were used to analyze a population structure of this high-risk and agriculturally important fungal pathogen. The following conclusions were drawn:

- In whole-genome sequences of ten selected *Bgh* isolates, 84,471 candidate ISBP, 65,535 candidate SNP and 9,068 candidate SSR markers were identified, confirming the high abundance of transposable elements in the *Bgh* genome.

- Design of primer pairs for SNPs exhibited the highest success rate in comparison to development of RJMs and SSRs.

- The success rate of SSR markers was significantly higher than previously reported for *Bgh*. Moreover, SSR markers used in this work yielded a substantial number of polymorphic alleles per primer pair.

- 1 RJM, 14 SSR markers and 3 primer pairs for Sanger sequencing (enabling scoring of 8 SNPs in total) were added to the panel of DNA markers specific for *Bgh*.

- A collection of 232 *Bgh* isolates was screened for 158 polymorphisms (141 SSR alleles and 17 SNPs), leading to unambiguous identifications of 97.4 % of *Bgh* isolates in the collection. DNA markers, specifically SSRs and SNPs, have the potential to unambiguously identify number of *Bgh* isolates originating in a geographically limited region.

- Based on DNA polymorphisms, high diversity of *Bgh* populations in two subsequent seasons was unraveled and substantial inter-seasonal variability was suggested.

- *Bgh* isolates originated at 17 and 15 geographical locations in 2014 and 2015, respectively but no correlation between genotype and geographical origin was revealed.

- In combination with phenotype data, DNA markers can contribute to decipher *Bgh* population structure, dynamics and sources of diversity in *Bgh* populations, including spore migration.

8 References

Agrama, H.A. and Tuinstra, M.R. (2003) Phylogenetic diversity and relationships among sorghum accessions using SSRs and RAPDs. *African Journal of Biotechnology* **2**, 334–340.

Akkaya, M.S., Bhagwat, A.A. and Cregan, P.B. (1992) Length Polymorphisms of Simple Sequence Repeat DNA in Soybean. *Genetics* **132**, 1131–1139.

Álvarez, I. and Wendel, J.F. (2003) Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* **29**, 417–434.

Amselem, J., Vigouroux, M., Oberhaensli, S., et al. (2015) Evolution of the EKA family of powdery mildew avirulence-effector genes from the ORF 1 of a LINE retrotransposon. *BMC Genomics* **16**.

Batley, J., Barker, G., O'Sullivan, H., Edwards, K.J. and Edwards, D. (2003) Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. *Plant Pathology* **132**, 84–91.

Bennetzen, J.L. (2000) Transposable element contributions to plant gene and genome evolution. *Plant Molecular Biology* **42**, 251–269.

Bindschedler, L. V, Burgis, T. a, Mills, D.J.S., Ho, J.T.C., Cramer, R. and Spanu, P.D. (2009) In planta proteomics and proteogenomics of the biotrophic barley fungal pathogen Blumeria graminis f. sp. hordei. *Molecular & Cellular Proteomics* **8**, 2368–2381.

Bindschedler, L. V, McGuffin, L.J., Burgis, T. a, Spanu, P.D. and Cramer, R. (2011) Proteogenomics and in silico structural and functional annotation of the barley powdery mildew Blumeria graminis f.sp. hordei. *Methods* **54**, 432–441.

BluGen - the Blumeria Sequencing Project: www.blugen.org. Accessed on 27 April, 2016.

Both, M., Csukai, M., Stumpf, M.P.H. and Spanu, P.D. (2005) Gene expression profiles of blumeria graminis indicate dynamic changes to primary metabolism during development of an obligate biotrophic pathogen. *The Plant Cell* **17**, 2107–2122.

Brown, J.K.M. (1994) Chance and selection in the evolution of barlev mildew. *Trend in Microbiology* **2**, 470–475.

Brown, J.K.M. and Hovmøller, M.S. (2002) Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science* **297**, 537–541.

Bushnell, W.R. and Allen, P.J. (1962) Induction of disease symptoms in barley by powdery mildew. *Plant Physiology* **37**, 50–59.

Caffier, V., Brädle, U.E. and Wolfe, M.S. (1999) Genotypic diversity in barley powdery mildew populations in northern France. *Plant Pathology* **48**, 582–587.

Casa, A.M., Brouwer, C., Nagel, A., Wang, L., Zhang, Q., Kresovich, S. and Wessler, S.R. (2000) The MITE family Heartbreaker (Hbr): Molecular markers in maize. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 10083–10089.

Catanzariti, A.-M., Dodds, P.N. and Ellis, J.G. (2007) Avirulence proteins from haustoria-forming pathogens. *FEMS Microbiology Letters* **269**, 181–188.

Cavalli-Sforza, L.L. (1998) The DNA revolution in population genetics. *Trends in Genetics* **14**, 60–65.

Damgaard, C. and Giese, H. (1996) Genetic variation in Danish populations of Erysiphe graminis f.sp. hordei: estimation of gene diversity and effective population size using RFLP data. *Plant Pathology* **45**, 691–696.

Dean, R., Kan, J.A.L. van, Pretorius, Z.A., et al. (2012) The Top 10 fungal pathogens in molecular plant pathology. *Molecular Plant Pathology* **13**, 414–430.

Delmotte, F., Bucheli, E. and Shykoff, J.A. (1999) Host and parasite population structure in a natural plant-pathogen system. *Heredity* **82**, 300–308.

DePristo, M.A., Banks, E., Poplin, R.E., et al. (2011) A framework for variant discovery and genotyping ussing next-generation DNA sequencing data. *Nature Genetics* **43**, 589–595.

Devos, K.M., Ma, J., Pontaroli, A.C., Pratt, L.H. and Bennetzen, J.L. (2005) Analysis and mapping of randomly chosen bacterial artificial chromosome clones from hexaploid bread wheat. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 19243–19248.

Dreiseitl, A. (2003) Adaptation of Blumeria graminis f. sp. hordei to barley resistance genes in the Czech Republic in 1971 – 2000. *Plant, Soil and Environment* **49**, 241–248.

Dreiseitl, A. (2011) Changes in the Population of Blumeria graminis f.sp. hordei in the Czech Republic from 2009 to 2010. *Plant Protection Science* **47**, 43–51.

Dreiseitl, A. (2015) Changes in virulence frequencies and higher fitness of simple pathotypes in the Czech population of Blumeria graminis f.sp. hordei. *Plant Protection Science* **51**, 67–73.
Dreiseitl, A. (2010) Odolnost odrůd a její využití k snížení škodlivosti padlí ječmene 1st ed., Kroměříž: Agrotest fyto, s.r.o.
Dreiseitl, A. (2005) Severity of Powdery Mildew on Winter Barley in the Czech Republic in 1976–2005. *Plant Protection Science* **43**, 77–85.
Dreiseitl, A., Dinoor, A. and Kosman, E. (2006) Virulence and Diversity of Blumeria graminis f.sp. hordei in Israel and in the Czech Republic. *Plant Disease* **90**, 1031–1038.
Dreiseitl, A., Fowler, R.A. and Platz, G.J. (2013) Pathogenicity of Blumeria graminis f. sp. hordei in Australia in 2010 and 2011. *Plant Protection Science* **42**, 713–721.
Dreiseitl, A. and Pickering, R.A. (1999) Pathogenicity of Blumeria graminis f. sp. hordei in New Zealand in 1997. *New Zealand Journal of Crop and Horticultural Science* **27**, 273–280.
Dreiseitl, A. and Steffenson, B.J. (2000) Structures of barley mildew populations in the Czech Republic and North Dakota in 1995. *Cereal Rusts and Powdery Mildews Bulletin [http://www.crpmb.org/2000/0608dreiseitl/]*.
Dreiseitl, A. and Wang, J. (2007) Virulence and diversity of Blumeria graminis f.sp. hordei in East China. *European Journal of Plant Pathology* **117**, 357–368.
Dutech, C., Enjalbert, J., Fournier, E., Delmotte, F., Barrès, B., Carlier, J., Tharreau, D. and Giraud, T. (2007) Challenges of microsatellite isolation in fungi. *Fungal Genetics and Biology* **44**, 933–949.
Estoup, A., Tailliez, C., Cournet, J.-M. and Solignac, M. (1995) Size homoplasy and mutational processes of interrupted microsatellites in two bee species, Apis mellifera and Bombus terrestris (Apidae). *Molecular Biology and Evolution* **12**, 1074–1084.
Etter, P.D., Bassham, S., Hohenlohe, P.A., Johnson, E.A. and Cresko, W.A. (2011) SNP Discovery and genotyping for Evolutionary Genetics using RAD Sequencing Orgogozo, V. and Rockman, M. V., eds. *Methods in Molecular Biology* **772**, 157–178.
Fazekas, A., Steeves, R. and Newmaster, S. (2010) Improving sequencing quality from PCR products containing long mononucleotide repeats. *BioTechniques* **48**, 277–285.
Feechan, A., Jermakow, A.M., Torregrosa, L., Panstruga, R. and Dry, I.B. (2008) Identification of grapevine MLO gene candidates involved in susceptibility to powdery mildew. *Functional Plant Biology* **35**, 1255–1266.
Felsenstein, J. (2005) PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle. *Cladistics* **5**.
Feschotte, C., Jiang, N. and Wessler, S.R. (2002) Plant transposable elements: where genetics meets genomics. *Nature Reviews Genetics* **3**, 329–341.
Filippi, C. V, Aguirre, N., Rivas, J.G., et al. (2015) Population structure and genetic diversity characterization of a sunflower association mapping population using SSR and SNP markers. *Fungal Genetics and Biology* **15**.
Flor, H.H. (1971) Current status of the gene-for-gene concept. *Annual Review of Phytopathology* **9**, 275–296.
Giese, H., Hippe-Sanwald, S., Somerville, S. and Waller, J. (1997) Erysiphe graminis. In Plant Relationships. (Deising, H.B., ed), pp. 55–79. Berlin Heidelberg: Springer-Verlag.
Giese, H., Justesen, A.F., Pedersen, C. and Christiansen, S.K. (1999) Different nuclear elements and avirulence genes have evolved independently in the barley powdery mildew fungus, Blumeria graminis f.sp. hordei. *Molecular Plant Pathology* **on-line**.
Giraldo, M.C. and Valent, B. (2013) Filamentous plant pathogen effectors in action. *Nature Reviews Microbiology* **11**, 800–814.
Glawe, D.A. (2008) The powdery mildews: A review of the world's most familiar (yet poorly known) plant pathogens. *Annual Review of Plant Phytopathology* **46**, 27–51.
Goldstein, D.B. (2001) Islands of linkage disequilibrium. *Nature Genetics* **29**, 109–111.
Goodwin, S.B., Lee, T. a J. van der, Cavaletto, J.R., Lintel Hekkert, B. Te, Crane, C.F. and Kema, G.H.J. (2007) Identification and genetic mapping of highly polymorphic microsatellite loci from an EST database of the septoria tritici blotch pathogen Mycosphaerella graminicola. *Fungal Genetics and Biology* **44**, 398–414.
Hacquard, S., Kracher, B., Maekawa, T., Vernaldi, S., Schulze-Lefert, P. and Loren van Themaat, E. Ver (2013) Mosaic genome structure of the barley powdery mildew pathogen and conservation of transcriptional programs in divergent hosts. *Proceedings of the National Academy of Sciences of the United States of America* **110**, E2219–2228.

Hayward, A., Mason, A.S., Dalton-Morgan, J., Zander, M., Edwards, D. and Batley, J. (2012) SNP discovery and applications in Brassica napus. *Journal of Plant Biotechnology* **39**, 1–12.

Chakravarthi, B.K. and Naravaneni, R. (2006) SSR marker based DNA fingerprinting and diversity study in rice (Oryza sativa L.). *African Journal of Biotechnology* **5**, 684–688.

Chalupska, D., Lee, H.Y., Faris, J.D., Evrard, A., Chalhoub, B., Haselkorn, R. and Gornicki, P. (2008) Acc homoeoloci and the evolution of wheat genomes. *Procee* **105**, 9691–9696.

Chen, X. and Sullivan, P.F. (2003) Single nucleotide polymorphism genotyping: biochemistry, protocol, cost and throughput. *The Pharmacogenomics Journal* **3**, 77–96.

Chiocchetti, A., Bernardo, I., Daboussi, M., et al. (1999) Detection of Fusarium oxysporum f.sp. dianthi in carnation tissue by PCR amplification of transposon insertions. *Phytopathology* **89**, 1169–1175.

Inuma, T., Khodaparast, S.A. and Takamatsu, S. (2007) Multilocus phylogenetic analyses within Blumeria graminis, a powdery mildew fungus of cereals. *Molecular Phylogenetics and Evolution* **44**, 741–751.

Jensen, H.R., Dreiseitl, A., Sadiki, M. and Schoen, D.J. (2013) High diversity, low spatial structure and rapid pathotype evolution in Moroccan populations of Blumeria graminis f.sp. hordei. *European Journal of Plant Pathology* **136**, 323–336.

Jørgensen, J.H. (2013) Erysiphe graminis, powdery mildew oc cereals and grasses. In Genetics of Plant Pathogenic Fungi. (Sidhu, G.S., ed), pp. 137–157. Burlington: Elsevier Science.

Kalendar, R., Flavell, A.J., Ellis, T.H.N., Sjakste, T., Moisy, C. and Schulman, A.H. (2011) Analysis of plant diversity with retrotransposon-based molecular markers. *Heredity* **106**, 520–530.

Kalendar, R. and Schulman, A.H. (2006) IRAP and REMAP for retrotransposon-based genotyping and fingerprinting. *Nature Protocols* **1**, 2478 – 2484.

Kemen, E. and Jones, J.D.G. (2012) Obligate biotroph parasitism: can we link genomes to lifestyles? *Trends in Plant Science* **17**, 448–457.

Koch, G. and Köhler, W. (1990) Isozyme Variation and Genetic Distances of Erysiphe graminis DC. Formae Speciales. *Journal of Phytopathology* **129**, 89–101.

Koch, G. and Köhler, W. (1991) Isozyme Variation and Genetic Diversity of the European Barley Powdery Mildew Population. *Journal of Phytopathology* **131**, 333–344.

Koichiro, T., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* **30**, 2725–2729.

Kokina, I. and Rashal, I. (2012) Results of monitoring of the population of Blumeria graminis f.sp. hordei in Latvia in 2009-2010. *Proceedings of the Latvian Academy of Sciences. Section B. Natural, Exact, and Applied Sciences.* **66**, 41–47.

Komínková, E. (2013) *Molekulární charakterizace patogenu Blumeria graminis*. Palacký University in Olomouc.

Komínková, E. (2011) *Padlí travní a studium jeho diverzity v rámci České republiky*. Palacký University in Olomouc.

Kumar, S., Banks, T.W. and Cloutier, S. (2012) SNP Discovery through Next-Generation Sequencing and Its Applications. *International Journal of Plant Genomics* **2012**.

Kunoh, M., Itoh, O., Kohno, M. and Ishizaki, H. (1979) Are primary germ tubes of conidia unique to Erysiphe graminis? *Annals of the Phytopathological Society of Japan* **45**, 675–682.

Kusch, S., Ahmadinejad, N., Panstruga, R. and Kuhn, H. (2014) In silico analysis of the core signaling proteome from the barley powdery mildew pathogen (Blumeria graminis f. sp. hordei). *BM* **15**.

Kůst, F. and Potměšilová, J. (2014) *Situační a výhledová zpráva: Obiloviny*, Praha.

Kwok, P. and Chen, X. (2003) Detection of Single Nucleotide Polymorphisms. *Current Issues in Molecular Biology* **5**, 43–60.

Leff, B., Ramankutty, N. and Foley, J. a. (2004) Geographic distribution of major crops across the world. *Global Biogeochemical Cycles* **18**, 1–27.

Li, H. and Durbin, R. (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595.

Li, R., Li, Y., Fang, X., Yang, H., Wang, J., Kristiansen, K. and Wang, J. (2009) SNP detection for massively parallel whole-genome resequencing. *Genome Research* **19**, 1124–1132.

Mazaheri, M., Kianian, P.M.A., Mergoum, M., et al. (2014) Transposable Element Junctions in Marker Development and Genomic Characterization of Barley. *The Plant Genome* **7**, 1–35.

McDermott, J.M., Brädle, U., Dutly, F., Haemmerli, U.A., Keller, S., Müller, K.E. and Wolfe, M.S. (1994) Genetic Variation in Powdery Mildew of Barley: Development of RAPD, SCAR, and VNTR Markers. *Phytopathology* **84**, 1316–1321.

McDonald, B.A. and Linde, C. (2002) Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Plant Phytopathology* **40**, 349–79.

Mcdonald, B.A. and Mcdermott, J.M. (1993) Population Genetics of Plant Pathogenic Fungi. *BioScience* **43**, 311–319.

Mori, Y., Sato, Y. and Takamatsu, S. (2000) Evolutionary analysis of powdery mildew fungi using nucleotide sequences of the nuclear ribosomal DNA. *Mycologia* **92**, 74–93.

Mori, Y., Sato, Y. and Takamatsu, S. (2015) Evolutionary analysis of the powdery mildew fungi using nucleotide sequences of the nuclear ribosomal DNA. *Mycologia* **92**, 74–93.

Morin, P.A., Luikart, G. and Wayne, R.K. (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution* **19**, 208–216.

Nielsen, K.A., Nicholson, R.L., Carver, T.L., Kunoh, H. and Oliver, R.P. (2000) First touch: An immediate response to surface recognition in conidia of Blumeria graminis. *Physiological and Molecular Plant Pathology* **56**, 63–70.

Nowara, D., Gay, A., Lacomme, C., Shaw, J., Ridout, C., Douchkov, D., Hensel, G., Kumlehn, J. and Schweizer, P. (2010) HIGS: host-induced gene silencing in the obligate biotrophic fungal pathogen Blumeria graminis. *The Plant Cell* **22**, 3130–3141.

Nybom, H., Weising, K. and Rotter, B. (2014) DNA fingerprinting in botany: past, present, future. *Investigative genetics* **5**, 1–35.

O'Dell, M., Wolfe, M.S., Flavell, R.B., Simpson, C.G. and Summers, R.W. (1989) Molecular variation in populations of Erysiphe graminis on barley, oats and rye. *Plant Pathology* **38**, 340–351.

O'Hara, R.B. and Brown, J.K.M. (1998) Movement of barley powdery mildew within field plots. *Plant Pathology* **47**, 394–400.

O'Hara, R.B. and Brown, J.K.M. (1997) Spatial aggregation of pathotypes of barley powdery mildew. *Plant Pathology* **46**, 969–977.

Oberhaensli, S., Parlange, F., Buchmann, J.P., Jenny, F.H., Abbott, J.C., Burgis, T. a, Spanu, P.D., Keller, B. and Wicker, T. (2011) Comparative sequence analysis of wheat and barley powdery mildew fungi reveals gene colinearity, dates divergence and indicates host-pathogen co-evolution. *Fungal Genetics and Biology* **48**, 327–334.

Opalski, K.S., Schultheiss, H., Kogel, K.-H. and Hückelhoven, R. (2005) The receptor-like MLO protein and the RAC/ROP family G-protein RACB modulate actin reorganization in barley attacked by the biotrophic powdery mildew fungus Blumeria graminis f.sp. hordei. *The Plant Journal* **41**, 291–303.

Panstruga, R. and Dodds, P.N. (2009) Terrific protein traffic: The mystery of effector protein delivery by filamentous plant pathogens. *Science* **324**, 748–750.

Pariaud, B., Ravigné, V., Halkett, F., Goyeau, H., Carlier, J. and Lannou, C. (2009) Aggressiveness and its role in the adaptation of plant pathogens. *Plant Pathology* **58**, 409–424.

Parlange, F., Oberhaensli, S., Breen, J., Platzer, M., Taudien, S., Šimková, H., Wicker, T., Doležel, J. and Keller, B. (2011) A major invasion of transposable elements accounts for the large size of the Blumeria graminis f.sp. tritici genome. *Functional & Integrative Genomics* **11**, 671–677.

Paux, E., Faure, S., Choulet, F., et al. (2010) Insertion site-based polymorphism markers open new perspectives for genome saturation and marker-assisted selection in wheat. *Plant Biotechnology Journal* **8**, 196–210.

Pedersen, C., Rasmussen, S.W. and Giese, H. (2002) A genetic map of Blumeria graminis based on functional genes, avirulence genes, and molecular markers. *Fungal Genetics and Biology* **35**, 235–246.

Pedersen, C., Ver, E., Themaat, L. Van, et al. (2012) Structure and evolution of barley powdery mildew effector candidates. *BMC Genomics* **13**.

Pessina, S., Pavan, S., Catalano, D., Gallotta, A., Visser, R.G.F. and Bai, Y. (2014) Characterization of the MLO gene family in Rosaceae and gene expression analysis in Malus domestica. *BMC Genomics* **15**, 618.

Powell, W., Machray, G.C. and Provar, J. (1996) Polymorphism revealed by simple sequence repeats. *Trends in Plant Science* **1**, 215–222.

Pryce-Jones, E., Carver, T. and Gurr, S.J. (1999) The roles of cellulase enzymes and mechanical force in host penetration by Erysiphe graminis f. sp. hordei. *Physiological and Molecular Plant Pathology* **55**, 175–182.

Rafalski, A. (2002) Applications of single nucleotide polymorphisms in crop genetics. *Current Opinion in Plant Biology* **5**, 94–100.

Raffaele, S. and Kamoun, S. (2012) Genome evolution in filamentous plant pathogens: why bigger can be better. *Nature Reviews Microbiology* **10**, 417–430.

Rambaut, A. FigTree v 1.4.2.: http://tree.bio.ed.ac.uk/software/figtree/. Accessed on April 27, 2016.
Ridout, C.J. (2009) Profiles in Pathogenesis and Mutualism: Powdery Mildews. In Plant Relationships. (Deisig, H.B., ed), pp. 51–68. Berlin Heidelberg: Springer-Verlag.
Ridout, C.J., Skamnioti, P., Porritt, O., Sacristan, S., Jones, J.D.G. and Brown, J.K.M. (2006) Multiple avirulence paralogues in cereal powdery mildew fungi may contribute to parasite fitness and defeat of plant resistance. *The Plant Cell* **18**, 2402–2414.
Riju, A. and Arunachalam, V. (2010) Electronic Sorting of SNP/Indel Sites in Expressed Sequence Tag Libraries of Cocoa (Theobroma cacao L.). *Genes, Genomes and Genomics* **4**, 37–40.
Robinson, A.J., Love, C.G., Batley, J., Barker, G. and Edwards, D. (2004) Simple sequence repeat marker loci discovery using SSR primer. *Bioinformatics* **20**, 1475–1476.
Russell, J.N. and Mikulasova, J. (2015) *Grain and Feed Annual 2015*, Prague.
Sacristán, S. and García-Arenal, F. (2008) The evolution of virulence and pathogenicity in plant pathogen populations. *Molecular Plant Pathology* **9**, 369–384.
Sacristán, S., Vigouroux, M., Pedersen, C., Skamnioti, P., Thordal-Christensen, H., Micali, C., Brown, J.K.M. and Ridout, C.J. (2009) Coevolution between a family of parasite virulence effectors and a class of LINE-1 retrotransposons. *PloS One* **4**, e7463.
Seeholzer, S., Tsuchimatsu, T., Jordan, T., et al. (2010) Diversity at the Mla powdery mildew resistance locus from cultivated barley reveals sites of positive selection. *Molecular Plant-Microbe Interactions* **23**, 497–509.
Seko, Y., Bolay, A., Kiss, L., Heluta, V., Grigaliunaite, B. and Takamatsu, S. (2008) Molecular evidence in support of recent migration of a powdery mildew fungus on Syringa spp. into Europe from East Asia. *Plant Pathology* **57**, 243–250.
Shattuck-Eidens, D.M., Bell, R.N., Neuhausen, S.L. and Helentjaris, T. (1990) DNA Sequence Variation Within Maize and Melon: Observations From Polymerase Chain Reaction Amplification and Direct Sequencing. *Genetics* **126**, 207–217.
Schlötterer, C. (2004) The evolution of molecular markers — just a matter of fashion? **5**, 3–9.
Schmidt, S.M., Kuhn, H., Micali, C., Liller, C., Kwaaitaal, M. and Panstruga, R. (2014) Interaction of a Blumeria graminis f. sp. hordei effector candidate with a barley ARF-GAP suggests that host vesicle trafficking is a fungal pathogenicity target. *Molecular Plant Pathology* **15**, 535–549.
Schulze-Lefert, P. and Panstruga, R. (2011) A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. *Trends in Plant Science* **16**, 117–125.
Schwarzbach, E. (1979) No TitleA High Throughput Jet Trap for Collecting Mildew Spores on Living Leaves. *Journal of Phytopathology* **94**, 165–171.
Schweizer, P. (2007) Nonhost resistance of plants to powdery mildew—New opportunities to unravel the mystery. *Physiological and Molecular Plant Pathology* **70**, 3–7.
Sillero, J.C., Fondevilla, S., Davidson, J., Patto, M.C.V., Warkentin, T.D., Thomas, J. and Rubiales, D. (2006) Screening techniques and sources of resistance to rusts and mildews in grain legumes. *Euphytica* **147**, 255–272.
Skamnioti, P., Pedersen, C., Al-Chaarani, G.R., Holefors, A., Thordal-Christensen, H., Brown, J.K.M. and Ridout, C.J. (2008) Genetics of avirulence genes in Blumeria graminis f.sp. hordei and physical mapping of AVR(a22) and AVR(a12). *Fungal Genetics and Biology* **45**, 243–252.
Spanu, P.D. (2012) The genomics of obligate (and nonobligate) biotrophs. *Annual Review of Phytopathology* **50**, 741–751.
Spanu, P.D., Abbott, J.C., Amselem, J., et al. (2010) Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science* **330**, 1543–1546.
Stukenbrock, E.H., Banke, S., Zala, M., Mcdonald, B. a. and Oliver, R.P. (2005) Isolation and characterization of EST-derived microsatellite loci from the fungal wheat pathogen Phaeosphaeria nodorum. *Molecular Ecology Notes* **5**, 931–933.
Sunnucks, P. (2000) Efficient genetic markers for population biology. *Trends in Ecology & Evolution* **15**, 199–203.
Takamatsu, S. (2013a) Molecular phylogeny reveals phenotypic evolution of powdery mildews (Erysiphales, Ascomycota). *Journal of General Plant Pathology* **79**, 218–226.
Takamatsu, S. (2013b) Origin and evolution of the powdery mildews (Ascomycota, Erysiphales). *Mycoscience* **54**, 75–86.
Takamatsu, S. (2004) Phylogeny and evolution of the powdery mildew fungi (Erysiphales , Ascomycota) inferred from nuclear ribosomal DNA sequences. *Mycoscience* **45**, 147–157.

Taylor, J.W., Jacobson, D.J. and Fisher, M.C. (1999) The evolution of asexual fungi: Reproduction, speciation and classification. *Annual Review of Phytopathology* **37**, 197–246.

Thomas, S.W., Rasmussen, S.W., Glaring, M.A., Rouster, J.A., Christiansen, S.K. and Oliver, R.P. (2001) Gene identification in the obligate fungal pathogen Blumeria graminis by expressed sequence tag analysis. *Fungal Genetics and Biology* **33**, 195–211.

Tratwal, A. and Bocianowski, J. (2014) Blumeria graminis f. sp. hordei virulence frequency and the powdery mildew incidence on spring barley in the Wielkopolska province. *Journal of Plant Protection Research* **54**, 28–35.

Troch, V., Audenaert, K., Wyand, R. a, Haesaert, G., Höfte, M. and Brown, J.K.M. (2014) Formae speciales of cereal powdery mildew: close or distant relatives? *Molecular Plant Pathology* **15**, 304–314.

Tucker, M.A., Moffat, C.S., Ellwood, S.R., Tan, K.C., Jayasena, K. and Oliver, R.P. (2015) Development of genetic SSR markers in Blumeria graminis f. sp. hordei and application to isolates from Australia. *Plant Pathology* **64**, 337–343.

Untergrass, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. and Rozen, S.G. (2012) Primer3 - new capabilities and interfaces. *Nucleic Acid Research* **40**, e115.

Usami, T., Ishigaki, S., Takashina, H., Matsubara, Y. and Amemiya, Y. (2007) Cloning of DNA fragments specific to the pathotype and race of Verticillium dahliae. *Journal of General Plant Pathology* **73**, 89–95.

Varshney, R.K., Graner, A. and Sorrells, M.E. (2005) Genic microsatellite markers in plants: features and applications. *Trends in Biotechnology* **23**, 48–55.

Varshney, R.K., Marcel, T.C., Ramsay, L., et al. (2007) A high density barley microsatellite consensus map with 775 SSR loci. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* **114**, 1091–1103.

Wang, B., Tan, H.-W., Fang, W., Meinhardt, L.W., Mischke, S., Matsumoto, T. and Zhang, D. (2015) Developing single nucleotide polymorphism (SNP) markers from transcriptome sequences for identification of longan (Dimocarpus longan) germplasm. *Horticulture Research* **2**.

Wang, M., Xue, F., Yang, P., Duan, X., Zhou, Y., Shen, C., Zhang, G. and Wang, B. (2014) Development of SSR Markers for a Phytopathogenic Fungus, Blumeria graminis f.sp. tritici, Using a FIASCO Protocol. *Journal of Integrative Agriculture* **13**, 100–104.

Wang, W., Wen, Y., Berkey, R. and Xiao, S. (2009) Specific targeting of the Arabidopsis resistance protein RPW8.2 to the interfacial membrane encasing the fungal Haustorium renders broad-spectrum resistance to powdery mildew. *The Plant Cell* **21**, 2898–2913.

Wanjugi, H., Coleman-derr, D., Huo, N., Kianian, S.F., Luo, M., Wu, J., Anderson, O. and Gu, Y.Q. (2009) Rapid development of PCR-based genome- specific repetitive DNA junction markers in wheat. *Genome* **52**, 576–587.

Weising, K., Nybom, H., Pfenninger, M., Wolff, K. and Kahl, G. (2005) DNA Fingerprinting in Plants: Principles, Methods and Applications 2nd ed., Boca Raton, FL, USA: CRC Press.

Wicker, T. TREP: Transposable Element Platform: http://botserv2.uzh.ch/home/wicker/TREP/TREP_homepage.php

Wicker, T., Oberhaensli, S., Parlange, F., et al. (2013) The wheat powdery mildew genome shows the unique evolution of an obligate biotroph. *Nature Genetics* **45**, 1092–1096.

Wicker, T., Sabot, F., Hua-van, A., et al. (2007) A unified classification system for ekaryotic transposable elements. *Nature Reviews Genetics* **8**, 973–982.

Wolfe, M.S. and McDermott, J.M. (1994) Population genetics of plant pathogen interactions: The example of the Erysiphe graminis-Hordeum vulgare pathosystem. *Annual Review of Phytopathology* **32**, 89–113.

Wright, A.J., Carver, T.L.W., Thomas, B.J., Fenwick, N.I.D., Kunoh, H. and Nicholson, R.L. (2000) The rapid and accurate determination of germ tube emergence site by Blumeria graminis conidia. *Physiological and Molecular Plant Pathology* **57**, 281–301.

Wyand, R.A. and Brown, J.K.M. (2003) Genetic and forma specialis diversity in Blumeria graminis of cereals and its implications for host-pathogen co-evolution. *Molecular Plant Pathology* **4**, 187–198.

Yadav, C.B., Bonthala, V.S., Muthamilarasan, M., Pandey, G., Khan, Y. and Prasad, M. (2015) Genome-wide development of transposable elements-based markers in foxtail millet and construction of an integrated database. *DNA Research* **22**, 79–90.

You, F.M., Wanjugi, H., Huo, N., Lazo, G.R., Luo, M.-C., Anderson, O.D., Dvorak, J. and Gu, Y.Q. (2010) RJPrimers: unique transposable element insertion junction discovery and PCR primer design for marker development. *Nucleic Acid Research* **38**, W313–W320.

Yuan, M., Gong, L., Meng, R., Li, S., Dang, P., Guo, B. and He, G. (2010) Development of trinucleotide (GGC)n SSR markers in peanut (Arachis hypogaea L.). *Electronic Journal of Biotechnology* **13**.

Zhang, W., Pedersen, C., Kwaaitaal, M., et al. (2012) Interaction of barley powdery mildew effector candidate CSEP0055 with the defence protein PR17c. *Molecular Plant Pathology* **13**, 1110–1119.

Zhang, Z., Henderson, C., Perfect, E., Carver, T.L.W., Thomas, B.J., Skamnioti, P. and Gurr, S.J. (2005) Of genes and genomes , needles and haystacks: Blumeria graminis and functionality. *Molecular Plant Pathology* **6**, 561–575.

Zhu, J., Wang, J., Jia, Q., Yang, J., Zhou, Y., Lin, F., Hua, W. and Shang, Y. (2010) Pathotypes and Genetic Diversity of Blumeria graminis f. sp. hordei in the Winter Barley Regions in China. *Agricultural Sciences in China* **9**, 1787–1798.

Zimin, A. V, Marçais, G., Puiu, D., Roberts, M., Salzberg, S.L. and Yorke, J.A. (2013) The MaSuRCA genome assembler. *Bioinformatics* **29**, 2669–2677.

9 Abbreviations

AFLP	amplified fragment length polymorhism
APS	ammonium persulfate
ATM	amplicon tagment mix (Nextera XT DNA Library Prep)
BEC	*Blumeria* effector candidate
Bgh	*Blumeria graminis* f.sp. *hordei*
Bgt	*Blumeria graminis* f.sp. *tritici*
CASP	cleaved amplified polymorhic sequence
CSEP	candidate secreted effector proteins
ff. spp.	*formae speciales*
f. sp.	*forma specialis*
GPD	glyceraldehyde-3-phosphate dehydrogenase
EDTA	ethylenediaminetetraacetic acid disodium salt dihydrate
EKA	effectors with similarity to avirulence proteins AVR_{k1} and AVR_{a10}
EST	expressed sequence tag
HSP	heat-shock protein
ISBP	insertion site-based polymorphism
LINE	long interspersed element
LTR	long terminal repeat
MITE	miniature inverted–repeat transposable element
NGS	next-generation sequencing
ORF	open reading frame
PAV	presence absence variation
RAPD	random amplified polymorphic DNA
RFLP	restriction fragment length polymorphism
RJM	repeat junction marker
SCAR	sequence characterized amplified region
SINE	short interspersed element
SNP	single nucleotide polymorphism
SSR	simple sequence repeat
TE	transposable element
TEMED	N,N,N',N'-tetramethylethylenediamine
VNTR	variable number of tandem repeats