Faculty of Science

Palacký University Olomouc

Department of Cell Biology and Genetics

&

Institute of Experimental Botany of the Czech Academy of Sciences

Centre of the Region Haná for Biotechnological and Agricultural Research

Eva Janáková

Genetics of host-pathogen interactions in the cereals-powdery mildew fungus (*Blumeria graminis*) pathosystem

Dissertation

Supervisor: Mgr. Miroslav Valárik, Ph.D.

Olomouc 2019

I declare that this dissertation has been composed solely by myself and has not been submitted for any other degree or professional qualification. I confirm that the results presented has formed part of jointly-authored publications, I played a major role in their preparation and execution of the experiments. Due references have been provided on all supporting literatures and resources.

Olomouc,

…….

Mgr. Eva Janáková

ACKNOWLEDGEMENTS

I am grateful to my supervisor, Mgr. Miroslav Valárik, Ph.D., for providing me with an interesting research topic to focus on during my graduate studies and for giving me opportunities to improve my qualification in the fields of genomics and bioinformatics. Moreover, I would like to express my gratitude to Prof. Ing. Jaroslav Doležel, DrSc., the head of Centre of Plant Structural and Functional Genomics, for enabling me to work in such a world-class research facility and to Doc. Ing. Antonín Dreiseitl, CSc., my external consultant, for giving me professional advice and help whenever I needed them and for promoting an international collaboration which considerably enriched my research experience. My thanks belong also to Ing. Marie Seifertová, Radka Tušková, Eva Jahnová and Helena Tvardíková for their invaluable technical assistance which allowed me to spare more time to immerse myself in the science. I would like to acknowledge my other colleagues as well and last but not least, I highly appreciate all the support I obtained from my family over the years.

This work benefited from interactions and support promoted by COST Action FA1208 and was funded by grant award LO1204 from the National Program of Sustainability I, by the Czech Science Foundation (award 14-07164S) and by institutional research funding IUT 193 of the Estonian Ministry of Education and Research. Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum provided under the program "Projects of Large Research, Development and Innovation Infrastructures" (CESNET LM2015042) is greatly appreciated. The Ph.D. candidate obtained a financial help to participate in short-term research fellowships from the Erasmus+ program and from the foundation "Nadání Josefa, Marie and Zdeňky Hlávkových".

BIBLIOGRAPHICAL IDENTIFICATION

Author: Mgr. Eva Janáková

Title: Genetics of host-pathogen interactions in the cereals-powdery mildew fungus (*Blumeria graminis*) pathosystem

Type of Thesis: Dissertation

Department: Department of Cell Biology and Genetics

Supervisor: Mgr. Miroslav Valárik, Ph.D.

The Year of Presentation: 2019

Abstract:

Increasing human population along with climate change represent a challenge to meet growing demand of food crop production. Cereals including wheat, the most widely cultivated crop, have a major importance in global food security. Apart from other factors, production is threatened by diverse biotic stress agents. Fungal pathogen *Blumeria graminis* causes a powdery mildew disease on cereals which can result in significant yield losses. To date, control measures have consisted predominantly in application of fungicides or deployment of major resistance (*R*) genes in grown cultivars. However, these strategies do not fully meet requirements for sustainable and durable crop protection. Although employing disease resistance is highly desirable, it is necessary to seek durable sources of resistance or to rationally deploy appropriate combinations of *R* genes.

This thesis presents a concise summary of state-of-the-art knowledge on plant-pathogen interactions with emphasis laid on the pathosystem of cereals and powdery mildew fungus. First, it contains general information about studied organisms and especially about their genomes. Second, it deals with different types of plant-pathogen interactions resulting in resistance and finally, the most important aspects of current approaches used for identification of interacting components are discussed.

As a follow-up to the provided theoretical background, two research projects are presented. The first one was conducted on host plant represented by bread wheat. Due to the limited gene pool of modern elite cultivars, introgressions from related species have been acknowledged as a rich source of genetic diversity. A powdery mildew resistance QTL had been therefore previously introgressed from tetraploid *Triticum militinae* to *T. aestivum* chromosome arm 4AL. The main goal of this project was map-based cloning of underlying gene(s) to elucidate mechanism of this partial, broad-spectrum resistance and facilitate its deployment in breeding. Additionally, the second project focusing on the pathogen genetic diversity and the repertoire of its avirulence effectors was addressed in parallel. The results of this research offer a promising perspective of pathogen-informed strategies for sustainable powdery mildew resistance in cereals. Moreover, findings acquired within both projects contribute to our knowledge on the complex problematics of plant-pathogen interactions and disease resistance.

Keywords: cereals, bread wheat, barley, powdery mildew fungus, resistance genes, avirulence, pathogen effectors, map-based cloning, genome-wide association study, introgression, diversity

Number of Pages: 86

Number of appendices: VIII

Language: English

BIBLIOGRAFICKÁ IDENTIFIKACE

Autor: Mgr. Eva Janáková

Název práce: Genetický základ interakcí hostitele a patogenu v patosystému obilovin a padlí travního (*Blumeria graminis*)

Typ práce: Disertační

Pracoviště: Katedra buněčné biologie a genetiky

Vedoucí práce: Mgr. Miroslav Valárik, Ph.D.

Rok obhajoby práce: 2019

Abstrakt:

Narůstající lidská populace spolu se změnou klimatu představují výzvu vzhledem k nutnosti uspokojit rostoucí požadavky na produkci plodin, které jsou zdrojem naší obživy. Obiloviny zahrnující i pšenici, nejrozšířenější plodinu na světě, mají pro zajištění dostatku potravin klíčový význam. Jejich produkce je ohrožena mimo jiné různými původci biotického stresu. Houbový patogen *Blumeria graminis* způsobuje chorobu obilovin známou jako „padlí“, která může vést k významnému snížení výnosu. Dosud převažující kontrolní opatření spočívají v aplikaci fungicidů a ve využití rasově specifických genů rezistence (*R* geny) v pěstovaných kultivarech. Tyto strategie však úplně nesplňují požadavky na udržitelnou a trvalou ochranu plodin. Přestože je využití rezistence vůči chorobám velmi žádoucí, je nutné hledat zdroje trvalé rezistence, případně vhodně kombinovat *R* geny.

Tato práce stručně shrnuje současné poznatky o interakcích hostitele a patogenu s důrazem kladeným na patosystém obilovin a padlí travního. V první řadě obsahuje obecné informace o studovaných organismech a zvláště o jejich genomech. Dále se práce zabývá různými typy interakcí hostitele a patogenu, které vedou k rezistenci. Kromě toho jsou zde probrány i nejdůležitější aspekty přístupů, které jsou v současnosti používány pro identifikaci interagujících komponent.

Na uvedená teoretická východiska následně navazují dva prezentované výzkumné projekty. První z nich se zabývá hostitelskou rostlinou, v tomto případě pšenicí. Vzhledem ke zúžené genetické základně moderních kultivarů jsou introgrese z příbuzných druhů cenným zdrojem genetické diverzity. Za tímto účelem byl do chromozomového ramene 4AL pšenice seté přenesen lokus pocházející z tetraploidního druhu *Triticum militinae* poskytující kvantitativní rezistenci k padlí travnímu. Hlavním cílem tohoto projektu bylo poziční klonování zodpovědného genu/genů, které umožní objasnit mechanismus této částečné, širokospektrální rezistence a usnadní její využití při šlechtění. Druhý z projektů řešený paralelně byl pak zaměřen na genetickou diverzitu patogenu a na soubor jeho efektorů avirulence. Výsledky tohoto výzkumu nabízí slibnou perspektivu udržitelné strategie rezistence k padlí travnímu obilovin založenou na znalosti populace patogenu. Kromě toho poznatky získané v rámci obou projektů přispívají k našemu porozumění složité problematiky interakcí hostitele a patogenu a odolnosti vůči chorobám.

Klíčová slova: obiloviny, pšenice, ječmen, padlí travní, geny rezistence, avirulence, efektory patogenu, poziční klonování, celogenomová asociační studie, introgrese, diverzita

Počet stran: 86

Počet příloh: VIII

Jazyk: Angličtina

CONTENTS

1 INTRODUCTION

According to the latest forecast released by the United Nations, the world's population is expected to reach 9.8 billion people by 2050. Naturally, such a prospect raises major concerns especially about global food security. Among all sources of human nutrition, cereal crops play an irreplaceable role and it is therefore absolutely crucial to maximize an effort towards keeping pace with growing demand. However, increasing production of certain crop is more complex than just increasing the yield. Throughout the world, agriculture crops are permanently threatened by a wide range of biotic and abiotic stresses which have an adverse effect on agricultural productivity. Besides the impact of climate change associated with weather extremes, plant pathogens represent a major factor resulting in potentially large yield losses. In the past century, a famous story of success known as The Green Revolution was launched by Norman Borlaug who originally pursued wheat resistant to stem rust. But finally, it resulted in dramatically improved global production of wheat and rice by changing agricultural practice together with creating high-yielding varieties. Nowadays, we are facing a similar challenge. However, we are also witnessing such a rapid progress in plant genomics that could have been hardly imagined a few decades ago. Due to the boom of next-generation sequencing technologies, obtaining a complete genomic sequence became almost a routine. A massive development in bioinformatics and computing technologies made sequence assembly feasible even in large, polyploid genomes rich in repetitive elements, such as the one of wheat, the most widely planted crop. With genomic sequences at hand, the next step consists in its deciphering to identify gene pathways involved in agronomically important traits. Nevertheless, it is crucial to realize that whenever talking about plant resistance to pathogens, we refer to a highly intricate system comprising three basic components – host, pathogen and environment. While the environmental factors are relatively easy to understand, a deeper knowledge of the pathogen including mechanisms involved in virulence and interactions with its host is of key importance. The ultimate aim consists in developing pathogen-informed strategies for sustainable and broad-spectrum resistance.

Since the area of plant-pathogen interactions is vast, it is not feasible for this dissertation to entirely cover its many different aspects. The theoretical part is therefore focused on general features of plant immunity, resistance mechanisms and genes involved in plant-pathogen interactions. In addition, different approaches available to identify these genes are presented. The experimental part is focused on pathosystem constituted by two major cereal crops in Europe, wheat and barley, and the fungal pathogen *Blumeria graminis* which causes powdery mildew disease on several cultivated cereal species and wild grasses. First, a map-based cloning has been performed to identify a gene improving broad-spectrum resistance of wheat to powdery mildew. Second, the diversity of isolates belonging to *Blumeria graminis* f. sp. *hordei* has been evaluated and these isolates were subsequently used for genome-wide association study to identify potential avirulence effectors. In the context of other studies, the results can contribute to shifting the borders of our knowledge on mechanisms involved in plant-pathogen interaction. Unveiling the whole complexity is a prerequisite to achieve effective, durable and environment-friendly crop protection.

2 LITERATURE REVIEW

2.1 Cereal crops

Cereal crops also called grain crops comprise a group of grasses of the family *Poaceae* (formerly *Gramineae*) grown for their small, edible seeds (Chapman and Carter 1976). The "true" cereals include maize, rice, wheat, barley, rye, triticale, oats, sorghum, millets and some other, not widely known species grown on a local scale. Apart from cereals *sensu stricto*, a group of dicotyledonous plants called pseudocereals has been defined based on their production of starch-rich seeds similar to cereals. This section covers minor crops such as amaranth, quinoa or buckwheat (Schoenlechner *et al.* 2008). In the concept of Food and Agriculture Organization of the United Nations (FAO), cereals have a more general meaning referring to crops harvested only for dry grain. The global importance of cereals is well-illustrated by the fact that they account for the majority of production of the crop sector of agriculture (FAO, 2015). According to the data from 2014, maize has the highest production worldwide (over one billion tonnes) which accounts for 36.8 % of total cereal production. It is followed by rice and wheat with 26.3 % and 25.9 % shares, respectively. The production of remaining cereals ranking among the "top five", barley and sorghum, rather lags behind with 5.1 % and 2.4 % shares, respectively. However, when focused on Europe, wheat is by far the most important of all cereals. It accounts for 47.2 % of total cereal production with 249 million tonnes harvested in 2014. Maize ranks second and barley third contributing 24.4 % and 17.7 %, respectively (FAO, 2017). Besides the significance of cereals as a principal component of human diet, it is worth mentioning that they played a substantial role in the onset of agriculture which gave rise to a number of civilizations (Halperin 1936; Lev-Yadun *et al.* 2000; Salamini *et al.* 2002).

Considering the specialization of this thesis, the following subchapters very briefly introduce two major cereal crops, wheat and barley, with an emphasis placed on different aspects of their genomes. In case of wheat, a short account of approaches employed to decipher its complex genome is given.

2.1.1 Wheat

Wheat is the most widely cultivated crop covering an area of 220.4 million hectares (FAO, 2017). It is grown in a wide range of environments, predominantly in temperate zones but it also tolerates warmer regions (Awika 2011). As a staple food, it provides about 20 % of total calories and proteins consumed worldwide which highlights its key position in global food security (Shiferaw *et al.* 2013). Wheat grains are used in a wide range of products, their processing and final utilization depend largely on properties such as grain hardness, protein content or dough strength. The wheat group (genera *Aegilops* and *Triticum*) includes a few dozen species which can hybridize with one another resulting in formation of new allopolyploid species with an advantage of genome plasticity associated with their great evolutionary success (Dubcovsky and Dvorak 2007; Feldman and Levy 2012). Today, world wheat production is practically dependent on only two species, both fully domesticated and free-threshing. Hexaploid bread wheat, *Triticum aestivum* L., accounts for 95 % of total production and the remaining share is made up of tetraploid durum wheat, *T. turgidum* ssp. *durum* (Bramel 2017). The domestication of wheat dates back to about 10,000 years ago in the Fertile Crescent of the Middle East. There, close relatives of modern wheat, diploid einkorn (*T. monococcum*; genome A^mA^m) and tetraploid emmer (*T. turgidum* ssp. *dicoccum*; genome AABB) became staple crops for early civilization (Faris 2014). However, the lineage leading to hexaploid bread wheat begun earlier, about 300,000 – 500,000 years ago when wild diploid wheat (*T. urartu*; genome A^uA^u) spontaneously hybridized with the B genome progenitor, presumably a relative of goat grass (*Aegilops speltoides*, genome SS) resulting in wild emmer wheat (*T. turgidum* ssp. *dicoccoides*, genome A^uA^uBB). After its domestication and development of cultivated emmer, it hybridized with another goat grass, *A. tauschii* (genome DD) creating an ancestral allohexaploid wheat, *T. aestivum* (genome AABBDD, 2n = 6x = 42, Fig. 1; reviewed in Peng *et al.* 2011).

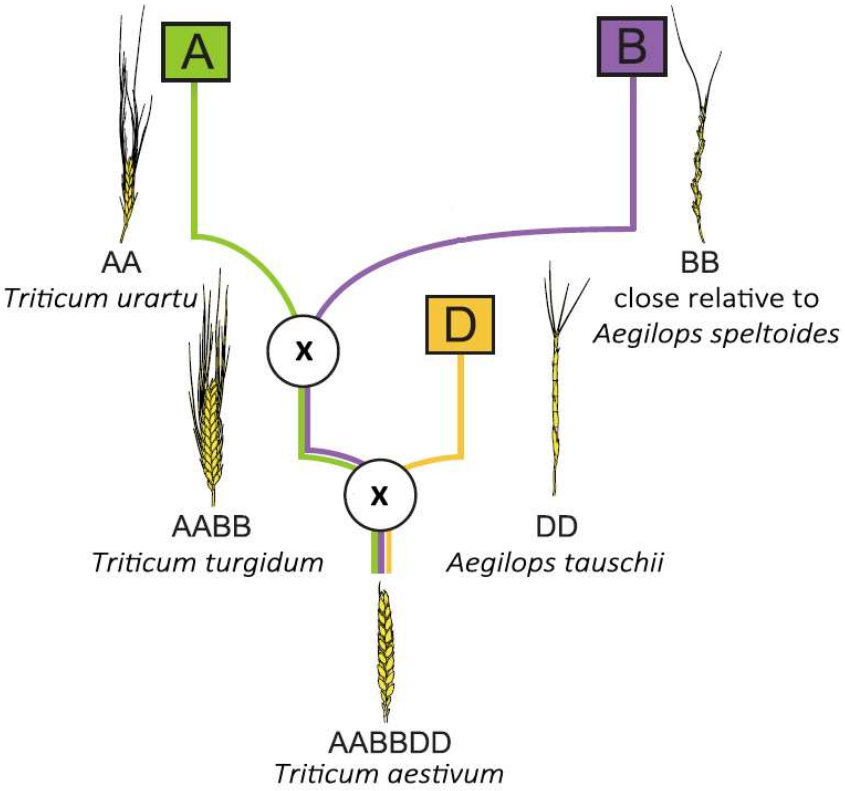


Fig. 1 The evolutionary model of modern hexaploid bread wheat originating from two hybridization events which involved three diploid lineages related to each other (indicated by different colors and letters). Modified from Marcussen *et al.* 2014.

T. aestivum is characterized by a huge nuclear genome of ~17 Gb/1C (Bennett and Smith 1976; Eilam *et al.* 2008) which combines three subgenomes of three ancestral species related with one another (Marcussen *et al.* 2014) and makes any genetic analysis challenging. Efforts to decipher the bread wheat genome were additionally hampered by an extremely high proportion of repetitive elements estimated to constitute 90 % of genome content (Li *et al.* 2004).

The polyploidy of wheat containing three homoeologous genomes allows it to tolerate aneuploidy and deletions of various extend. Due to this fact, special cytogenetic stock could be developed to facilitate physical mapping on wheat chromosomes or chromosome arms. Sears (1966) created 42 nulli-tetrasomic lines where a loss of certain pair of chromosomes was compensated by addition of a pair of homoeologs. After that, another lines of great practical importance, ditelosomics, were developed (Sears and Sears 1978). All of the materials belonged to *T. aestivum* cv. Chinese Spring and thus

determined the privileged position of this variety as a model for wheat research. Later on, ditelosomic or double ditelosomic series became a key material for a successful strategy reducing the complexity of wheat genome based on chromosome flow sorting (Gill *et al.* 1999; Doležel *et al.* 2012) till the establishment of the FISHES approach which enables sorting practically any chromosome from different wheat varieties (Giorgi *et al.* 2013). Besides aneuploid stocks, segmental deletion lines were developed based on induction of chromosomal breaks by crossing normal cv. Chinese Spring parent with a line carrying a monosomic addition of an alien gametocidal chromosome from *Aegilops cylindrica*, *A. triuncialis* or a specific chromosomal segment from *A. speltoides* (Endo and Gill 1996). All of the mentioned lines have been frequently used for assigning DNA markers or genes to individual wheat chromosomes, chromosome arms or sub-arm region (Gill *et al.* 2004; Qi *et al.* 2004). To achieve physical mapping on a finer scale, construction of chromosome-specific Bacterial Artificial Chromosome (BAC) libraries in wheat was introduced (Šafář *et al.* 2004) along with an efficient high-throughput approach of BAC clones fingerprinting and assembly into contigs (Luo *et al.* 2003) which was successfully demonstrated on wheat 3B BAC library (Paux *et al.* 2008).

Research on wheat genome was further facilitated by comparative genetics due to relatively conserved gene order (collinearity) across grass species (Gale and Devos 1998). After sequencing of smaller genomes of rice (Matsumoto *et al.* 2005), sorghum (Paterson *et al.* 2009) and *Brachypodium distachyon* (Vogel *et al.* 2010), they become model plants for cultivated grasses with more complex genomes. With an ongoing advance and improved availability of Next Generation Sequencing (NGS), another step forward was accomplished by obtaining draft genome sequences of two wheat progenitors, *T. urartu* and *A. tauschii* (Ling *et al.* 2013; Jia *et al.* 2013, respectively).

In spite of all above-mentioned resources, the necessity of sequencing complete bread wheat genome was obvious. To accomplish this goal, the International Wheat Genome Sequencing Consortium (IWGSC; http://www.wheatgenome.org/) was established in 2005. Originally, it adopted a strategy of dissecting complex wheat genome by a chromosome-based approach followed by constructing physical maps of 21 wheat chromosomes to determine a Minimal Tiling Path (MTP) and its subsequent sequencing by NGS technologies (Feuillet and Eversole 2007). The success of this

strategy was demonstrated by a 774.4 Mb-pseudomolecule of *T. aestivum* cv. Chinese Spring chromosome 3B (Choulet *et al.* 2014). The availability of whole chromosome sequence enabled quantification of different parameters which provide an insight into wheat genome in general. 85.5 % of 3B chromosome was reported to be composed of Transposable Elements (TEs), predominantly Long Terminal Repeat (LTR)-retrotransposons, and 5,326 full protein-coding genes were identified. In addition, anchoring of genetic map constructed from cv. Chinese Spring x cv. Renan population on 3B pseudomolecule clarified distribution of meiotic recombinations. All crossing over events were localized into only 13 % of the chromosome including largely distal regions. Similarly, a gradient of gene density was observed increasing from centromere towards telomeres and distal regions were shown to be enriched in non-syntenic genes related to adaptation. They originated supposedly via gene duplications or translocations enabled by double-strand break repair or gene capture by TEs, especially by CACTA DNA transposons (Choulet *et al.* 2014).

Another step forward reached by the IWGSC was a chromosome-based draft sequence of wheat with the assembly representing 61% of its genome (IWGSC 2014). Above all, this resource greatly facilitated targeted marker development and gene cloning. Recently, a shift in the original IWGSC strategy was induced by the NRGene company (http://www.nrgene.com/) whose DeNovoMAGIC™ software brought a revolution into whole-genome shotgun assembly of large and polyploid genomes. As a result, the release of IWGSC Reference Sequence v1.0 assembly and annotation was accomplished several years earlier than expected (IWGSC 2018). A high-quality reference genome sequence is an invaluable resource and a key to elucidate function of bread wheat genome, to understand the molecular basis of many important agronomic traits and finally, to accelerate wheat breeding.

2.1.2 Barley

Cultivated barley, *Hordeum vulgare* ssp. *vulgare* is a versatile crop adapted to a wide range of cultivation areas including northern parts of Asia, Europe and North America. It is more tolerant of low temperatures than other cereals and its predominant use

includes livestock feed, human consumption and malt in beverages. One of its minor application consists in using roasted grains as a coffee substitute (FAO; Horsley *et al.* 2009). The progenitor of cultivated barley, *H. vulgare* ssp. *spontaneum* (wild barley), was presumably domesticated in the Upper Jordan Valley (Mascher *et al.* 2016). However, at least two independent domestications of barley have been suggested with the second one supposedly taking place outside the Fertile Crescent (Morrell and Clegg 2007). Besides classification according to the grain use, barley is differentiated based on several other criteria. There are spring or winter types, two-row or six-row and hulled or hulless (naked) barley (Baik and Ullrich 2008).

Barley is a diploid species (2n = 2x = 14) with a haploid genome size of ~5.1 Gb (Doležel *et al.* 1998). An intensive research on barley genome organized by The International Barley Sequencing Consortium (IBSC) bore fruit by developing a whole-genome physical map anchored to a high-resolution genetic map together with providing a whole-genome shotgun sequence assembly and extensive RNA-seq data (IBSC 2012). Recently, a major milestone in barley genomics was achieved by the IBSC, a map-based reference genome sequence of spring, six-row cv. Morex was finally completed and published (Mascher *et al.* 2017). The high-quality assembly including pericentromeric regions was acquired due to chromosome conformation capture mapping, the same approach was employed to elucidate the three-dimensional organization of chromatin in interphase nuclei. A comprehensive analysis of the reference enabled revealing different features of the barley genome. 80.8 % of the sequence was identified to be derived from TEs. Surprisingly, only 10 % of them were found to be complete. Regarding the gene complement, 39,734 high-confidence genes were identified. Interestingly, genes families related to defense response and disease resistance were found to be over-represented in barley and preferentially located in distal regions of chromosomes. Similarly as in wheat, the barley reference genome sequence represents a solid platform for further research and applications such as identification of genes underlying different phenotype traits or next generation breeding (Barabaschi *et al.* 2016).

2.2 *Blumeria graminis*

Powdery mildews, a group of ascomycete fungi belonging to the order *Erysiphales*, are major plant pathogens worldwide. They earned their name due to a massive production of conidia resulting in typical appearance of affected plants. Being able to infect aerial parts of almost 10,000 angiosperm species including the most economically important ones ranks them among serious biotic stress agents. The attention focused on their control made them models for research on host-pathogen interactions despite their obligate biotrophic nature which complicates cultivation of these fungi (Glawe *et al.* 2008). Typically, powdery mildews show strict host specialization. *Blumeria graminis* (formerly *Erysiphe graminis*), the grass powdery mildew fungus, occurs only on certain wild and domesticated species of the family *Poaceae*. The extend of yield losses caused by *B. graminis* depends on a number of different factors including aggressiveness of pathogen genotype, level of resistance of host variety, environmental conditions and crop management practices. In temperate regions, powdery mildew of barley can decrease the yield by 5-20 %, under certain circumstances by as much as 40 % (Chaure *et al.* 2000). In wheat, yield reductions of 6-34 % were reported (Leath and Bowen 1989). During the infection, the pathogen acquires nutrients by forming specialized feeding structures, haustoria, in living epidermal plant cells. An important characteristic of *B. graminis* is a mixed reproduction system with prevailing haploid stage. Epidemic spread is dependent on asexual, wind-borne spores called conidia which are produced in the amount of up to 200,000 from a single colony and can overcome distances as long as several hundred kilometers in a single season. Sexual spores are produced within fruiting bodies, cleistothecia, to ensure both creating diversity of pathogen genotypes and surviving adverse environmental conditions (Fig. 2; Zhang *et al.* 2005).

Fig. 2 Fungal pathogen *Blumeria graminis*. (A) Basic morphological structures produced by the pathogen include nutrient-absorbing organs (haustoria), specialized hyphal branches producing asexual spores called conidia (conidiophores) and fruiting bodies (cleistothecia) securing sexual reproduction and survival in adverse conditions. Illustration reproduced from https://alchetron.com/Blumeria-graminis. (B) Bread wheat leaf infected by the pathogen. Black dots within the white to brownish fungal colonies represent cleistothecia. (C) Detail of conidiophores of *Blumeria graminis*.

2.2.1 Classification, host range and evolution

Due to host range specialization of *B. graminis*, a classification of this species was based on the *formae speciales* (ff. spp.) concept. Originally, the pathogen was divided into eight distinct ff. spp., each of them was supposed to infect only one host genus. Four of them are specialized to cereals including *B. graminis* f. sp. *tritici* (*Bgt*) on wheat, f. sp. *hordei* (*Bgh*) on barley, f. sp. *secalis* on rye and f. sp. *avenae* on oat and the remaining ff. spp. parasite on wild grasses. However, several studies revealed that host ranges can include even plants from more than one genus (summarized in Wyand and Brown 2003). Before the dawn of NGS era, phylogenetic analyses between different ff. spp were conducted on house-keeping genes or rDNA regions. Such results based on single or several loci often contradicted one another and the questions of host-pathogen co-evolution and divergence time remained in dispute for a long time (Wyand and Brown 2003, Takamatsu and Matsuda 2004, Inuma *et al.* 2007). Later, co-evolution of *Bgt* and *Bgh* with their respective hosts was confirmed based on large-scale sequence data from intergenic regions including orthologous TEs (Oberhaensli *et al.* 2011). Nevertheless, these results together with earlier studies were again questioned by Troch *et al.* (2014) who objected improper use of plant mutation rates for pathogen completely exposed to

sunlight and other radiation which could result in considerable overestimation of divergence times and thus incorrectly deduced co-evolution. In addition, the authors found that taxonomic classification of *B. graminis* into ff. spp. is not consistent with the fungus phylogeny and suggested that the *forma specialis* concept is not suitable for *B. graminis* infecting wild grasses and it should be therefore preserved only for forms growing on cultivated cereals. Moreover, using term "adaptation" was preferred to "strict host specialization" which was common until then. Finally, the release of draft genome sequences of *Bgh* (Spanu *et al.* 2010) and *Bgt* (Wicker *et al.* 2013) enabled comparison of respective ff. spp. and also provided an insight into their large and repeat-rich genomes. Surprisingly, the estimated genome sizes and TE content differed significantly in both ff. spp. with ~120 Mbp and 64 % TEs versus ~180 Mbp and >90 % TEs reported in *Bgh* and *Bgt*, respectively. To eventually reconstruct the evolutionary history of *B. graminis*, Menardo *et al.* (2017a) sequenced multiple isolates collected on eight different host species. Phylogenomic analysis together with coalescent-based approach revealed different phenomena such as host-pathogen co-evolution, host jumps or host range expansion and fast radiation. In addition, horizontal gene flow between recently diverged lineages was detected. These results highlight the necessity of genomic data to unveil the complex evolutionary history. One case supporting above-mentioned findings is a recently emerged *B. graminis* f. sp. *triticale* causing powdery mildew on triticale which was found to originate from a hybridization between *Bgt* and *B. graminis* f. sp. *secalis* (Menardo *et al.* 2016).

2.2.2 Diversity

The modern agriculture practices based on intensive production of genetically identical plants on large and dense crop areas favor massive development of pathogen populations (Wolfe and McDermott 1994). Distribution of gene diversity within and among individuals in such populations is greatly influenced by pathogen reproduction and mating system (McDonald and Linde 2002). Due to the ability of sexual reproduction, *B. graminis* is able to produce many new genotypes and successful clones can be subsequently spread by asexual reproduction characterized by extremely high

number of spores and long-distance dispersal. Such a mixed reproduction system is considered to be the most prone to launch an epidemic on susceptible crops.

To monitor pathogen populations and evaluate its diversity within individual ff. spp., phenotype assessment including virulence spectra and fungicide resistance had been the only criterion available prior to the advent of molecular biology techniques. Identification of different virulences is enabled by host differentials containing known resistance alleles. Due to the fast evolution and spread of pathogen, changes of avirulence gene frequencies in studied population are a highly dynamic process. Pathogen surveys therefore result in important observations applicable to disease management strategies. For example, population survey on a European scale revealed increasing virulence complexity from west to east. This phenomenon can be explained by prevailing westerly winds which carry pathogen spores to east and expose them to different host genotypes and thus a selection pressure to increase virulence complexity (Limpert *et al.* 1999). However, a higher fitness of pathotypes with fewer virulences was suggested based on regional survey repeated in several successive seasons (Dreiseitl 2015). In case of geographically distant populations, differences in virulence complexity can be substantially influenced by agricultural practices as demonstrated by Dreiseitl (2014).

From the last decade of the 20th century on, development of first and second generation of DNA markers enabled supplementing phenotype data with data on genetic diversity of pathogen populations (Brown 1996). Listing all available DNA-based marker techniques is beyond the scope of this short review, a comprehensive summary is provided e.g. by Agarwal *et al.* (2008); Semagn *et al.* (2006). Applying DNA markers on *B. graminis* samples from Europe, Asia and North America revealed lack of correlation between geographical and genetic distances. This finding can be explained by long-distance dispersal of *B. graminis* spores (Wyand and Brown 2003).

In the genomic era, a new exciting perspective of resequencing more isolates to study genetic diversity both between and within ff. spp became available. Wicker *et al.* (2013) compared genomic data of three additional *Bgt* isolates of different geographical or temporal origin to the draft genome reference and identified 537 deletions longer than

500 bp. Interestingly, some of these deletions were found to include genes. Moreover, ~114 to ~161 thousand Single Nucleotide Polymorphisms (SNPs) were detected between the reference sequence and resequenced isolates. The highly uneven distribution of SNPs throughout the genome revealed a mosaic of different haplogroups. The same observation was reported by Hacquard *et al.* (2013) who performed a similar analysis with *Bgh* isolates.

2.2.3 Control of powdery mildew disease

Intensive cultivation of cereals provides favorable conditions for the powdery mildew fungus. In spite of its obligate biotrophic nature, survival of the pathogen over winter is ensured by green bridges of host plants due to the continuity of autumn- and spring-sown cultivars (Wolfe and McDermott 1994). Traditionally, chemical control, breeding for resistance in the host and crop management practices have been employed to control powdery mildews (Chaure *et al.* 2000).

Since the introduction of novel effective fungicides in the 1980s, their widespread use has become common practice (Jørgensen *et al.* 2014). Besides limited sustainability, a significant problem of chemical control consists in fungicide resistance. From the early history of fungicide application on, mutants with reduced sensitivity or even insensitivity appeared shortly after introduction of new type of fungicide (Wolfe 1984). The three groups of fungicides used against *B. graminis* include benzimidazoles, triazoles and strobilurins. For all of them, a mutation causing insensibility is known (http://eurowheat.au.dk/). To deal with the problems of fungicides, Wolfe (1984) suggested integrating strategies based on both disease resistance and fungicides. Indeed, increasing crop resistance to fungal diseases has been recognized as a means of reducing the need for fungicide application (Loyce *et al.* 2008). The ultimate goal should be growing durably resistant varieties with no required disease treatment (Wolfe 1984).

Breeding crop plants for resistance represents the most effective and environment-friendly approach to control any disease. Resistance genes have been therefore deployed by plant breeders for almost a century (Dodds and Rathjen 2010). However, the breakdown of major resistance genes (see chapter **2.3.1.1**) after their extensive

distribution over a large area has become a commonly observed phenomenon (Wolfe and McDermott 1994). Its basis can be well explained by the high evolutionary potential of *B. graminis* which poses a substantial risk of breaking down resistance conferred by single major-effect genes. Considering the nature of this pathogen, resistance-breeding strategies based preferably on quantitative resistance (see chapter **2.3.1.3**) or alternatively on development of cultivar mixtures and multilines have been recommended (McDonald and Linde 2002).

The problem of insufficient crop diversity in present-day agricultural ecosystems can be addressed by exploiting available genetic resources, either collections of landraces or wild relatives. The available gene pool provides a rich and valuable source of novel powdery mildew resistances potentially applicable in breeding (e.g. Ames *et al.* 2015; Dreiseitl 2017; Hysing *et al.* 2007; Li *et al.* 2016). Although introgression breeding based on exploiting traits identified in wild relatives of secondary or tertiary gene pools is challenging (Boyd *et al.* 2013), improved technologies for genetics and genomics together with modern breeding approaches have been constantly increasing our opportunities to make use of these resources (Mondal *et al.* 2016; Tester and Langridge 2010). Since introduction of disease resistance by interspecific hybridization is a slow process with drawbacks such as linkage drag and limited or completely suppressed recombination, a strategy based on genetically modified cassette of several resistance genes provides a tempting alternative (Wulff and Moscou 2014). Rapidly progressing technologies of genetic engineering make creating pyramids of major resistance genes or alleles feasible. Clearly, a prerequisite of this approach is prior identification of such genes. However, commercial use of genetically modified crops has been complicated by public non-acceptance resulting from poor understanding and by regulatory restrictions, especially within the European Union or China (Huang *et al.* 2016; Wolt *et al.* 2016). Due to the precision of currently available genome editing techniques, product-based regulation is recommended instead of technology-based regulation which is still very frequent (Huang *et al.* 2016).

Besides chemical control and strategies based on breeding for durable resistance, certain cultural measures can reduce the disease prevalence. For wheat, these crop management practices are summarized in Jørgensen *et al.* (2014). Generally, lower risk

of powdery mildew disease is associated with early sowing, low crop densities and low input of nitrogen. However, some practices can have contradictory effect on different pathogens resulting in various dilemmas. Besides that, optimizing the economic yield represents a principal consideration for farmers. Another option consists in increasing the overall agroecosystem diversity at different scales (e.g. field, farm, landscape) and ensuring its dynamic changes over time and space to slow the rate of pathogen adaptation. This can be achieved either by growing populations composed of many genotypes or by selecting for populations composed of many resistance-conferring genes (McDonald 2014). In spite of the advantage granted by growing variety mixtures, this strategy was not commercially successful in agricultural practice (Wolfe 1984).

2.3 Plant-pathogen interactions

Any plant pathogen which happens to collide into any plant has to deal with its innate immune system. The first line of plant defense is responsible for the fact that most plants are resistant to most fungal pathogens. Such a basic incompatibility also called nonhost resistance is caused by specialization of a pathogen to narrow host range. The basic principle consists in recognition of Pathogen-Associated Molecular Patterns (PAMPs) by plant transmembrane Pattern Recognition Receptor (PRRs). Usually, PAMPs are indispensable and conserved molecules of the pathogen, such as chitin in case of fungi. The result, PAMP-Triggered Immunity (PTI), is able to stop colonization of non-adapted pathogen. Typically, PRRs are considered to be transmembrane proteins with extracellular Leucine-Rich Repeat (LRR) domain and intracellular kinase domain. To be able to obtain nourishment from the host and complete their life cycles, adapted pathogens have developed a specialized mechanisms to suppress PTI. The components involved in promoting successful infection are called effectors and in general, they contribute to pathogen virulence. They can act either inside the host cell or extracellularly, their structure and function are highly variable and yet insufficiently explored. Suppression of PTI by effectors is designated Effector-Triggered Susceptibility (ETS). However, plants have developed another line of defense which is localized predominantly inside the cell and recognizes pathogen effectors. Importantly, this

mechanism is applicable only to biotrophic or hemi-biotrophic pathogens. In the process, a given effector termed as Avirulence (AVR) protein is directly or indirectly (through an accessory protein) recognized by particular plant receptor in a highly specific manner and Effector-Triggered Immunity (ETI) is induced. These receptors are encoded by *R* genes and very often, they belong among the Nucleotide-Binding (NB)-LRR proteins. ETI is faster and stronger than PTI and induces resistance which is usually associated with a Hypersensitive Response (HR; Fig. 3). A typical feature of pathogen effectors is their fast evolution or even potential loss caused by high level of natural selection to evade recognition by the second layer of plant defense. Moreover, it has been suggested that some effectors are able to suppress ETI triggered by other effectors. This results in a strong selection pressure posed upon the host to evolve receptor with new recognition specificities to restore ETI. Thus, interactions of a host and adapted pathogen result in co-evolution in a form of ongoing evolutionary arms race (Jones and Dangl 2006; Dodds and Rathjen 2010). The co-evolution of plant-pathogen interactions can be illustrated by a renowned and widely used zigzag model (Fig. 4). However, Pritchard and Birch (2014) highlighted limitations of this model due to restricted scope and oversimplification and suggested replacing it by more complex dynamic and predictive models.

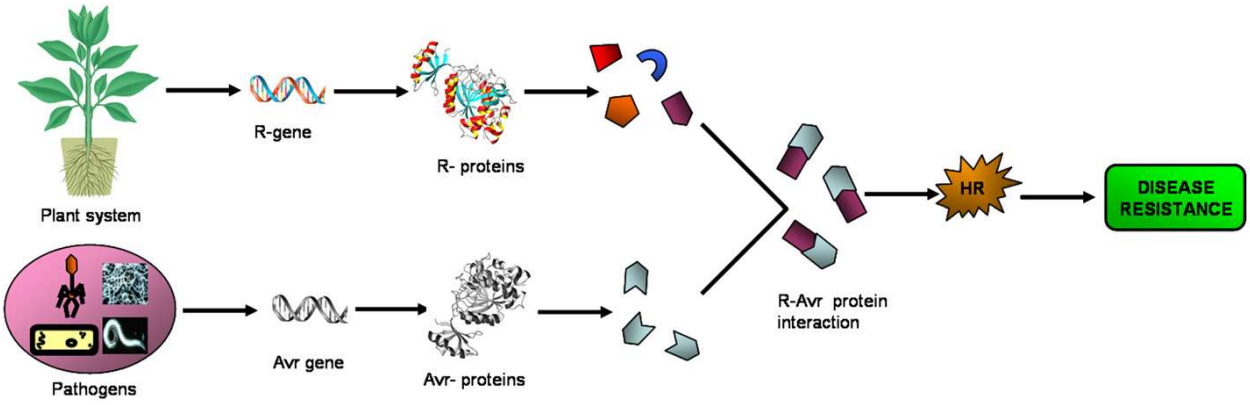


Fig. 3 Plant-pathogen interaction resulting in disease resistance (Gururani *et al.* 2012). R = resistance, Avr = avirulence, HR = hypersensitive response.

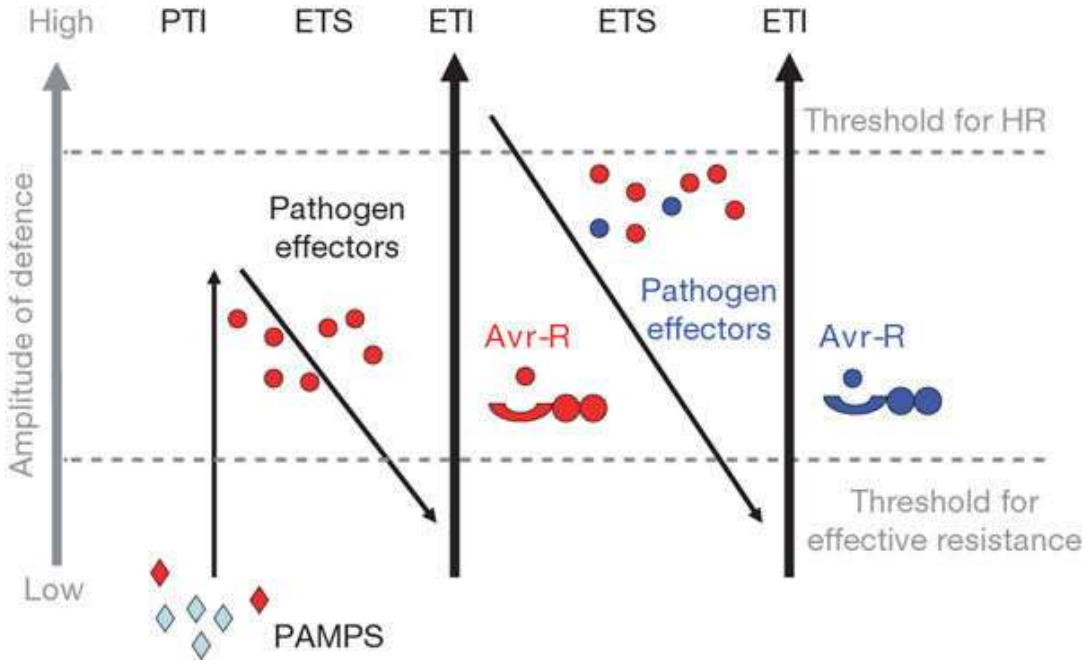


Fig. 4 A zigzag co-evolutionary model of plant-pathogen interactions. First, plant pattern recognition receptors recognize pathogen-associated molecular patterns (PAMPs) and induce PAMP-triggered immunity (PTI). After that, pathogen employs effectors to mediate suppression of PTI which results in effector-triggered susceptibility (ETS). In the next step, plant resistance (R) protein can specifically recognize certain effector in this case called avirulence (Avr) protein which leads to activation of effector-triggered immunity (ETI). ETI often crosses the threshold of hypersensitive cell death (HR). Finally, pathogen can lose or modify the recognized effector or gain new effector able to suppress ETI. A selection pressure working on plant can result in new R allele recognizing modified or new effector and ETI is triggered again (Jones and Dangl 2006).

2.3.1 Plant resistance genes

Genetically-based plant resistance manifests itself in different ways and can be classified based on several criteria including resistance specificity, effect size or type of inheritance. Usually, two main categories of resistance are used. First, quantitative resistance (also called horizontal resistance) is usually characterized by race non-specificity, small effect and quantitative inheritance. In contrast, major-gene resistance (also called vertical resistance) is considered to be race-specific with large effect and simple inheritance. The later type results in the most explored form of interaction between plant and pathogen, the Gene-For-Gene (GFG) concept (Flor 1971). The hypothesis states that for each gene conditioning resistance in the host, there is a corresponding gene in the parasite conditioning pathogenicity. As a result, each

component in the host can be identified only by its counterpart in the pathogen and vice versa. Resistance was found to be dominant as well as avirulence, presence of resistance allele together with a corresponding avirulence allele in host-pathogen interaction leads to incompatible response with no disease. However, losing any of these interacting components results in compatible response and disease outbreak. This is in agreement with modern ETI and ETS concept. The practical implications of this finding were huge, however, employing major resistance genes in breeding was shown to have only short-term effect due to fast evolution occurring in pathogen populations as demonstrated by the zigzag model. As a consequence of "boom and bust cycle" describing frequencies of resistance and virulence alleles in a population over time (Fig. 5), alternative strategies such as increasing the genetic diversity of crops and breeding for durable quantitative resistance have been proposed (Brown and Tellier 2011). In comparison with major-gene resistance, other types of resistance are still far from being thoroughly explored and well understood. Nevertheless, the need of durable disease resistance in crop breeding generated increased attention in this field and as a result, new findings have gradually been contributing to our knowledge. Next subchapters aim to provide a very brief general overview of main resistance categories with a more detailed insight into components of each resistance class efficient against powdery mildew fungus which were identified in wheat and barley to date.

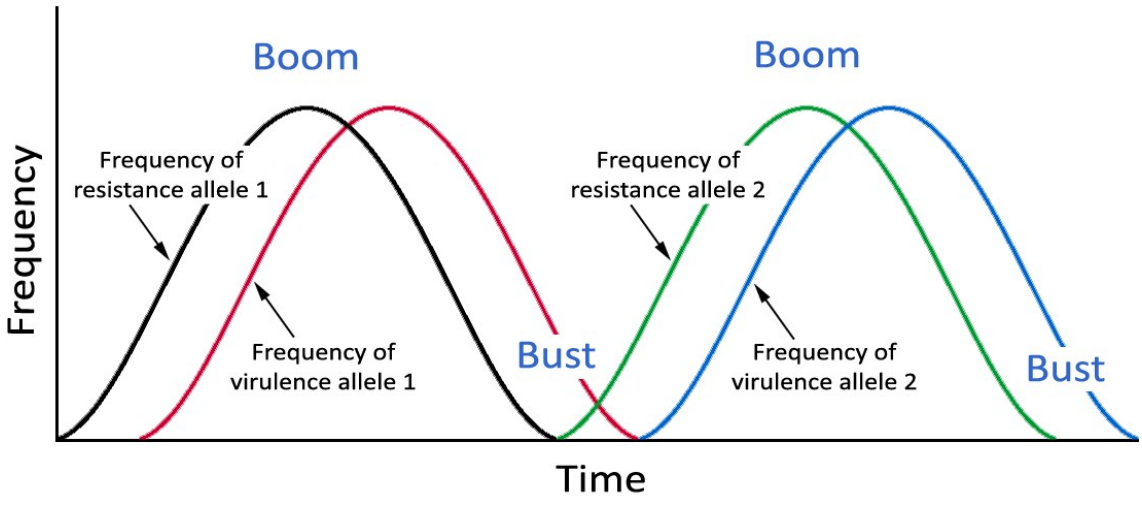


Fig. 5 A phase-diagram representing the boom-and-bust cycle based on oscillation in frequencies of major resistance alleles in the host population and corresponding virulence alleles in the pathogen population. Modified from https://www.apsnet.org.

2.3.1.1 Major-gene resistance

A progress in gene cloning revealed that most *R* genes encode NB-LRR protein containing variable amino- and carboxy-terminal domains. These genes were found to be extremely abundant in plant genomes and they are frequently found in clusters. Generally, they show high levels of inter- and intraspecific variation generated presumably by unequal crossing over or gene conversion. The plant NB-LRR proteins were originally classified into two major subfamilies based on the presence of Toll/Interuleukin-1 Receptor (TIR) or Coiled-Coil (CC) motifs in the N-terminal domain. Unlike CC-NB-LRR proteins, TIR-NB-LRR proteins have not been detected in cereal species. The N-terminal domain is suggested to be involved in protein-protein interactions while the LRR domain is probably responsible for recognition specificity together with a regulatory function in signal transduction and seems to be under diversifying selection. The NB domain presumably activates downstream signaling through conformational changes induced by the exchange of ADP for ATP. Some TIR-NB-LRR proteins were found to carry a C-terminal domain with a nuclear localization signal and a WRKY motif. Generally, NB-LRR proteins are assumed to be localized in cytoplasm due to lack of signal peptide or transmembrane domain and they can induce a range of plant defense responses including oxidative burst, mitogen-associated protein kinase cascade, expression of pathogenesis-related genes or the hypersensitive response (reviewed in DeYoung and Innes 2006; McHale *et al.* 2006). Later on, a summary of Gururani *et al.* (2012) showed that the variability of R proteins is considerably higher. Based on their classification, there are altogether eight classes of R proteins defined by their domain compositions, majority of them containing the LRR domain (Fig. 6). However, some of them might have different functions than true R proteins mediating ETI.

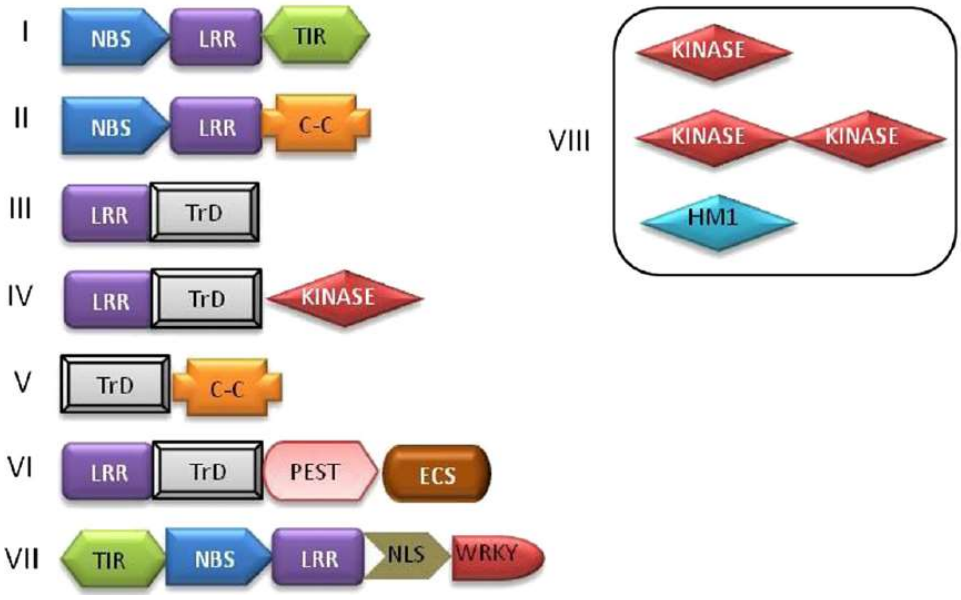


Fig. 6 Major classes of plant resistance genes according to Gururani *et al.* (2012). Classification is based on the arrangement of the functional domains of respective proteins (LRR – Leucine rich repeats; NBS – Nucleotide-binding site; TIR – Toll/Interleukin-1 receptors; C-C – Coiled coil; TrD – Transmembrane domain; PEST – Protein degradation domain (proline-glycine-serine-threonine); ECS – Endocytosis cell signaling domain; NLS – Nuclear localization signal; WRKY – Amino acid domain; HM1 – *Helminthosporium carbonum* toxin reductase enzyme.

Considering only major resistance genes/alleles in cereals conferring powdery mildew resistance, several of them have been cloned to date. The *Mla* locus located on short arm of barley chromosome 1H harbors approximately 30 resistance specificities identified by genetic studies (Jørgensen 1994, Kintzion *et al.* 1995). Wei *et al.* (1999) enabled cloning of different *Mla* alleles by physically delimiting the locus into an interval of 240 kbp containing a cluster of NB-LRR resistance gene homologues which belong to three distinct families. Subsequently, identification of *Mla1* (Zhou *et al.* 2001), *Mla6* (Halterman *et al.* 2001), *Mla12* (Shen *et al.* 2003), *Mla13* (Halterman *et al.* 2003), *Mla7* and *Mla10* (Halterman and Wise 2004) followed. All of the allelic variants were found to encode R proteins of the CC-NB-LRR class and besides MLA1 and MLA7, remaining R proteins required the zinc-binding protein RAR1 (Shirasu *et al.* 1999) for disease resistance signaling (Jørgensen 1996). Seeholzer *et al.* (2010) identified 13 additional validated *Mla* alleles and suggested direct binding of effectors to MLA receptors. In wheat, race-specific resistance to wheat powdery mildew is regulated by the *Pm* genes. Analogously to barley *Mla* locus, a number of different resistance specificities were

mapped to wheat *Pm3* locus on short arm of chromosome 1A. A total of ten alleles were designated *Pm3a* to *Pm3j* (Hsam and Zeller 2002) and *Pm3b* encoding a member of the CC-NB-LRR protein class became the first one to be cloned (Yahiaoui *et al.* 2004). Srichumpa *et al.* (2005) identified three additional alleles *Pm3a*, *Pm3d* and *Pm3f*, all of them showed high sequence conservation with *Pm3b*. Finally, Yahiaoui *et al.* (2006) completed the set of cloned *Pm3* alleles. However, no unique allele of *Pm3* was identified in case of *Pm3h*, *Pm3i* and *Pm3j* and the resistance of respective wheat lines was therefore attributed to additional loci. Later on, nine new functional alleles were isolated by screening more than two thousand bread wheat gene bank accessions (Bhullar *et al.* 2009; Bhullar *et al.* 2010). Intriguingly, rye *Pm8* gene was found to be an orthologue of *Pm3* allelic series (Hurni *et al.* 2013) and when present in wheat genome as a part of 1BL·1RS translocation from rye, *Pm8*-mediated resistance is suppressed by functional *Pm3* gene (Hurni *et al.* 2014). Similarly, Stirnweis et al. (2014) showed that pyramiding of different pairs of *Pm3* alleles can result in suppression of *Pm3*-based resistance in laboratory experiments. However, this effect was not confirmed in field trials (Koller *et al.* 2018). Most recently, *Pm2* gene localized on chromosome 5D and encoding a CC-NB-LRR protein was cloned in hexaploid bread wheat (Sánchez-Martín *et al.* 2016) and *Pm60* encoding the same type of protein was identified on chromosome arm 7AL of *T. urartu* (Zou *et al.* 2017). Despite being cloned and characterized in a wild grass, the gene was shown to be functional even in cultivated bread wheat.

2.3.1.2 *mlo*-based resistance

A unique case of monogenic, broad-spectrum powdery mildew resistance in barley is mediated by recessive loss-of-function alleles of the barley *Mildew resistance Locus O* (*Mlo*). This resistance does not match the GFG system and has been durable since its deployment in breeding in 1980s. Its molecular mechanism consists in formation of cell wall appositions which prevent the fungus from entering the host cell. In addition, an attempt of pathogen penetration induces enhanced epidermal and mesophyll oxidative burst with higher hydrogen peroxide levels than those observed in wild-type *Mlo* genotypes. On the other hand, *mlo*-mediated resistance has a negative effect on yield

due to accelerated leaf senescence observable as spontaneous mesophyll cell death (Jørgensen 1992, Piffanelli *et al.* 2002). Cloning the wild-type *Mlo* gene revealed that it encodes a novel-class protein anchored to plant plasma membrane by seven Transmembrane (TM) helices. Furthermore, close homologues to *Mlo* were found in rice and *Arabidopsis thaliana* suggesting conserved function among monocots and dicots (Fig. 7). Considering properties of proteins belonging to the *Mlo* family, they are reminiscent of the G-protein-coupled receptors (Büschges *et al.* 1997; Devoto *et al.* 1999). However, association of G proteins with the MLO system was not confirmed and MLO activity was proved to be enhanced by calmodulin binding (Kim *et al.* 2002).

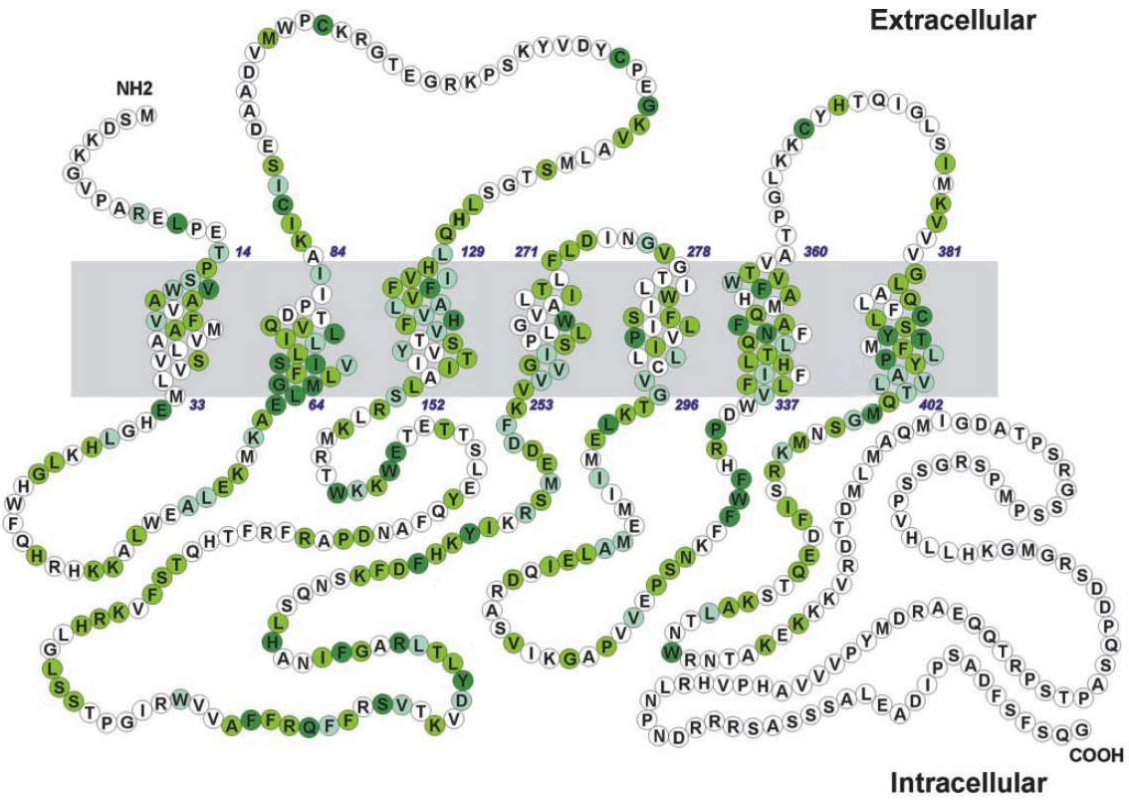


Fig. 7 A common 7-TM scaffold topology of the *Mlo* family. The plasma membrane is illustrated by grey color, circles with letters represent amino acids, numbers indicate amino acid positions. A color code highlights conserved and variable domains among family members (dark green – invariant residues; light green – conservative changes; green – non-conservative changes). Reproduced from Devoto *et al.* 1999.

The function of *mlo*-mediated resistance in barley was found to be dependent on two additional loci designated *Required for mlo-specified resistance1* (*Ror1*) and *Ror2*

(Freialdenhoven *et al.* 1996). Collins *et al.* (2003) isolated *Ror2* encoding a syntaxin gene homologous to *Penetration1* (*Pen1*) gene of *A. thaliana* involved in nonhost resistance. In addition, different experiments carried out on barley and *A. thaliana* suggest that *mlo*-based resistance is actually a consequence of deregulated PTI (Hückelhoven and Panstruga 2011) and it was hypothesized to share identical mechanism with nonhost resistance (Humphry *et al.* 2006). Recently, transgenic knock-out mutation of three *Mlo* orthologues in bread wheat was reported resulting in heritable broad-spectrum resistance to powdery mildew (Wang *et al.* 2014). Enhanced powdery mildew resistance in wheat was achieved also by non-transgenic Targeted Induced Lesions IN Genomes (TILLING) technology (Acevedo-Garcia *et al.* 2017). Due to the fact that *mlo*-based resistance has already been described in many other plant species, it is depicted as a universal mechanism to control powdery mildew disease in angiosperms (Kusch and Panstruga 2017).

2.3.1.3 Quantitative resistance

Quantitative Resistance (QR) is a collective term encompassing mechanisms of diverse biological and molecular basis. It can be described as phenotypically incomplete host plant resistance causing a reduction of disease which is mediated by the joint effect of several genes. Sometimes, it is also called partial resistance, adult-plant resistance, field resistance, slow-rusting or slow-mildewing etc. (Niks *et al.* 2015; St Clair 2010). However, QR is not strictly demarcated and categorization of some resistance genes might not be unequivocal if we consider different aspects (Niks *et al.* 2015). The classification presented here is therefore somewhat simplified. Based on work in several crop species, a relation between QR and PTI has been proposed (Boyd *et al.* 2013). Generally, the common characteristics of QR include durability, broad-spectrum specificity and lack of HR. Due to its properties, this type of resistance has become favored by some current cereal breeding programs (Krattinger and Keller 2016). Traits with complex inheritance can be genetically dissected by Quantitative Trait Loci (QTL) mapping (Sehgal *et al.* 2016). However, subsequent cloning of identified loci is often substantially hampered by arduous phenotyping. Most probably, this explains the fact

that the number of identified genes involved in QR is significantly lower than the number of cloned *R* genes. With only a few diverse representatives of QR-associated genes available to date, we can still anticipate new mechanisms to be discovered in the future. Cloning of new QTL will allow optimizing their combinations with respect to underlying mechanisms (Boyd *et al.* 2013).

In 2009, Krattinger *et al.* (2009a) cloned a partial, adult-plant wheat leaf rust resistance gene *Lr34* which encodes a putative ATP-Binding Cassette (ABC) transporter. Importantly, the same gene controls partial resistance to stripe rust (*Yr18*) and powdery mildew (*Pm38*) and it is associated with a leaf tip necrosis phenotype which facilitated the cloning process. Another example of successfully cloned QR-involved gene is *Yr36* conferring partial non race-specific resistance to stripe rust at relatively high temperatures and originating from wild emmer wheat. The corresponding protein comprises a kinase and a putative START lipid-binding-domain (Fu *et al.* 2009). Most recently, the wheat *Lr67* gene conferring partial resistance to leaf rust, stem rust (*Sr55*), stripe rust (*Yr46*) and powdery mildew (*Pm46*) has been identified. Again, the phenotype includes also leaf tip necrosis. In this case, the LR67 protein is a predicted hexose transporter (Moore *et al.* 2015).

2.3.1.4 Nonhost resistance

Nonhost Resistance (NHR) is the most common type of plant resistance and it is exhibited by an entire plant species to all members of a certain pathogen species (Heath 2000). Mysore and Ryu (2004) proposed a classification of NHR against bacteria, fungi and oomycetes into two distinct types. The first one, type I, does not result in any visible symptoms while the second one, type II, is associated with a hypersensitive response. A single nonhost plant species can show both types of NHR on different pathogen species. On the other hand, a single pathogen species can triggers both types depending on plant species. In type I, pathogen is first confronted by preformed plant barriers (such as antimicrobial compound or cell wall). After that, it must face inducible plant defense responses triggered by plant recognition of general pathogen elicitors. These responses include cell wall thickening or lignification, papilla formation, accumulation of

phenolics or different compounds and induction of pathogenesis-related genes. This description corresponds to PTI as mentioned earlier. The type II producing a nonhost HR phenotypically reminds an incompatible gene-for-gene interaction. The nonhost pathogen can surmount preformed and general elicitor-induced plant defense responses presumably due to production of detoxifying enzymes. However, pathogen elicitors are recognized by the plant surveillance system and defense reaction leading to HR is triggered. Results of several experiments suggest that both host and nonhost resistances can share a common pathway (reviewed in Mysore and Ryu 2004). Schulze-Lefert and Pastruga (2011) gathered an evidence for the involvement of NB-LRR genes in NHR and proposed a unifying model which states that both NB-LRR- and PRR-triggered immunity are involved in NHR in a way that the relative contribution of PTI increases and the relative contribution of NB-LRR protein-triggered immunity decreases with growing phylogenetic divergence time between two plant species. Thus, NHR depends on the same components that operate also in host immunity and according to the model, type I is characteristic for species distantly related to the host while type II with HR occurs in closely related species.

Due to its strength, durability and broad-spectrum nature, a better understanding of NHR has become highly desirable. It has been observed that NHR of barley against *Bgt* is associated with hydrogen peroxide accumulation in cell wall appositions beneath attempted penetration sites and in penetrated cells undergoing HR (Hückelhoven *et al.* 2001). However, clarifying the genetic basis underlying NHR is a complex task. One approach to study determinants of NHR is screening of different accessions of a nonhost species for some degree of susceptibility to a particular heterologous pathogen followed by accumulation of genes for susceptibility in a single genotype usable for QTL mapping. This approach has been demonstrated in barley – heterologous rusts pathosystem (Atienza *et al.* 2004; Jafary *et al.* 2006; Jafary *et al.* 2008) and recently also in barley – wheat powdery mildew pathosystem (Aghnoum and Niks 2009; Romero *et al.* 2015; Romero *et al.* 2018). Moreover, an approach of transient-induced gene silencing of candidate genes for NHR and QR in barley against the powdery mildew fungus yielded 96 genes with a significant effect (Douchkov *et al.* 2014). One of them, *Rnr8*, encoding a transmembrane receptor-like kinase with a LRR and malectin domain was shown to

contribute to nonhost resistance of barley to the non-adapted *Bgt.* Notably, the same gene also improved quantitative host resistance of wheat to the adapted powdery mildew fungus (Rajamaran *et al.* 2016).

2.3.2 Pathogen effectors

Plant pathogens through different kingdoms secrete proteins or other molecules to modify host cell structure and function. The purpose of these modifications consists in promoting infection (virulence factors and toxins) but they can also activate defense responses (avirulence factors and elicitors) or both depending on host susceptibility. Generally, these secreted molecules are termed effectors. Their mode of action can be complex, one effector can have more than one host targets. Frequently, effectors participate in suppression of plant immunity but they can also alter host plant behavior and morphology. Uncovering molecular function of effectors is crucial for understanding the mechanisms involved in pathogenicity (Hogenhout *et al.* 2008).

Recent boom of whole-genome sequencing projects has revealed a remarkable trend observed in genomes of filamentous plant pathogens belonging to the groups of fungi and oomycetes. Generally, their genomes are expanded due to a high proportion of TEs. Moreover, they show a mosaic-like character with uneven evolutionary rates. Large repertoires of genes encoding effector proteins are frequently associated with gene sparse, repeat rich compartments which presumably act as a cradle for adaptive evolution according to the "two-speed genome" model (reviewed by Dong *et al.* 2015).

While some interesting findings gradually help to uncover the function of bacterial effectors (reviewed in Dodds and Rathjen 2010), our knowledge on eukaryotic effectors, their function and mechanism of delivery into host cells is still scarce. It seems that the repertoire of eukaryotic effectors is vast and extremely poor in conserved sequence motifs (Fig. 8). Among a few documented exception, there is the RXLR amino acid motif typically occurring in oomycete effectors (Birch *et al.* 2006) and the N-terminal effector motif Y/F/WxC discovered in *Bgh* and wheat rust fungi (Godfrey *et al.* 2010). Besides this conserved motif, a number of *Bgh* genes expressed in haustoria share small size, N-terminal signal peptides and a highly similar exon-intron structure suggesting

a common origin. Based on these characteristics, the Y/F/WxC proteins were suggested to constitute a novel class of effectors from haustoria-producing pathogenic fungi.

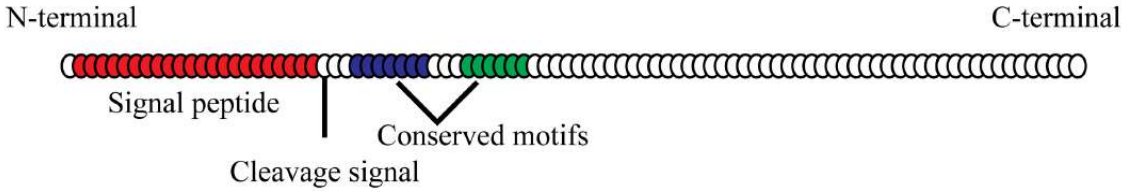


Fig. 8 Typical protein organization of fungi/oomycetes effectors with signal peptide, cleavage site and conserved domain present toward the N-terminus (reproduced from Sonah *et al.* 2016)

In silico genomic prediction complemented by expression analysis which were conducted in barley powdery mildew fungus (Pedersen *et al.* 2012; Spanu *et al.* 2010) identified nearly five hundred Candidate Secreted Effector Proteins (CSEPs). Interestingly, structure similar to ribonucleases was detected throughout the entire superfamily and candidate effector genes were found to be associated with retrotransposon-like repetitive DNA. Moreover, most CSEPs were predominantly expressed in haustoria. The prediction of CSEPs was subsequently fine-tuned by Kusch *et al.* (2014). Similarly, an analysis in the wheat powdery mildew fungus yielded 437 *Blumeria*-specific CSEPs. Together with candidate effector proteins without a signal peptide, a set of six hundred putative effector genes was defined (Wicker *et al.* 2013). In both ff. spp., positive selection was observed as a general trend acting upon candidate effector genes. Most recently, Menardo *et al.* (2017b) conducted an extensive analysis in five *B. graminis* lineages and identified a total of 2798 candidate effector genes. Besides *Bgh* and *Bgt* isolates with available genome sequences, genomes of *B. graminis* infecting oat, *Lolium* and *Poa* were sequenced and mined. A general trend observed across all predicted effector families was a high turnover rate contributing to the fast diversification of effector repertoires in these lineages and presumably to their adaptation to different hosts.

To date, several functional studies have been undertaken to elucidate the role of *Bgh* CSEPs in virulence. A number of CSEPs was proved to be required for normal

haustoria formation rate (Aguilar *et al.* 2015; Ahmed *et al.* 2015; Ahmed *et al.* 2016; Pennington *et al.* 2016), some of them were found to contribute to the fungal penetration as well (Ahmed *et al.* 2016). Less typically, Zhang *et al.* (2012) validated a CSEP involved in plant defense suppression mediated probably by interaction with barley pathogenesis-related protein families. Moreover, several other host targets of different CSEPs were identified including small heat shock proteins with chaperone activity (Ahmed *et al.* 2015), glutathione–S-transferase or malate dehydrogenase (Pennington *et al.* 2016). All of these functional CSEPs act as virulence factors contributing to the infection success.

Another significant progress has been recently achieved in identification of *B. graminis* effectors recognized by *R* gene-encoded receptors, i.e. avirulence proteins. Bourras *et al.* (2015) cloned the *Bgt* avirulence effector designated $AvrPm3^{a2/f2}$ which is recognized by two different alleles of wheat *Pm3* resistance gene (*Pm3a* and *Pm3f*). This study also demonstrated that the *Pm3* race-specific resistance is controlled by multiple interacting loci in the fungus, avirulence genes and avirulence gene suppressors. Moreover, it was shown that different *Pm3* alleles can recognize distinct *Avr* factors. Parlange *et al.* (2015) cloned another component of this system modifying the *Pm3*/*AvrPm3* interactions, presumably by suppression mechanism. Such a striking complexity discovered in the wheat-powdery mildew pathosystem has led to formulation of updated GFG hypothesis (Bourras *et al.* 2016). The extension of original model contains a Suppressor of avirulence (*Svr*) besides the two basic components, *Avr* and *R* genes. Resistance is mediated by an interaction involving a resistance gene allele with a corresponding allele-specific avirulence effector and an allele-unspecific *Svr* encoded by pathogen. In a presence of active SVR, recognition of AVR effector by its specific R protein does not induce a defense response. The AVR protein can therefore keep its potential function in virulence and pathogen still avoids ETI induced by AVR recognition. In addition to avirulence determinants involved in interactions with *Pm3*, Praz *et al.* (2017) cloned another *Bgt* avirulence gene, *AvrPm2* encoding an effector specifically recognized by wheat *Pm2*. Interestingly, *Pm2* was found to recognize also a close homologue of *AvrPm2* from rye powdery mildew fungus. Furthermore, the study of Lu *et al.* (2016) reported identification of Avr_{a1} and Avr_{a13} effector genes of barley

powdery mildew fungus which encode proteins recognized by *Mla1* and *Mla13* alleles of the barley *Mla* locus, respectively. Notably, all of these avirulence factors belong among predicted CSEPs and barley Avr_{a13} is a member of the same effector family characterized by structural homologies to fungal ribonucleases as wheat *AvrPm2* (Praz *et al.* 2017).

2.4 Identification of interacting components

Biotic stresses including biotrophic fungal pathogens represent a major constraint to crop production. It is therefore important to elucidate the mechanisms of crop-pathogen interaction including identification of interacting components and molecular basis of their function and to apply this knowledge in agricultural strategies. Fortunately, the current era of available high-quality genome sequences has eliminated a substantial barrier that researchers had to confront during the last decades. Indeed, newly identified resistance genes or effectors has been springing up like mushrooms after the rain in the last few years and gene cloning is no longer considered to be a herculean task even in genomes as large and complex as the one of bread wheat. The following subchapters describe several approaches frequently employed in wheat, barley and the respective powdery mildew fungi including recent innovations which significantly accelerate the research.

2.4.1 Map-based cloning

Map-based or positional cloning is a standard procedure applicable for genetically mapped traits. It means that this approach requires prior establishment of a mapping population, a feasible task in case of cereals but not an easy one when dealing with the powdery mildew fungus. However, mapping populations of *B. graminis* have been already successfully employed for map-based gene cloning (Bourras *et al.* 2015; Parlange *et al.* 2015). A conventional map-based cloning of a major-effect locus consists of several basic steps (Fig. 9). Briefly, two parents differing in the trait of interest are a basis for development of a mapping population. In general, the more individuals form this

population the better in terms of mapping resolution. In the next step, a genetic map is developed by integrating phenotype and molecular data obtained by screening the population with appropriate polymorphic markers. Next, to span the locus of interest physically, a BAC library (earlier a cosmid library or a yeast artificial chromosome library) is required. First, it is screened by the closest flanking markers and then, additional markers are developed based on BAC sequences and they are subsequently used for another round of BAC library screening. This approach called chromosome walking results in establishment of a contig spanning the target interval. After that, candidate genes are identified in sequence of the interval and finally, they need to be validated (summarized in Krattinger *et al.* 2009b). A significant time and labor reduction occurs in case a physical map of appropriate BAC library is already available. A successful example is set by physical maps of bread wheat cv. Chinese Spring chromosomes or chromosome arms (IWGSC 2018; https://urgi.versailles.inra.fr/gb2/gbrowse/wheat_phys_pub/). On the other hand, the standard map-based cloning procedure gets more difficult when dealing with QTL. First, it is necessary to dissect them using populations of double haploids, recombinant inbred lines, near isogenic lines or multiparent advanced generation intercross (Wulff and Moscou 2014) since the locus of interest needs to be mendelized prior to its cloning. Despite possible complications caused by partial phenotype, even QTL can be cloned in complex genome as demonstrated by Uauy *et al.* (2006).

Before whole genome sequences became available for cereal crops with large genomes, conserved gene order along chromosome segments (so called synteny) of grass species (Gale and Devos 1998) had been frequently exploited as a source of genetic markers for fine-mapping stage of map-based cloning (e.g. Fu *et al.* 2009, Griffiths *et al.* 2006; Krattinger *et al.* 2009a; Liu *et al.* 2008; Perovic *et al.* 2004; Saintenac *et al.* 2013; Vu *et al.* 2010; Yeo *et al*. 2017). For wheat and barley, sequences of model organisms including especially rice (Matsumoto *et al*. 2005) and *Brachypodium* (Vogel *et al.* 2010) were of great help. However, the synteny on micro-level is not always perfect (Gale and Devos 1998). The high number of wheat and barley genes mapping to non-syntenic regions in model organisms is partly accounted for by TE-driven gene movement generating numerous pseudogenes (Wicker *et al.* 2011). Sometimes, even macro-collinearity can be disrupted

(Bossolini *et al.* 2007; Foote *et al.* 1997; Gallego *et al.* 1998; Lu *et al.* 2006). Despite its limitations, a synteny-based approach was upgraded into a linear gene order model provided by the so-called GenomeZipper (Hernandez *et al.* 2012; Mayer *et al.* 2011) which enabled more efficient marker development (e.g. Dawson *et al.* 2015; Lüpken *et al.* 2014; Randhawa *et al.* 2014; Staňková *et al.* 2015; Tucker *et al.* 2017).

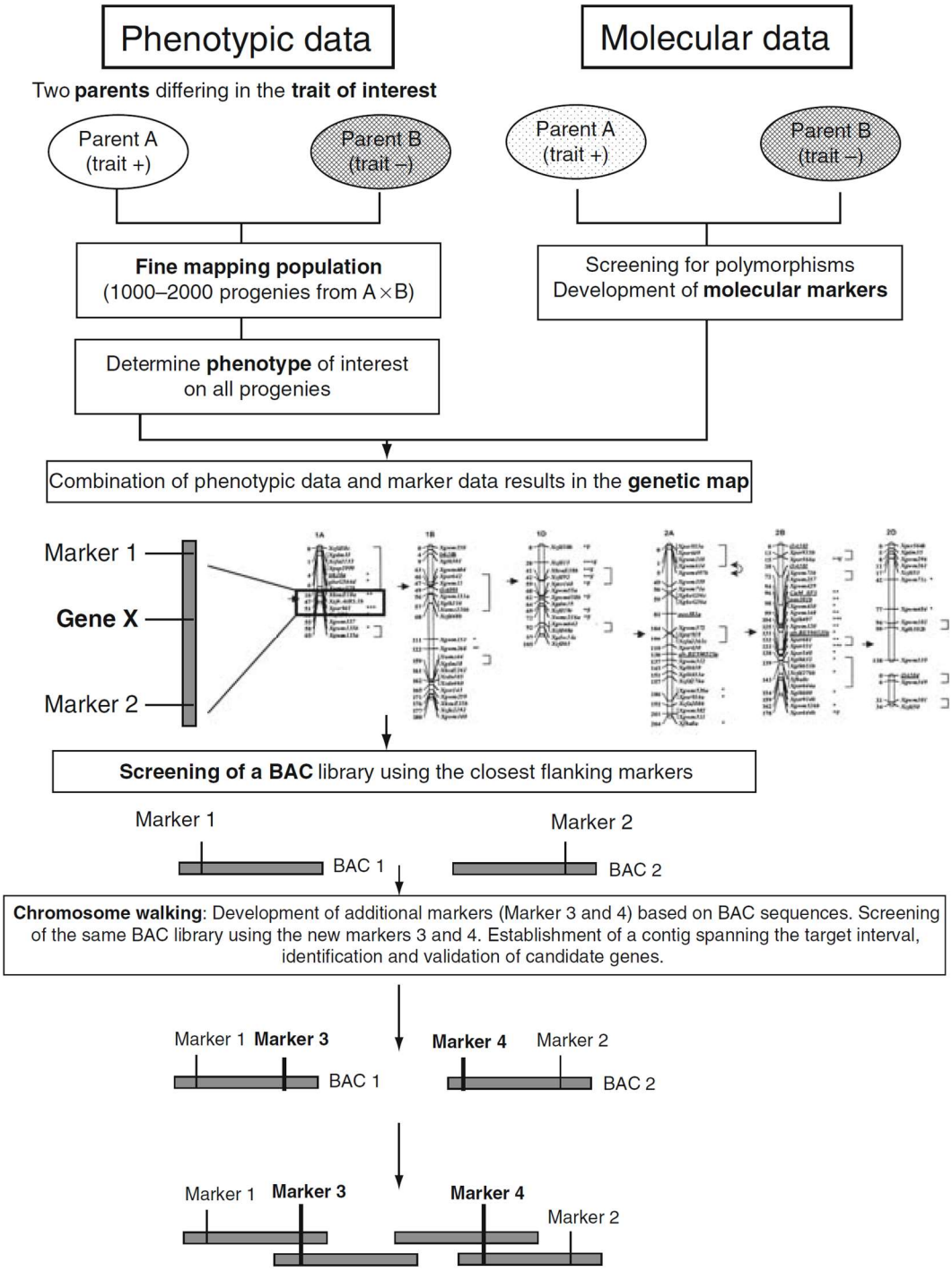


Fig. 9 Schematic demonstration of a basic strategy of the map-based cloning approach. Reproduced from Krattinger *et al.* (2009b).

A substantial hindrance to map-based cloning originates from location of target gene in low-recombining region. Lack of recombination can be caused either by centromeric (or pericentromeric) position (as shown in Lüpken *et al.* 2014; Qi and Gill 2001), presumed structural rearrangements (Hanemann *et al.* 2009) or by low homology resulting from alien translocation (e.g. Cao *et al.* 2011). In a special case of suppressed recombination between wheat and alien chromatin, a possible solution consists in employing plant material with homozygous deletion of the *Ph1* locus, a major locus controlling chromosome pairing in polyploid wheat (Riley and Chapman 1958) which has been localized within wheat chromosome 5B (Griffiths *et al.* 2006). Due to the *ph1* mutation, induction of homoeologous recombination has been accomplished (Bonafede *et al.* 2007; Millet *et al.* 2014; Mullan *et al.* 2009).

Another major obstacle in map-based cloning is represented by uneven coverage of genomes by BAC libraries which results in gaps. Uncomplete physical map of the target locus is therefore not a rare problem (e.g. Hanemann *et al.* 2009; Liang *et al.* 2016; Vu *et al.* 2010). A possible solution consists in additional BAC library employed to bridge the gap (as demonstrated by Yahiaoui *et al.* 2004). Besides an absence of target region in a BAC library, chromosome walking might be complicated or even stopped by large regions composed of TEs (Griffiths *et al.* 2006). However, this problem can be addressed by development of specific markers spanning TE junctions (Paux *et al.* 2006). Recently, an alternative and faster approach to obtain a sequence of target locus has been developed. Instead of using BAC libraries, it consists in targeted chromosome-based cloning via long-range linkage. This approach couples chromosome flow sorting with Chicago long-range linkage (Putnam *et al.* 2016), a new method to assemble shotgun reads (Thind *et al.* 2017). Resulting high-quality *de novo* chromosome assembly of any cultivar represents an excellent platform for gene cloning achieved at unprecedented speed.

The final step of map-based cloning is represented by functional validation of candidate genes. Standard approaches include mutagenesis followed by screening for loss-of-function mutants, transient expression and transformation of candidate gene into the background of organism with contrasting phenotype. Chemical mutagenesis for creating a TILLING population can be performed by Ethyl Methanesulfonate (EMS;

Faris *et al.* 2010; Feuillet *et al.* 2003; Fu *et al.* 2009; Mago *et al.* 2004; Mascher *et al.* 2014; Moore *et al.* 2015; Saintenac *et al.* 2013) or by sodium azide (Krattinger *et al.* 2009a) to obtain point mutations or small deletions. However, this approach is unable to produce mutants in loci with a complex structure as documented for the case of *Ph1* by Griffiths *et al.* (2006). Other options include mutagenesis induced by ionizing radiation which can cause large-scale deletions or chromosomal rearrangements (Mago *et al.* 2004; Moore *et al.* 2015), UV-mutagenesis (Parlange *et al.* 2015) or fast-neutron bombardment (Faris *et al.* 2010). A transgenic complementation experiment performed by particle bombardment (Feuillet *et al.* 2003; Fu *et al.* 2009; Huang *et al.* 2003; Liu *et al.* 2014; Saintenac *et al.* 2013) or *Agrobacterium tumefaciens*-mediated transformation (Horvath *et al.* 2003; Mago *et al.* 2015; Periyannan *et al.* 2013) represent additional level of evidence. Other standardly employed options include ectopic expression in model organisms (Bourras *et al.* 2015; Praz *et al.* 2017), virus-induced gene silencing (Liu *et al.* 2014; Periyannan *et al.* 2013) or RNA Interference (RNAi)-induced gene silencing (Rawat *et al.* 2016). Furthermore, indirect evidence can be provided by association mapping performed on large number of accessions differing in studied phenotype (Faris *et al.* 2010). Apart from these traditional techniques, availability of novel approaches based on modified, sequence-specific nucleases (Belhaj *et al.* 2015; Sprink *et al.* 2015) holds promise for efficient and targeted genome editing to verify function of candidate genes and facilitate their practical application in the future.

2.4.2 Mutagenesis-based approaches

Besides reverse genetics-based strategies employing mutagenesis for validation of candidate gene function, approaches based on forward genetics can be applicable as well. As shown in Gill *et al.* (2016), mutagenesis and selection of mutants with desired phenotype can become a basis for genetic mapping leading to identification of additional signaling components in certain pathway via map-based cloning described in the previous subchapter. However, even this strategy cannot surmount the biological limitations caused by lack of recombination. Recently, a completely novel approach called "mutational genomics" which is independent of recombination has appeared

(Wulff and Moscou 2014). The first, mutagenesis stage of this approach is identical with classical forward genetics. In the next step, whole genomes or their parts of several independent mutants are sequenced and candidate genes are identified based on co-localization of multiple mutations. Steuernagel *et al.* (2016) presents "MutRenSeq" (Mutagenesis and *R* gene Enrichment Sequencing), a method combining chemical mutagenesis with exome capture and sequencing for rapid *R* gene cloning. A necessary premise for this approach is the fact that most *R* genes encode proteins with NB-LRR domains and it is therefore not useful if we seek a gene of unknown structure. However, another approach called "MutChromSeq" (Sánchez-Martín *et al.* 2016) can identify any gene and it has been successfully employed in both wheat and barley. Instead of exome capture, the complexity reduction strategy of MutChromSeq consists in sequencing flow-sorted mutant chromosomes (Fig. 10). A limitation of this approach resides in its requirements. Besides the necessary knowledge of which chromosome contains the target gene, the species needs to be amenable to mutagenesis and very importantly, the gene must be associated with a clear phenotype.

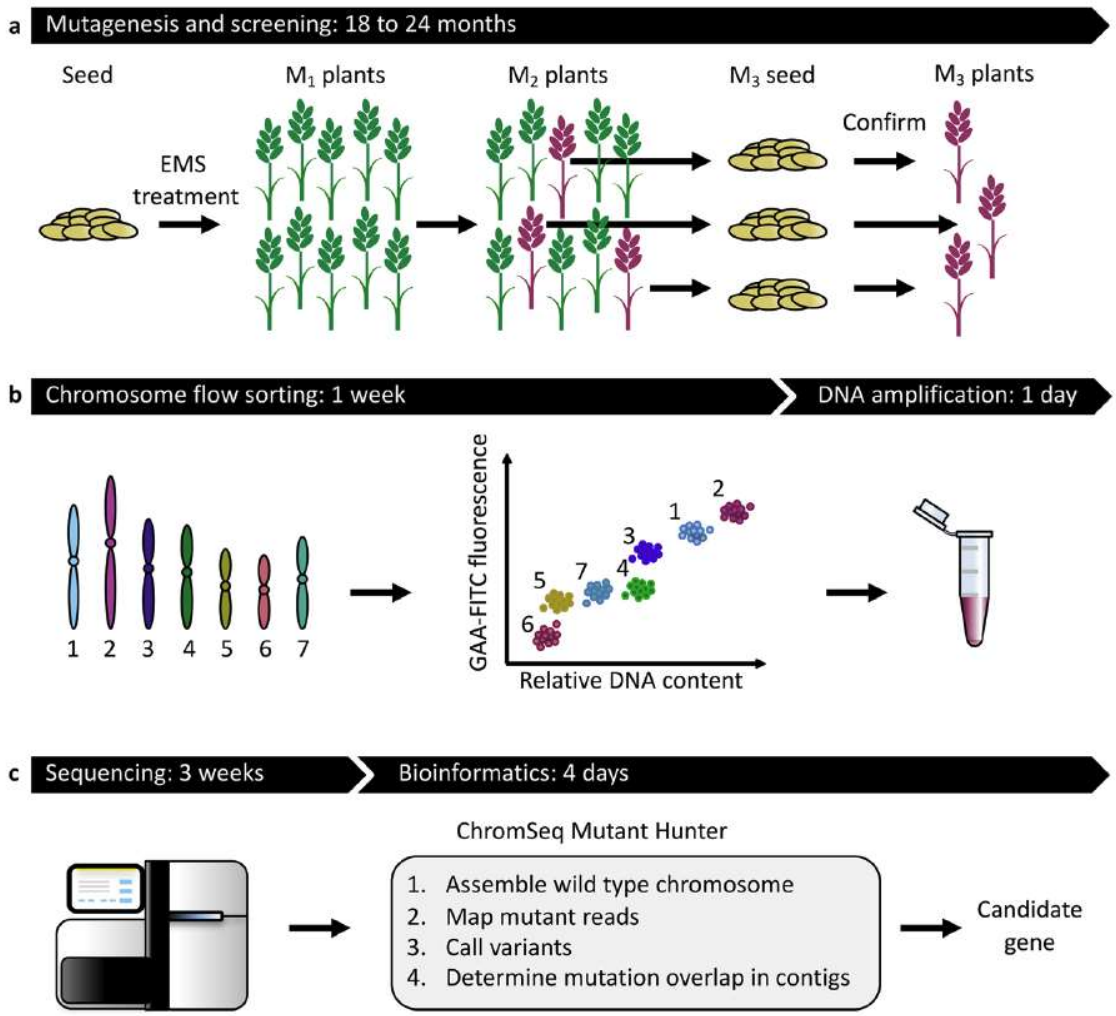


Fig. 10 Workflow of gene isolation in plants by the MutChromSeq strategy. (a) Mutagenesis of seeds (EMS = ethyl methanesulfonate), screening for homozygous mutants in M_2 generation, progeny test in M_3 generation. (b) Flow sorting of target chromosome (GAA – microsatellite probe, FITC – fluorescein isothiocyanate), DNA amplification. (c) Sequencing of wild type and mutant chromosomes, sequence comparison, candidate gene identification. Reproduced from Sánchez-Martín *et al.* (2016).

2.4.3 Genome-wide association studies

The origin of Genome-Wide Association Study (GWAS) dates back to over a decade ago and it is tightly linked with human research. This powerful strategy enabling identification of causal genetic variants for many diseases was triggered by the HapMap project (Int HapMap Consortium 2003) along with great progress in DNA analysis technology which facilitated large-scale genotyping experiments (Uitterlinden 2016). Instead of a typical human GWAS based on a case-control design, GWAS in crops

detects natural variation underlying complex traits in a population of diverse varieties or landraces (Fig. 11; Huang and Han 2014). Like traditional genetic mapping, this approach is based on linkage disequilibrium. However, GWAS is able to address restraints including insufficient resolution or limited genetic diversity within segregating biparental populations because it benefits from polymorphisms and recombination events accumulated within large number of individuals over many generations (Bartoli and Roux 2017; Boyd *et al.* 2013).

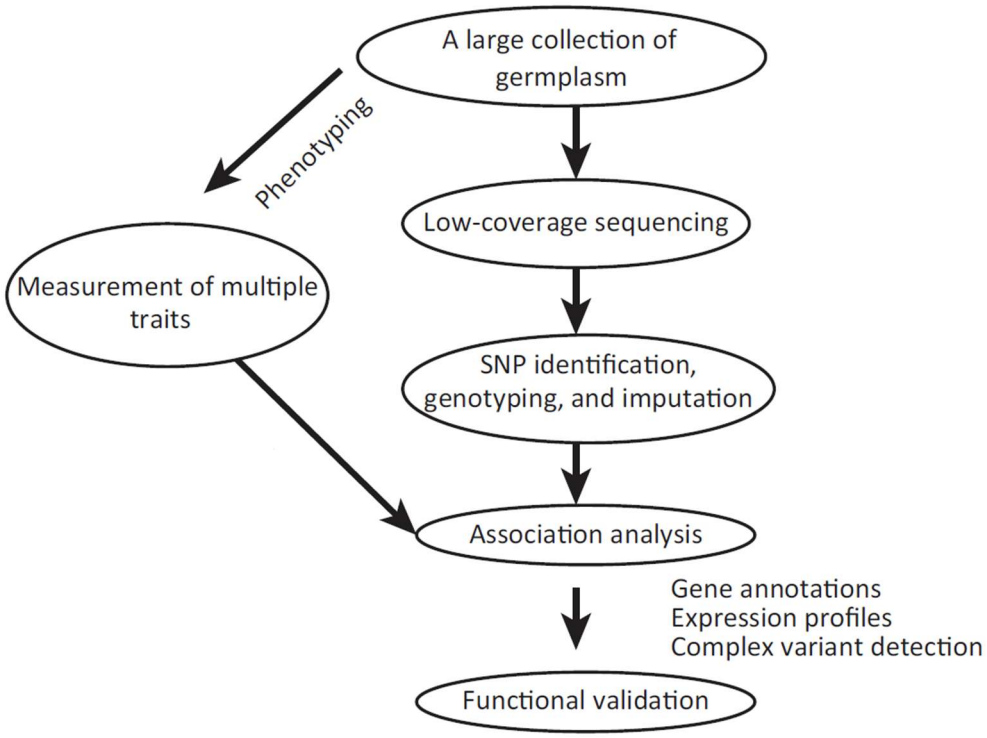


Fig. 11 General scheme of a sequence-based genome-wide association study. Reproduced from Huang *et al.* (2013).

Until recently, applicability of the GWAS approach has been technically difficult and thus limited in organisms with large genomes (Wulff and Moscou 2014). Indeed, map-based cloning has been predominating approach used for wheat and barley although the situation might gradually change. Due to the completion of reference genome sequences for many crops and the relative availability of resequencing whole genomes, dissecting the genetic basis of important traits in these crop species is

becoming a more feasible task (Huang and Han 2014). Moreover, in some crop species including wheat and barley, resequencing can be partially substituted by available high-throughput genotyping platforms (Gupta *et al.* 2013; Rasheed *et al.* 2017). GWAS analyses performed on barley to date analyzed a sufficient number of accessions, however, the number of markers ranged only from several hundreds to several thousands (e.g. Burlakoti *et al.* 2017; Fan *et al.* 2016; Pasam *et al.* 2012; Reinert *et al.* 2016; Wang *et al.* 2012). Considering the genome size of barley, such a marker density provides limited coverage and these studies therefore often resulted in detecting QTL rather than in identification of candidate gene/s underlying studied phenotypes. Consequently, implementing GWAS in wheat has been even more challenging (Huang and Han 2014). However, recent progress and available resources launched release of many studies performed even on hexaploid bread wheat using e.g. 90k iSELECT Illumina chip or 35k Affymetrix chip (Alomari *et al.* 2017; Liu N *et al.* 2017; Sukumaran *et al.* 2016; Valluru *et al.* 2017), DArT-seq markers (Kaur *et al.* 2017; Liu YX *et al.* 2017) or combination of DArT-seq and Axion Wheat 660 SNP Affymetrix array (Zhou *et al.* 2017). A genotyping-by-sequencing approach employed by Arruda *et al.* (2016) made use of wheat cv. Chinese Spring chromosome survey sequences (IWGSC 2014) as a pseudo-reference genome to align reads and identify SNP markers. Although the number of markers in above-mentioned studies reached only tens of thousands, some of them resulted even in identification of putative candidate genes (Alomari *et al.* 2017; Kaur *et al.* 2017; Liu YX *et al.* 2017; Sukumaran *et al.* 2016).

Importantly for this thesis, Liu N *et al.* (2017) carried out GWAS in a set of U.S. winter wheats to identify genomic regions responsible for powdery mildew resistance and using a total of 21,600 SNP markers, this study revealed several new QTL. To discover novel sources of powdery mildew resistance in barley, Ames *et al.* (2015) mined a collection of 314 wild accessions and performed a GWAS using 10,508 DArT-Seq, DArT and SNP markers. 15 independent QTL associated with resistance were identified and seven of them were found to be novel. Lately, Bengtsson *et al.* (2017) employed GWAS in a Nordic spring barley panel of 169 lines and cultivars which were genotyped by 5,556 markers of the barley iSelect SNP chip based on the Illumina Infinium 9K assay. One novel powdery mildew resistance QTL was detected and the responsible gene was

predicted to be the 26S proteasome regulatory subunit RPN1 which had been previously showed to be involved in innate immunity in *A. thaliana* (Yao *et al.* 2012).

Altogether, GWAS on diverse wheat or barley accessions is a powerful tool to elucidate genomic basis of various traits including disease resistance, however, higher density of markers is necessary to achieve higher resolution. In case chromosome location of the target locus is known, another possibility is open to increase efficiency and save resources. Coupling chromosome flow-sorting with low-coverage resequencing and subsequent mapping of reads on available high-quality reference sequences has a potential to provide sufficient marker coverage and facilitate direct identification of genes of interest.

GWAS in organisms with no SNP genotyping chips available is completely dependent on sequencing (Power *et al.* 2016). Although the genomes of powdery mildew fungi show a substantial size expansion in comparison with their relatives (Spanu *et al.* 2010), resequencing their 100-200 Mbp genomes is still significantly less costly than resequencing genomes of cereal crops which are dozens of times bigger. Despite this advantage, the use of GWAS in *B. graminis* is still in the beginnings. Praz *et al.* (2017) carried out both map-based cloning and GWAS to identify *Bgt* avirulence effector *AvrPm2*. For GWAS, genomic sequences of 60 isolates were included and over 400,000 SNPs were tested. As a result, the best correlated SNP was localized 50 kbp from *AvrPm2*. Due to the fact that the ~10 kbp region surrounding this effector was found to be deleted in virulent isolates, it was not possible to capture directly the target gene. Since multiple gene duplications and deletions are a common feature among candidate effector genes of *B. graminis* lineages (Menardo *et al.* 2017b), this phenomenon can be expected when performing effector-seeking GWAS. Instead of GWAS on genomic data, Lu *et al.* (2016) applied a transcriptome-wide association study on 16 *Bgh* isolates with *Avr* genes corresponding to different *Mla* alleles. Over 60,000 SNPs were employed for the analysis and this approach enabled them to identify two avirulence effectors, Avr_{a1} and $Avra_{13}$. While SNPs were identified in Avr_{a1}, *Mla13*-virulent isolates carried a 326-bp insertion in the Avr_{a13} gene which could not be directly detected by the association analysis. These two examples demonstrate a potential of this approach to identify the whole set of *B. graminis* avirulence determinants. Nevertheless, one possible bottleneck

might consist in their subsequent functional validation. Avirulence function can be tested by transient *Agrobacterium*-mediated transformation of *Nicotiana benthamiana* (Bourras *et al.* 2015; Praz *et al.* 2017) or barley (Lu *et al.* 2016) or by particle bombardment assays (Bourras *et al.* 2015). Unfortunately, a direct functional evidence for contribution of *Avrs* to virulence is still hard to obtain due to lacking transformation protocols for mildews (Ellis *et al.* 2009; Nowara *et al.* 2010). However, the RNAi-based approach of host-induced gene silencing has been described in *B. graminis* (Nowara *et al.* 2010).

In general, standard methods of functional validation for genes identified by GWAS, map-based cloning or any other approach do not differ and they are summarized in subchapter dealing with map-based cloning.

3 OBJECTIVES

This thesis focuses on plant-pathogen interactions with respect to the pathosystem constituted by important cereal crops, wheat and barley, and a pathogenic fungus *Blumeria graminis*. It is supposed to provide a brief summary on these organisms, plant immunity in general, types of resistance responses and involved genetic determinants uncovered to date along with basic aspects of strategies used for their identification. The experimental work follows two lines which are complementary to each other – the plant line and the pathogen line. Based on this basic structure, objectives can be divided into two separate parts:

I. Employing a map-based cloning approach in hexaploid bread wheat (*Triticum aestivum*) to identify gene/s conferring an improved, broad-spectrum powdery mildew resistance which had been previously mapped to an introgression from the tetraploid *T. militinae* located on the long arm of chromosome 4A (Jakobson *et al*. 2006; Jakobson *et al*. 2012). This task requires performing various methods of genetics and genomics including recombination mapping using different mapping populations, physical mapping supported by chromosome-specific BAC libraries, sequencing and sequence analysis, development of molecular markers and analysis of RNA expression.

II. Evaluating genetic diversity within a set of *Blumeria graminis* f. sp. *hordei* isolates and its subsequent application for identification of candidate pathogen effectors involved in avirulence interactions. First, a strategy based on designing new molecular markers was selected to obtain information on pathogen diversity within a single *forma specialis*. The aim of the next step consisted in whole genome resequencing of isolates followed by bioinformatic analysis of sequence data to obtain a large set of SNP markers and finally employing this set for a genome-wide association study with available phenotype data.

4 RESULTS

4.1 Publications

4.1.1 Divergence between bread wheat and *Triticum militinae* in the powdery mildew resistance *QPm.tut-4A* locus and its implications for cloning of the resistance gene

Janáková E, Jakobson I, Peusha H, Abrouk M, Škopová M, Šimková H, Šafář J, Vrána J, Järve K, Doležel J, Valárik M

Theoretical and Applied Genetics: accepted for publication

https://doi.org/10.1007/s00122-018-3259-3

IF 3.930 (2017/2018)

APPENDIX I

ABSTRACT

A segment of *Triticum militinae* chromosome 7G harbors a gene(s) conferring powdery mildew resistance which is effective at both the seedling and the adult plant stages when transferred into bread wheat (*T. aestivum*). The introgressed segment replaces a piece of wheat chromosome arm 4AL. An analysis of segregating materials generated to positionally clone the gene highlighted that in a plant heterozygous for the introgression segment, only limited recombination occurs between the introgressed region and bread wheat 4A. Nevertheless, 75 genetic markers were successfully placed within the region, thereby confining the gene to a 0.012 cM window along the 4AL arm. In a background lacking the *Ph1* locus, the localized rate of recombination was raised 33-fold, enabling the reduction in the length of the region containing the resistance gene to a 480 kbp stretch harboring 12 predicted genes. The substituted segment in the reference sequence of bread wheat cv. Chinese Spring is longer (640 kbp) and harbors 16 genes. A comparison of the segments' sequences revealed a high degree of divergence with respect to both their gene content and nucleotide sequence. Of the 12 *T. militinae* genes, only four have a homolog in cv. Chinese Spring. Possible candidate genes for the resistance have been identified based on function predicted from their sequence.

4.1.2 Genetic diversity of *Blumeria graminis* f. sp. *hordei* in Central Europe and its comparison with Australian population

Komínková E, Dreiseitl A, Malečková E, Doležel J, Valárik M

Plos One (2016) 11: e0167099

DOI: 10.1371/journal.pone.0167099

IF 2.806

APPENDIX II

ABSTRACT

Population surveys of *Blumeria graminis* f. sp. *hordei* (*Bgh*), a causal agent of more than 50% of barley fungal infections in the Czech Republic, have been traditionally based on virulence tests, at times supplemented with non-specific Restriction fragment length polymorphism or Random amplified polymorphic DNA markers. A genomic sequence of *Bgh*, which has become available recently, enables identification of potential markers suitable for population genetics studies. Two major strategies relying on transposable elements and microsatellites were employed in this work to develop a set of Repeat junction markers, Single sequence repeat and Single nucleotide polymorphism markers. A resolution power of the new panel of markers comprising 33 polymorphisms was demonstrated by a phylogenetic analysis of 158 *Bgh* isolates. A core set of 97 Czech isolates was compared to a set 50 Australian isolates on the background of 11 diverse isolates collected throughout the world. 73.2% of Czech isolates were found to be genetically unique. An extreme diversity of this collection was in strong contrast with the uniformity of the Australian one. This work paves the way for studies of population structure and dynamics based on genetic variability among different *Bgh* isolates originating from geographically limited regions.

4.1.3 The *in silico* identification and characterization of a bread wheat/*Triticum militinae* introgression line

Abrouk M, Balcárková B, Šimková H, Komínková E, Martis M, Jakobson I, Timofejeva L, Rey E, Vrána J, Kilian A, Järve K, Doležel J, Valárik M

Plant Biotechnology Journal (2017) 15:249-256

DOI: 10.1111/pbi.12610

IF 6.305

APPENDIX III

ABSTRACT

The capacity of the bread wheat (*Triticum aestivum*) genome to tolerate introgression from related genomes can be exploited for wheat improvement. A resistance to powdery mildew expressed by a derivative of the cross bread wheat cv. Tähti × *T. militinae* (*Tm*) is known to be due to the incorporation of a *Tm* segment into the long arm of chromosome 4A. Here, a newly developed *in silico* method termed RICh (rearrangement identification and characterization) has been applied to characterize the introgression. A virtual gene order, assembled using the GenomeZipper approach, was obtained for the native copy of chromosome 4A; it incorporated 570 4A DArTseq markers to produce a zipper comprising 2,132 loci. A comparison between the native and introgressed forms of the 4AL chromosome arm showed that the introgressed region is located at the distal part of the arm. The *Tm* segment, derived from chromosome 7G, harbors 131 homoeologs out of the 357 genes present on the corresponding region of Chinese Spring 4AL. The estimated number of *Tm* genes transferred along with the disease resistance gene was 169. Characterizing the introgression's position, gene content and internal gene order should facilitate not only gene isolation, but may also be informative with respect to chromatin structure and behavior studies.

4.1.4 The utility of flow sorting to identify chromosomes carrying a single copy transgene in wheat

Cápal P, Endo TR, Vrána J, Kubaláková M, Karafiátová M, Komínková E, Mora-Ramírez I, Weschke W, Doležel J

Plant Methods (2016) 12: 24

DOI: 10.1186/s13007-016-0124-8

IF 3.51

APPENDIX IV

ABSTRACT

Background: Identification of transgene insertion sites in plant genomes has practical implications for crop breeding and is a stepping stone to analyze transgene function. However, single copy sequences are not always easy to localize in large plant genomes by standard approaches.

Results: We employed flow cytometric chromosome sorting to determine chromosomal location of barley sucrose transporter construct in three transgenic lines of common wheat. Flow-sorted chromosomes were used as template for PCR and fluorescence in situ hybridization to identify chromosomes with transgenes. The chromosomes carrying the transgenes were then confirmed by PCR using DNA amplified from single flow-sorted chromosomes as template.

Conclusions: Insertion sites of the transgene were unambiguously localized to chromosomes 4A, 7A and 5D in three wheat transgenic lines. The procedure presented in this study is applicable for localization of any single-copy sequence not only in wheat, but in any plant species where suspension of intact mitotic chromosomes suitable for flow cytometric sorting can be prepared.

4.2 Unpublished research

4.2.1 Utilizing long-term and extensive virulence monitoring data as a basis for *in silico* identification of candidate avirulence effectors of barley powdery mildew fungus

Janáková E, Menardo F, Dreiseitl A, Bourras S, Doležel J, Valárik M, Keller B

APPENDIX V

ABSTRACT

Sustainable control of plant pathogens based on disease resistance is highly desirable to reduce yield losses of cultivated crops. However, employing effective and durable resistance requires prior identification of plant and pathogen interacting components and understanding their interplay. Nowadays, affordable whole genome sequencing provides a basis for dissecting core genes involved in defense responses. Annotation of *Blumeria graminis* f. sp. *hordei* genome revealed a presence of nearly five hundred genes encoding putative candidate effectors. In an effort to predict effectors with potential avirulence function, we conducted a genome-wide association study on a set of 90 isolates employing 453,114 high-confidence SNP markers and 35 distinct phenotypes gathered over decades of virulence monitoring. Setting stringent criteria yielded ten highly significant associations and closer inspection detected a candidate effector for seven of them including *Mla1*, *Mla3*, *Mla10*, *Mla22*, *Mlg*, *Mlk1* and *MlRu2*. In three cases, the effector was suggested to act as a suppressor of avirulence. A single candidate with presumably dual function designated Avr_{a10}/Svr_{a22} indicated existence of complex interactions influencing avirulence in barley-powdery mildew fungus pathosystem. Complete or partial deletion of effector gene frequently observed among identified candidates exposed a potential pitfall of SNP-based association analysis. This work opens new perspectives for high-throughput identification of avirulence effectors in barley powdery mildew fungus.

4.3 List of conference abstracts

- Komínková E, Jakobson I, Peusha H, Abrouk M, Järve K, Doležel J, Valárik M (2017) Positinal cloning of powdery mildew resistance gene introgressed to bread wheat from *Triticum militinae*. In: Buerstmayr H, Lang-Mladek C, Steiner B, Michel S, Buerstmayr M, Lemmens M, Vollmann J, Grausgruber H (Eds.) Proceedings of the 13th International Wheat Genetics Symposium. Tulln, Austria; April 23-28, 2017; BOKU – University of Natural Resources and Life Sciences, Vienna; pp. 247 (APPENDIX VI)

- Komínková E, Malečková E, Vanžurová H, Dreiseitl A, Doležel J, Valárik M (2015) Molecular description of *Blumeria graminis* f. sp. *hordei* isolates. In: 14th International Cereal Rusts and Powdery Mildews Conference. Helsingør, Denmark; July 5-8, 2015; pp. 92 (APPENDIX VII)

- Komínková E, Klocová B, Abrouk M, Jakobson I, Peusha H, Šimková H, Šafář J, Järve K, Doležel J, Valárik M (2015) Positional cloning of powdery mildew resistance gene introgressed to bread wheat from *Triticum militinae*. In: Plant Biotechnology: Green for Good III. Olomouc, Czech Republic; June 15-18, 2015; pp. 63 (APPENDIX VIII)

5 GENERAL DISCUSSION

The first part of experimental work performed within this dissertation reports a significant progress that was made towards map-based cloning of the powdery mildew resistance QTL designated *QPm.tut-4A* and localized on 4AL chromosome arm of bread wheat. Finally, the underlying region originally mapped to 2.5-cM distance (Jakobson *et al.* 2012) was narrowed down to a 480 kb-segment of *T. militinae* origin harboring 12 predicted protein-coding gene. This accomplishment was substantially hindered by the fact that the genomic region introgressed from a member of bread wheat secondary gene pool exhibited a considerable sequence divergence accompanied even by partial loss of synteny in the region of interest. The obvious consequence for genetic mapping consists in suppression of recombination resulting in distortion of genetic distances and painful physical mapping. However, alien gene transfer from related species has a great potential for improvement of current elite bread wheat cultivars. In this project, we show that employing a mapping population with homozygous deletion of the *Ph1* locus (Riley and Chapman 1958) can significantly eliminate introgression-related problems by increasing the recombination frequency 33-fold. To our knowledge, this value represents the first available quantification of the *Ph1* locus effect. Another contribution of this research lies in presenting a highly efficient and fast marker development strategy based on survey sequences of flow sorted chromosomes and the virtual gene order approach called "GenomeZipper" (Mayer *et al.* 2011). The ease, specificity and success rate of this strategy resulted in its employment for developing chromosome arm-specific wheat markers in different projects. The reported progress in the *QPm.tut-4A*-cloning project provides a solid starting point for functional validation of identified candidate genes. The most relevant candidates are represented by a cluster of genes encoding putative CC-NB-(LRR) proteins. This rather unexpected finding suggests an intriguing eventuality of typical major *R* genes being involved in partial, broad-spectrum resistance. Such a scenario highlights the fragile border between different resistance types.

In the second project, attention was focused on the pathogenic fungus *Blumeria graminis* f. sp. *hordei* causing powdery mildew disease of barley. First, a set of molecular

markers was designed and tested on a large set of isolates representing two distinct populations. The results confirmed that repetitive sequences constituting a majority of *B. graminis* genome are a rich source of polymorphism and thus of practically unlimited number of easily obtainable markers. The second important implication was extremely high genetic diversity detected within population originating from geographically small area. This knowledge opened the door for using available isolates as a basis of genome-wide association study where diversity is a prerequisite. Having extensive data on virulence patterns of individual pathogen isolates at our disposal, an exciting opportunity to unveil avirulence determinants corresponding to high number of *R* genes simultaneously appeared. Indeed, ten significant associations with seven candidates for effectors involved in avirulence interactions represented a considerable step forward. Up to now, a maximum of two new avirulence effectors of this pathogen have been reported at the same time and their total number has been still poor. Moreover, detailed analysis of the results indicated that even more high-confidence associations could be potentially mined out of our data when proposed measures are adopted to optimize the analysis. Apart from the relatively high number of identified putative avirulence effectors, a striking piece of knowledge was provided by the discovery that almost half of them seem to act as suppressors of avirulence interaction. So far, the first one and the only suppressor within this species was identified in wheat powdery mildew fungus and it has been therefore impossible to estimate the extent of occurrence of such phenomenon in this pathogen. Based on our results, it seems that others than canonical gene-for-gene interactions are more widespread than anticipated. Naturally, these exciting results need to be experimentally validated before any solid conclusion can be drawn. However, especially candidate suppressors represent a challenge in this regard since no appropriate functional assay has been reported so far. Prospectively, once a complete repertoire of pathogen avirulence effectors is identified and validated, it may become a part of strategy aiming at durable disease resistance. Even though such effectors are involved in the process of race-specific, major gene resistance and its short durability is a many times proved fact, monitoring avirulence effectors' complement present in pathogen populations can support rational and long-life deployment of *R* genes and their pyramiding.

Altogether, both presented projects were seeking sustainable disease resistance as their ultimate goal. In spite of being focused on phylogenetically very distant organisms, their long-term coevolution resulted in a kind of interconnection causing the necessity of studying both of them at the same time to obtain a complete picture. Conducting research on each of this host-pathogen pair brought new pieces to the mosaic of our knowledge which will be hopefully utilized to achieve sustainable and high-yielding agriculture with minimized impact on the environment.

6 CONCLUSIONS

If we are supposed to keep pace with growing demand of food crops in the conditions of changing climate, mining crops' diversity present in their wild relatives is a must in spite of attached complications. Overcoming some of these obstacles in the process of map-based cloning of alien powdery mildew resistance QTL in bread wheat formed one of the key objectives of research realized within this dissertation. The outputs reported in original publication (Janáková *et al.*, accepted) demonstrate identification of 480 kb-physical region of the introgressed locus harboring 12 candidate genes potentially involved in the *QPm.tut-4A* resistance. Moreover, highly specific DNA markers developed to saturate the locus of interest facilitated identification of introgression origin within the donor genome. As discovered in Abrouk *et al.* (2017), the 4AL-segment in bread wheat introgression line was originally translocated from *T. militinae* chromosome 7G. Finally, the described marker development approach was applied also in original research of Cápal *et al.* (2016) thus proving its versatility. Generally, strategies employed in this project represent a promising support to map-based cloning of other genes introgressed to bread wheat.

To broaden our understanding of the bipartite process known as "disease resistance", the second part of presented research dealt with the fungal pathogen *Blumeria graminis* f. sp. *hordei*. First, genetic diversity within a large set of isolates was characterized using newly developed molecular markers and results were published in Komínková *et al.* (2016). The subsequent genome-wide association analysis aiming at identification of pathogen avirulence effectors (Janáková, unpublished) brought a progress to our understanding of plant-pathogen interactions in this pathosystem and also a potentially model strategy to be employed in other pathosystems. In conclusion, new pieces of knowledge obtained within both major parts of this dissertation contribute to the problematics of sustainable and durable protection of cereals from the powdery mildew fungus.

7 REFERENCES

Abrouk M, Balcárková B, Šimková H, Komínková E, Martis M, Jakobson I, Timofejeva L, Rey E, Vrána J, Kilian A, Järve K, Doležel J, Valárik M (2017): The *in silico* identification and characterization of a bread wheat/*Triticum militinae* introgression line. Plant Biotechnol J 15:249-256

Acevedo-Garcia J, Spencer D, Thieron H, Reinstädler A, Hammond-Kosack K, Phillips AL, Panstruga R (2017) *mlo*-based powdery mildew resistance in hexaploid bread wheat generated by a non-transgenic TILLING approach. Plant Biotechnol J 15:367-378

Agarwal M, Shrivastava N, Padh H (2008) Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Rep 27:617-631

Aghnoum R, Niks RE (2010) Specificity and levels of nonhost resistance to nonadapted *Blumeria graminis* forms in barley. New Phytol 185:275-284

Aguilar GB, Pedersen C, Thordal-Christensen H (2015) Identification of eight effector candidate genes involved in early aggressiveness of the barley powdery mildew fungus. Plant Pathol 65:953-958

Ahmed AA, Pedersen C, Schultz-Larsen T, Kwaaitaal M, Jørgensen HJL, Thordal-Christensen H (2015) Barley powdery mildew effector candidate CSEP inhibits chaperone activity of small heat shock protein. Plant Physiol 168:321-333

Ahmed AA, Pedersen C, Thordal-Christensen H (2016) The barley powdery mildew effector candidates CSEP0081 and CSEP0254 promote fungal infection success. Plos One 11:e0157586

Alomari DZ, Eggert K, von Wirén N, Pillen K, Röder MS (2017) Genome-wide association study of calcium accumulation in grains of European wheat cultivars. Front Plant Sci 8:1797

Ames N, Dreiseitl A, Steffenson BJ (2015) Mining wild barley for powdery mildew resistance. Plant Pathol 64:1396-1406

Arruda MP, Brown P, Brown-Guedira G, Krill AM, Thurber C, Merrill KR, Foresman BJ, Kolb FL (2016) Genome-wide association mapping of fusarium head blight resistance in wheat using genotyping-by-sequencing. Plant Genome-US 9. DOI: 10.3835/plantgenome2015.04.0028

Atienza SG, Jafary H, Niks RE (2004) Accumulation of genes for susceptibility to rust fungi for which barley is nearly a nonhost results in two barley lines with extreme multiple susceptibility. Planta 220:71-79

Awika JM (2011) Major cereal grains production and use around the world. In: Awika JM, Piironen V, Bean S (eds) Advances in cereal science: implications to food processing and health promotion. American Chemical Society, Washington, DC, USA, pp 1-13

Baik BK, Ullrich SE (2008) Barley for food: Characteristics, improvement and renewed interest. J Cereal Sci 48:233-242

Barabaschi D, Tondelli A, Desiderio F, Volante A, Vaccino P, Vale G, Cattivelli L (2016) Next generation breeding. Plant Sci 242:3-13

Bartoli C, Roux F (2017) Genome-wide association studies in plant pathosystems: toward an ecological genomics approach. Front Plant Sci 8:763

Belhaj K, Chaparro-Garcia A, Kamoun S, Patron NJ, Nekrasov V (2015) Editing plant genomes with CRISPR/Cas9. Curr Opin Biotech 32: 76-84

Bengtsson T, Åhman I, Manninen O, Reitan L, Christerson T, Jensen JD, Krusell L, Jahoor A, Orabi J (2017) A novel QTL for powdery mildew resistance in Nordic spring barley (*Hordeum vulgare* L. ssp. *vulgare*) revealed by genome-wide association study. Front Plant Sci 8:1954

Bennett MD, Smith JB (1976) Nuclear DNA amounts in angiosperms. Philos T Roy Soc B 274:227-274

Bhullar NK, Street K, Mackay M, Yahiaoui N, Keller B (2009) Unlocking wheat genetic resources for the molecular identification of previously undescribed functional alleles at the *Pm3* resistance locus. P Natl Acad Sci USA 106:9519-9524

Bhullar NK, Zhang ZQ, Wicker T, Keller B (2010) Wheat gene bank accessions as a source of new alleles of the powdery mildew resistance gene *Pm3*: a large scale allele mining project. BMC Plant Biol 10:88

Birch PRJ, Rehmany AP, Pritchard L, Kamoun S, Beynon JL (2006) Trafficking arms: oomycete effectors enter host plant cells. Trends Microbiol 14:8-11

Bonafede M, Kong LR, Tranquilli G, Ohm H, Dubcovsky J (2007) Reduction of a T*riticum monococcum* chromosome segment carrying the softness genes *Pina* and *Pinb* translocated to bread wheat. Crop Sci 47:821-828

Bossolini E, Wicker T, Knobel PA, Keller B (2007) Comparison of orthologous loci from small grass genomes *Brachypodium* and rice: implications for wheat genomics and grass genome annotation. Plant J 49:704-717

Bourras S, McNally KE, Ben-David R, Parlange F, Roffler S, Praz CR, Oberhaensli S, Menardo F, Stirnweis D, Frenkel Z, Schaefer LK, Flückiger S, Treier G, Herren G, Korol AB, Wicker T, Keller B (2015) Multiple avirulence loci and allele-specific effector recognition control the *Pm3* race-specific resistance of wheat to powdery mildew. Plant Cell 37:2991-3012

Bourras S, McNally KE, Müller MC, Wicker T, Keller B (2016) Avirulence genes in cereal powdery mildews: the gene-for-gene hypothesis 2.0. Front Plant Sci 7:241

Boyd LA, Ridout C, O'Sullivan DM, Leach JE, Leung H (2013) Plant-pathogen interactions: disease resistance in modern agriculture. Trends Genet 29:233-240

Bramel P (2017) Wheat genetic resources: global conservation and use for the future. In: Langridge P (ed) Achieving sustainable cultivation of wheat. Volume 1: Breeding, quality traits, pests and diseases. Burleigh Dodds Science Publishing, Sawston, Cambridge, UK, pp 3-24

Brown JKM (1996) The choice of molecular marker methods for population genetic studies of plant pathogens. New Phytol 133:183-195

Brown JKM, Tellier A (2011) Plant-parasite coevolution: bridging the gap between genetics and ecology. Annu Rev Phytopathol 49:345-367

Burlakoti RR, Gyawali S, Chao S, Smith KP, Horsley RD, Cooper B, Muehlbauer GJ, Neate SM (2017) Genome-wide association study of spot form of net blotch resistance in the Upper Midwest barley breeding programs. Phytopathology 107:100-108

Büschges R, Hollricher K, Panstruga R, Simons G, Wolter M, Frijters A, van Daelen R, van der Lee T, Diergaarde P, Groenendijk J, Töpsch S, Vos P, Salamini F, Schulze-Lefert P (1997) Cell 88:695-705

Cao AH, Xing LP, Wang XY, Yang XM, Wang W, Sun YL, Qian C, Ni JL, Chen YP, Liu DJ, Wang X, Chen PD (2011) Serine/threonine kinase *Stpk-V*, a key member of powdery mildew resistance gene *Pm21*, confers powdery mildew resistance in wheat. P Natl Acad Sci USA 108:7727-7732

Cápal P, Endo TR, Vrána J, Kubaláková M, Karafiátová M, Komínková E, Mora-Ramírez I, Weschke W, Doležel J (2016): The utility of flow sorting to identify chromosomes carrying a single copy transgene in wheat. Plant Methods 12:24

Collins NC, Thordal-Christensen H, Lipka V, Bau S, Kombrink E, Qiu JL, Hückelhoven R, Stein M, Freialdenhoven A, Sommerville SC, Schulze-Lefert P (2003) SNARE-protein-mediated disease resistance at the plant cell wall. Nature 425:973-977

Chapman SR, Carter LP (1976) Crop Production: Principles and Practices. W.H. Freeman and Co., San Francisco, USA, pp 247-458

Chaure P, Gurr SJ, Spanu P (2000) Stable transformation of *Erysiphe graminis*, an obligate biotrophic pathogen of barley. Nat Biotechnol 18:205-207

Choulet F, Alberti A, Theil S, Glover N, Barbe V, Daron J, Pingault L, Sourdille P, Couloux A, Paux E *et al.* (2014) Structural and functional partitioning of bread wheat chromosome 3B. Science 345: 1249721

Dawson AM, Ferguson JN, Gardiner M, Green P, Hubbard A, Moscou MJ (2015) Isolation and fine mapping of Rps6: an intermediate host resistance gene in barley to wheat stripe rust. Theor Appl Genet 129:831-843

Devoto A, Piffanelli P, Nilsson I, Wallin E, Panstruga R, von Heijne G, Schulze-Lefert P (1999) Topology, subcellular localization, and sequence diversity of the *Mlo* family in plants. J Biol Chem 274:34993-35004

DeYoung BJ, Innes RW (2006) Plant NBS-LRR proteins in pathogen sensing and host defence. Nat Immunol 7:1243-1249

Dodds PN, Rathjen JP (2010) Plant immunity: towards an integrated view of plant-pathogen interactions. Nat Rev Genet 11:539-548

Doležel J, Greilhuber J, Lucretti S, Meister A, Lysák MA, Nardi L, Obermayer R (1998) Plant genome size estimation by flow cytometry: inter-laboratory comparison. Ann Bot (Lond) 82:17-26

Doležel J, Vrána J, Šafář J, Bartoš J, Kubaláková M, Šimková H (2012) Chromosomes in the flow to simplify genome analysis. Funct Integr Genomic 12:397-416

Dong S, Raffaele S, Kamoun S (2015) The two-speed genomes of filamentous pathogens: waltz with plants. Curr Opin Genet Dev 35:57-65

Douchkov D, Lück S, Johrde A, Nowara D, Himmelbach A, Rajaraman J, Stein N, Sharma R, Kilian B, Schweizer (2014) Discovery of genes affecting resistance of barley to adapted and non-adapted powdery mildew fungi. Genome Biol 15:518

Dreiseitl A (2014) Pathogenic divergence of Central European and Australian populations of *Blumeria graminis* f. sp. *hordei*. Ann Appl Biol 165:364-372

Dreiseitl A (2015) Changes in virulence frequencies and higher fitness of simple pathotypes in the Czech population of *Blumeria graminis* f. sp. *hordei*. Plant Protect Sci 51:67-73

Dreiseitl A (2017) Heterogenity of powdery mildew resistance revealed in accessions of the ICARDA wild barley collection. Front Plant Sci 8:202

Dubcovsky J, Dvorak J (2007) Genome plasticity a key factor in the success of polyploid wheat under domestication. Science 316:1862-1866

Eilam T, Anikster Y, Millet E, Manisterski J, Feldman M (2008) Nuclear DNA amount and genome downsizing in natural and synthetic allopolyploids of the genera *Aegilops* and *Triticum*. Genome 51:616-627

Ellis JG, Rafiqi M, Gan P, Chakrabarti A, Dodds PN (2009) Recent progress in discovery and functional analysis of effector proteins of fungal and oomycete plant pathogens. Curr Opin Plant Biol 12:399-405

Endo TR, Gill BS (1996) The deletion stocks of common wheat. J Hered 87:295-307

Fan Y, Zhou GF, Shabala S, Chen ZH, Cai SG, Li CD, Zhou MX (2016) Genome-wide association study reveals a new QTL for salinity tolerance in barley (*Hordeum vulgare* L.). Front Plant Sci 7:946

FAO, Food and Agriculture Organization of the United Nations (2017) FAOSTAT – Food and agriculture data. http://www.fao.org/faostat/en/#data/. Accessed 18 September 2017

FAO, Food and Agriculture Organization of the United Nations (2015) FAO statistical pocketbook: World food and agriculture. Rome, Food and Agriculture Organization of the United Nations

Faris JD (2013) Wheat domestication: key to agricultural revolutions past and future. In: Tuberosa R, Graner A, Frison E (eds) Genomics of Plant Genetic Resources. Springer, Dordrecht, Netherlands

Faris JD, Zhang ZC, Lu HJ, Lu SW, Reddy L, Cloutier S, Fellers JP, Meinhardt SW, Rasmussen JB, Xu SS, Oliver RP, Simons KJ, Friesen TL (2010) A unique wheat disease resistance-like gene governs effector-triggered susceptibility to necrotrophic pathogens. P Natl Acad Sci USA 107:13544-13549

Feldman M, Levy AA (2012) Genome evolution due to allopolyploidization in wheat. Genetics 192:763-774

Feuillet C, Travella S, Stein N, Albar L, Nublat A, Keller B (2003) Map-based isolation of the leaf rust disease resistance gene *Lr10* from the hexaploid wheat (*Triticum aestivum* L.) genome. P Natl Acad Sci USA 100: 15253-15258

Feuillet C, Eversole K (2007) Physical mapping of the wheat genome: A coordinated effort to lay the foundation for genome sequencing and develop tools for breeders. Isr J Plant Sci 55:307-313

Flor HH (1971) Current status of gene-for-gene concept. Annu Rev Phytopathol 9:275-296

Foote T, Roberts M, Kurata N, Sasaki T, Moore G (1997) Detailed comparative mapping of cereal chromosome regions corresponding to the *Ph1* locus in wheat. Genetics 147:801-807

Freialdenhoven A, Peterhansel C, Kurth J, Kreuzaler F, Schulze-Lefert P (1996) Identification of genes required for the function of non-race-specific *mlo* resistance to powdery mildew. Plant Cell 8:5-14

Fu D, Uauy C, Distelfeld A, Blechl A, Epstein L, Chen XM, Sela H, Fahima T, Dubcovsky J (2009) A novel kinase-START gene confers temperature-dependent resistance to wheat stripe rust. Science 323:1357-1360

Gale MD, Devos KM (1998) Comparative genetics in the grasses. P Natl A Sci USA 95:1971-1974

Gallego F, Feuillet C, Messmer M, Penger A, Graner A, Yano M, Sasaki T, Keller B (1998) Comparative mapping of the two wheat leaf rust resistance loci *Lr1* and *Lr10* in rice and barley. Genome 41:328-336

Gill BS, Appels R, Botha-Oberholster AM, Buell CR, Bennetzen JL, Chalhoub B, Chumley F, Dvorak J, Iwanaga M, Keller B, Li WL, McCombie WR, Ogihara Y, Quetier F, Sasaki T (2004) A Workshop Report on Wheat Genome Sequencing: International Genome Research on Wheat Consortium. Genetics 168:1087-1096

Gill KS, Arumuganathan K, Lee JH (1999) Isolating individual wheat (*Triticum aestivum*) chromosome arms by flow cytometric analysis of ditelosomic lines. Theor Appl Genet 98:1248-1252

Gill U, Brueggeman R, Nirmala J, Chai Y, Steffenson B, Kleinhofs A (2016) Molecular and genetic characterization of barley mutants and genetic mapping of mutant *rpr2* required for *Rpg1*-mediated resistance against stem rust. Theor Appl Genet 129:1519-1529

Giorgi D, Farina A, Grosso V, Gennaro A, Ceoloni C, Lucretti S (2013) FISHIS: Fluorescence *In Situ* Hybridization in Suspension and chromosome flow sorting made easy. Plos One 8:e57994

Glawe (2008) The powdery mildews: A review of the world's most familiar (yet poorly known) plant pathogens. Annu Rev Phytopathol 46:27-51

Godfrey D, Böhlenius H, Pedersen C, Zhang Z, Emmersen J, Thordal-Christensen H (2010) Powdery mildew fungal effector candidates share N-terminal Y/F/WxC-motif. BMC Genomics 11:317

Griffiths S, Sharp R, Foote TN, Bertin I, Wanous M, Reader S, Colas I, Moore G (2006) Molecular characterization of *Ph1* as a major chromosome pairing locus in polyploid wheat. Nature 439:749-752

Gupta PK, Rustgi S, Mir RR (2013) Array-based high-throughput DNA markers and genotyping platforms for cereal genetics and genomics. In: Gupta PK, Varshney RK (eds) Cereal Genomics II, Springer, New York, NY, USA, pp 11-55

Gururani MA, Venkatesh J, Upadhyaya CP, Nookaraju A, Pandey SK, Park SW (2012) Plant disease resistance genes : Current status and future directions. Physiol Mol Plant P 78:51-65

Hacquard S, Kracher B, Maekawa T, Vernaldi S, Schulze-Lefert P, van Themaat EVL (2013) Mosaic genome structure of the barley powdery mildew pathogen and conservation of transcriptional programs in divergent hosts. P Natl Acad Sci USA 110:E2219-E2228

Halperin M (1936) Cereals and civilization. Sci Mon 42:355-364

Halterman DA, Zhou FS, Wi FS, Wise RP, Schulze-Lefert P (2001) The MLA6 coiled-coil, NBS-LRR protein confers *AvrMla6*-dependent resistance specificity to *Blumeria graminis* f. sp. *hordei* in barley and wheat. Plant J 25:335-348

Halterman DA, Wei FS, Wise RP (2003) Powdery mildew-induced *Mla* mRNAs are alternatively spliced and contain multiple upstream open reading frames. Plant Physiol 131:558-567

Halterman DA, Wise RP (2004) A single-amino acid substitution in the sixth leucine-rich repeat of barley MLA6 and MLA13 alleviates dependence on RAR1 for disease resistance signaling. Plant J 38:215-226

Hanemann A, Schweizer GF, Cossu R, Wisker T, Röder MS (2009) Fine mapping, physical mapping and development of diagnostic markers for the *Rrs2* scald resistance gene in barley. Theor Appl Genet 119:1507-1522

Heath MC (2000) Nonhost resistance and nonspecific plant defenses. Curr Opin Plant Biol 3:315-319

Hernandez P, Martis M, Dorado G, Pfeifer M, Galvez S, Schaaf S, Jouve N, Šimková H, Valárik M, Doležel J, Mayer KFX (2012) Next-generation sequencing and syntenic integration of flow-sorted arms of wheat chromosome 4A exposes the chromosome structure and gene content. Plant J 69:377-386

Hogenhout SA, Van der Hoorn RAL, Terauchi R, Kamoun S (2008) Emerging concepts in effector biology of plant-associated organisms. Mol Plant Microbe In 22:115-122

Horsley RD, Franckowiak JD, Schwarz PB (2009) Barley. In: Carena MJ (ed) Cereals, vol 3, Handbook of plant breeding. Springer, New York, NY, USA, pp 227-250

Horvath H, Rostoks N, Brueggeman R, Steffenson B, von Wettstein D, Kleinhofs A (2003) Genetically engineered stem rust resistance in barley using the *Rpg1* gene. P Natl Acad Sci USA 100:364-369

Hsam SLK, Zeller FJ (2002) Breeding for powdery mildew resistance in common wheat (*T. aestivum* L. em Thell.). In: Bélanger RR, Bushnell WR, Dik AJ, Carver TLW (eds) The Powdery Mildews: A comprehensive Treatise, APS Press, St. Paul, USA, pp 219-238

Huang L, Brooks SA, Li WL, Fellers JP, Trick HN, Gill BS (2003) Map-based cloning of leaf rust resistance gene *Lr21* from the large and polyploid genome of bread wheat. Genetics 164:655-664

Huang SW, Weigel D, Beachy RN, Li JY (2016) A proposed regulatory framework for genome-edited crops. Nat Genet 48:109-111

Huang XH, Lu TT, Han B (2013) Resequencing rice genomes: an emerging new era of rice genomics. Trends Genet 29:225-232

Huang XH, Han B (2014) Natural variations and genome-wide association studies in crop plants. Annu Rev Plant Biol 65:531-551

Hückelhoven R, Dechert C, Kogel KH (2001) Non-host resistance of barley is associated with a hydrogen peroxide burst at sites of attempted penetration by wheat powdery mildew fungus. Mol Plant Pathol 2:199-205

Hückelhoven R, Panstruga R (2011) Curr Opin Plant Biol 14:738-746

Humphry M, Consonni C, Panstruga R (2006) *mlo*-based powdery mildew immunity: silver bullet or simply non-host resistance? Mol Plant Pathol 7:605-610

Hurni S, Brunner S, Buchmann G, Herrer G, Jordan T, Krukowski P, Wicker T, Yahiaoui N, Mago R, Keller B (2013) Rye *Pm8* and wheat *Pm3* are orthologous genes and

show evolutionary conservation of resistance function against powdery mildew. Plant J 76:957-969

Hurni S, Brunner S, Stirnweis D, Herren G, Peddito D, McIntosh RA, Keller B (2014) The powdery mildew resistance gene *Pm8* derived from rye is suppressed by its wheat ortholog *Pm3*. Plant J 79:904-913

Hysing SC, Merker A, Liljeroth E, Boebner RMD, Zeller FJ, Hsam SLK (2007) Powdery mildew resistance in 155 Nordic bread wheat cultivars and landraces. Hereditas 144:102-119

IBSC (2012) A physical, genetic and functional sequence assembly of the barley genome. Nature 493:711-717

Inuma T, Khodaparast SA, Takamatsu S (2007) Multilocus phylogenetic analyses within *Blumeria graminis*, a powdery mildew fungus of cereals. Mol Phylogenet Evol 44:741-751

Int HapMap Consortium (2003) The International HapMap Project. Nature 426:789-796

IWGSC (2014) A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science 345: 1251788

IWGSC (2018) Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science 361: eaar7191. DOI: 10.1126/science.aar7191

Jafary H, Szabo LJ, Niks RE (2006) Innate nonhost immunity in barley to different heterologous rust fungi is controlled by sets of resistance genes with different and overlapping specificities. Mol Plant Microbe In 19:1270-1279

Jafary H, Albertazzi G, Marcel TC, Niks RE (2008) High diversity of genes for nonhost resistance of barley to heterologous rust fungi. Genetics 178:2327-2339

Jakobson I, Peusha H, Timofejeva L, Järve K (2006) Adult plant and seedling resistance to powdery mildew in a *Triticum aestivum* x *Triticum militinae* hybrid line. Theor Appl Genet 112:760-769

Jakobson I, Reis D, Tiidema A, Peusha H, Timofejeva L, Valárik M, Kladivová M, Šimková H, Doležel J, Järve K (2012) Fine mapping, phenotypic characterization and validation of non-race-specific resistance to powdery mildew in a wheat-*Triticum militinae* introgression line. Theor Appl Genet 125:609-623

Janáková E, Jakobson I, Peusha H, Abrouk M, Škopová M, Šimková H, Šafář J, Vrána J, Järve K, Doležel J, Valárik M: Divergence between bread wheat and *Triticum militinae* in the powdery mildew resistance *QPm.tut-4A* locus and its implications

for cloning of the resistance gene. Theoretical and Applied Genetics: accepted for publication. https://doi.org/10.1007/s00122-018-3259-3

Jia JZ, Zhao SC, Kong XY, Li YR, Zhao GY, He WM, Appels R, Pfeifer M, Tao Y, Zhang XY *et al.* (2013) *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. Nature 496:91-95

Jones JDG, Dangl JL (2006) The plant immune system. Nature 444:323-329

Jørgensen JH (1992) Discovery, characterization and exploitation of *Mlo* powdery mildew resistance in barley. Euphytica 63:141-152

Jørgensen JH (1994) Genetics of powdery mildew resistance in barley. Crit Rev Plant Sci 13:97-119

Jørgensen JH (1996) Effect of three suppressors on the expression of powdery mildew resistance genes in barley. Genome 39:492-498

Jørgensen LN, Hovnmøller MS, Hansen JG, Lassen P, Clark B, Bayles R, Rodemann B, Flath K, Jahn M, Goral T, Czembor J, Cheyton P, Maumene C, De Pope C, Ban R, Nielsen GC, Berg G (2014) IPM strategies and their dilemmas including an introduction to www.eurowheat.org. J Integr Agr 13:265-281

Kaur S, Zhang X, Mohan A, Dong HX, Vikram P, Singh S, Zhang ZW, Gill KS, Dhugga KS, Singh J (2017) Genome-wide association study reveals novel genes associated with culm cellulose content in bread wheat (*Triticum aestivum*, L.). Front Plant Sci 8:1913

Kim MC, Panstruga R, Elliot C, Müller J, Devoto A, Yoon HW, Park HC, Cho MJ, Schulze-Lefert P (2002) Calmodulin interacts with MLO protein to regulate defence against mildew in barley. Nature 416:447-450

Kintzios S, Jahoor A, Fischbeck G (1995) Powdery-mildew-resistance genes *Mla29* and *Mla32* in *H. spontaneum* derived winter-barley lines. Plant Breeding 114:265-266

Koller T, Brunner S, Herren G, Hurni S, Keller B (2018) Pyramiding of transgenic *Pm3* alleles in wheat results in improved powdery mildew resistance in the field. Theor Appl Genet, DOI: 10.1007/s00122-017-3043-9. [Epub ahead of print]

Komínková E, Dreiseitl A, Malečková E, Doležel J, Valárik M (2016): Genetic diversity of *Blumeria graminis* f. sp. *hordei* in Central Europe and its comparison with Australian population. Plos One 11: e0167099

Krattinger SG, Lagudah ES, Spielmeyer W, Singh RP, Huerta-Espino J, McFadden H, Bossolini E, Selter LL, Keller B (2009a) A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. Science 323:1360-1363

Krattinger S, Wicker T, Keller B (2009b) Map-based cloning of genes in *Triticeae* (wheat and barley). In: Muehlbauer G, Feuillet C (eds) Genetics and Genomics of the *Triticeae.* Plant Genetics and Genomics: Crops and Models, vol 7. Springer, New York, NY, USA, pp 337-357

Krattinger SG, Keller B (2016) Molecular genetics and evolution of disease resistance in cereals. New Phytol 212:320-332

Kusch S, Ahmadinejab N, Panstruga R, Kuhn H (2014) *In silico* analysis of the core signaling proteome from the barley powdery mildew pathogen (*Blumeria graminis* f. sp. *hordei*). BMC Genomics 15:843

Kusch S, Panstruga R (2017) *mlo*-based resistance: an apparently universal "weapon" to defeat powdery mildew disease. Mol Plant Microbe In 30:179-189

Leath S, Bowen KL (1989) Effects of powdery mildew, triadimenol seed treatment and triadimefon foliar sprays on yield of winter wheat in North Carolina. Phytopathology 79:152-155

Lev-Yadun S, Gopher A, Abbo S (2000) The cradle of agriculture. Science 288:1602-1603

Li GQ, Xu XY, Bai GH, Carver BF, Hunger R, Bonman JM (2016) Identification of novel powdery mildew resistance sources in wheat. Crop Sci 56:1817-1830

Li W, Zhang P, Fellers JP, Friebe B, Gill BS (2004) Sequence composition, organization, and evolution of the core *Triticeae* genome. Plant J 40:500-511

Liang JC, Fu BS, Tang WB, Khan NU, Li N, Ma ZQ (2016) Fine mapping of two wheat powdery mildew resistance genes located at the *Pm1* cluster. Plant Genome-US 9. DOI: 10.3835/plantgenome2015.09.0084

Limpert E, Godet F, Müller K (1999) Dispersal of cereal mildews across Europe. Agr Forest Meteorol 97:293-308

Ling HQ, Zhao SC, Liu DC, Wang JY, Sun H, Zhang C, Fan HJ, Li D, Dong LL, Tao Y *et al.* (2013) Draft genome of the wheat A-genome progenitor *Triticum urartu*. Nature 496:87-90

Liu N, Bai GH, Lin M, Xu XY, Zheng WM (2017) Genome wide association analysis of powdery mildew resistance in US winter wheat. Sci Rep-UK 7:11743

Liu SX, Pumphrey MO, Gill BS, Trick HN, Zhang JX, Doležel J, Chalhoub B, Anderson JA (2008) Towards positional cloning of *Fhb1*, a major QTL for Fusarium head blight resistance in wheat. Cereal Res Commun 26:307-316

Liu W, Frick M, Huel R, Nykiforuk CL, Wang XM, Gaudet DA, Eudes F, Conner RL, Kuzyk A, Chen Q, Kang ZS, Laroche A (2014) The stripe rust resistance gene Yr10 encodes an evolutionary-conserved and unique CC-NBS-LRR sequence in wheat. Mol Plant 7: 1740-1755

Liu YX, Lin Y, Gao S, Li ZY, Ma J, Deng M, Chen GY, Wei YM, Zheng YL (2017) A genome-wide association study of 23 agronomic traits in Chinese wheat landraces. Plant J 91:861-873

Loyce C, Meynard JM, Bouchard C, Rolland B, Lonnet P, Bataillon P, Bernicot MH, Bonnefoy M, Charrier X, Debote B, Demarquet T, Duperrier B, Félix I, Heddadj D, Leblanc O, Leleu M, Mangin P, Méausoone M, Doussinault G (2008) Interaction between cultivar and crop management effects on winter wheat diseases, lodging and yield. Crop Prot 27:1131-1142

Lu HJ, Faris JD (2006) Macro- and microcolinearity between the genomic region of wheat chromosome 5B containing the *Tsn1* gene and the rice genome. Funct Integr Genomic 6:90-103

Lu XL, Kracher B, Isabel ML Saur, Bauer S, Ellwood SR, Wise R, Yaeno T, Maekawa T, Schulze-Lefert P (2016) Allelic barley MLA immune receptors recognize sequence-unrelated avirulence effectors of the powdery mildew pathogen. P Natl Acad Sci USA 113:E6486-E6495

Luo MC, Thomas C, You FM, Hsiao J, Shu OY, Buell CR, Malandro M, McGuire PE, Anderson OD, Dvorak J (2003) High-throughput fingerprinting of bacterial artificial chromosomes using the SNaPshot labeling kit and sizing of restriction fragments by capillary electrophoresis. Genomics 82:378-389

Lüpken T, Stein N, Perovic D, Habekuß A, Serfling A, Krämer I, Hähnel U, Steurnagel B, Scholz U, Ariyadasa R, Martis M, Mayer K, Niks RE, Collins NC, Friedt W, Ordon F (2014) High-resolution mapping of the barley *Ryd3* locus controlling tolerance to BYDV. Mol Breeding 33:477-488

Mago R, Spielmeyer W, Lawrence GJ, Ellis JG, Pryor AJ (2004) Resistance for rye stem rust (*SrR*) and barley powdery mildew (*Mla*) are located in syntenic regions on short arm of chromosome. Genome 47:112-121

Mago R, Zhang P, Vautrin S, Šimková H, Bansal U, Luo MC, Rouse M, Karaoglu H, Periyannan S, Kolmer J, Jin Y, Ayliffe MA, Bariana H, Park RF, McIntosh R, Doležel

J, Bergès H, Spielmeyer W, Lagudah ES, Ellis JG, Dodds PN (2015) The wheat *Sr50* gene reveals rich diversity at a cereal disease resistance locus. Nat Plants 1:15186

Marcussen T, Sandve SR, Heier L, Spannagl M, Pfeifer M, Jakobsen KS, Wulff BBH, Steuernagel B, Mayer KFX, Olsen OA (2014) Ancient hybridizations among the ancestral genomes of bread wheat. Science 345:1250092

Mascher M, Gundlach H, Himmelbach A, Beier S, Twardziok SO, Wicker T, Radchuk V, Dockter C, Hedley PE, Russell J *et al.* (2017) A chromosome conformation capture ordered sequence of the barley genome. Nature 544:427-433

Mascher M, Jost M, Kuon JE, Himmelbach A, Aßfalg A, Beier S, Scholz U, Graner A, Stein N (2014) Mapping-by-sequencing accelerates forward genetics in barley. Genome Biol 15:R78

Mascher M, Schuenemann VJ, Davidovich U, Marom N, Himmelbach A, Hübner S, Korol A, David M, Reiter E, Riehl S, Schreiber M, Vohr SH, Green RE, Dawson IK, Russell J, Kilian B, Muehlbauer GJ, Waugh R, Fahima T, Krause J, Weiss E, Stein N (2016) Genomic analysis of 6,000-year-old cultivated grain illuminates the domestication history of barley. Nat Genet 48:1089-1093

Matsumoto T, Wu JZ, Kanamori H, Katayose Y, Fujisawa M, Namiki N, Mizuno H, Yamamoto K, Antonio BA, Baba T *et al.* (2005) The map-based sequence of the rice genome. Nature 436:793-800

Mayer KFX, Martis M, Hedley PE, Šimková H, Liu H, Morris JA, Steuernagel B, Taudien S, Roessner S, Gundlach H, Kubaláková M, Suchánková P, Murat F, Felder M, Nussbaumer T, Graner A, Salse J, Endo T, Sakai H, Tanaka T, Itoh T, Sato K, Platzer M, Matsumoto T, Scholz U, Doležel J, Waugh R, Stein N (2011) Unlocking the barley genome by chromosomal and comparative genomics. Plant Cell 23:1249-1263

McDonald BA (2014) Using dynamic diversity to achieve durable disease resistance in agricultural ecosystems. Trop Plant Pathol 39:191-196

McDonald BA, Linde C (2002) Pathogen population genetics, evolutionary potential and durable resistance. Annu Rev Phytopathol 40:349-379

McHale L, Tan X, Koehl P, Michelmore RW (2006) Plant NBS-LRR proteins: adaptable guards. Genome Biol 7:212

Menardo F, Praz C, Wyder S, Ben-David R, Bourras S, Matsumae H, McNally KE, Parlange F, Riba A, Roffler S, Schaefer LK, Shimizu KK, Valenti L, Zbinden H,

Wicker T, Keller B (2016) Hybridization of powdery mildew strains gives rise to pathogens on novel agricultural crop species. Nat Genet 48:201-205

Menardo F, Wicker T, Keller B (2017a) Reconstructing the evolutionary history of powdery mildew lineages (*Blumeria graminis*) at different evolutionary time scales with NGS data. Genom Biol Evol 9:446-456

Menardo F, Praz CR, Wicker T, Keller B (2017b) Rapid turnover of effectors in grass powdery mildew (*Blumeria graminis*). BMC Evol Biol 17:223

Millet E, Manisterski J, Ben-Yehuda P, Distelfeld A, Deek J, Wan A, Chen X, Steffenson BJ (2014) Introgression of leaf rust and stripe rust resistance from Sharon Goatgrass (*Aegilops sharonensis* Eig) into bread wheat (*Triticum aestivum* L.). Genome 57:309-316

Mondal S, Rutkovski JE, Velu G, Singh PK, Crespo-Herrera LA, Guzmán C, Bhavani S, Lan CX, He XY, Singh RP (2016) Harnessing diversity in wheat to enhance grain yield, climate resilience, disease and insect pest resistance and nutrition through conventional and modern breeding approaches. Front Plant Sci 7:991

Moore JW, Herrera-Foessel S, Lan CX, Schnippenkoetter W, Ayliffe M, Huerta-Espino J, Lillemo M, Viccars L, Milne R, Periyannan S, Kong XY, Spielmeyer W, Talbot M, Bariana H, Patrick JW, Dodds P, Singh R, Lagudah E (2015) A recently evolved hexose transporter variant confers resistance to multiple pathogens in wheat. Nat Genet 47:1494-1498

Morrell P, Clegg M (2007) Genetic evidence for a second domestication of barley (*Hordeum vulgare*) east of the Fertile Crescent. P Natl Acad Sci USA 104:3289-3294

Mullan DJ, Mirzaghaderi G, Walker E, Colmer TD, Francki MG (2009) Development of wheat-*Lophopyrum elongatum* recombinant lines for enhanced sodium 'exclusion' during salinity stress. Theor Appl Genet 119:1313-1323

Mysore KS, Ryu CM (2004) Nonhost resistance: how much do we know? Trends Plant Sci 9:97-104

Niks RE, Qi XQ, Marcel TC (2015) Quantitative resistance to biotrophic filamentous plant pathogens: concepts, misconceptions and mechanisms. Annu Rev Phytopathol 53: 445-470

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

Oberhaensli S, Parlange F, Buchmann JP, Jenny FH, Abbott JC, Burgis TA, Spanu PD, Keller B, Wicker T (2011) Comparative sequence analysis of wheat and barley powdery mildew fungi reveals gene collinearity, dates divergence and indicates host-pathogen co-evolution. Fungal Genet Biol 48:327-334

Parlange F, Roffler S, Menardo F, Ben-David R, Bourras S, McNally KE, Oberhaensli S, Stirnweis D, Buchmann G, Wicker T, Keller B (2015) Genetic and molecular characterization of a locus involved in avirulence of *Blumeria graminis* f. sp. *tritici* on wheat *Pm3* resistance alleles. Fungal Genet Biol 82:181-192

Pasam RK, Sharma R, Malosetti M, van Eeuwijk FA, Haseneyer G, Kilian B, Graner A (2012) Genome-wide association studies for agronomical traits in a world psring barley collection. BMC Plant Biol 12:16

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A *et al.* (2009) The sorghum bicolor genome and the diversification of grasses. Nature 457:551-556

Paux E, Roger D, Badaeva E, Gay G, Bernard M, Sourdille P, Feuillet C (2006) Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. Plant J 48:463-474

Paux E, Sourdille P, Salse J, Saintenac C, Choulet F, Leroy P, Korol A, Michalak M, Kianin S, Spielmeyer W, Lagudah E, Somers D, Kilian A, Alaux M, Vautrin S, Bergès H, Eversole K, Appels R, Šafář J, Šimková H, Doležel J, Bernard M, Feuillet C (2008) A physical map of the 1-gigabase bread wheat chromosome 3B. Science 332:101-104

Pedersen C, van Themaat EVL, McGuffin LJ, Abbott JC, Burgis TA, Barton G, Bindschedler LV, Lu XL, Maekawa T, Wessling R, Cramer R, Thordal-Christensen H, Panstruga R, Spanu PD (2012) Structure and evolution of barley powdery mildew effector candidates. BMC Genomics 13:694

Peng JH, Sun D, Nevo E (2011) Domestication evolution, genetics and genomics in wheat. Mol Breeding 28:281-301

Pennington HG, Gheorghe DM, Damerum A, Pliego C, Spanu PD, Cramer R, Bindschedler LV (2016) Interactions between the powdery mildew effector BEC1054 and barley proteins identify candidate host targets. J Proteome Res 15:826-839

Periyannan S, Moore J, Ayliffe M, Bansal U, Wang XJ, Huang L, Deal K, Luo MC, Kong XY, Bariana H, Mago R, McIntosh R, Dodds P, Dvorak J, Lagudah E (2013) The

gene *Sr33*, an ortholog of barley *Mla* genes, encodes resistance to wheat stem rust race Ug99. Science 341:786-788

Perovic D, Stein N, Zhang HN, Drescher A, Prasad M, Kota R, Kopahnke D, Graner A (2004) An integrated approach for comparative mapping in rice and barley with special reference to the *Rph16* resistance locus. Funct Integr Genomic 4:74-83

Piffanelli P, Zhou FS, Casais C, Orme J, Jarosch B, Schaffrath U, Collins NC, Panstruga R, Schulze-Lefert P (2002) The barley MLO modulator of defense and cell death is responsive to biotic and abiotic stress stimuli. Plan Physiol 129:1076-1085

Power RA, Parkhill J, de Oliveira T (2017) Microbial genome-wide association studies: lessons from human GWAS. Nat Rev Genet 18:41-50

Praz CR, Bourras S, Zeng FS, Sánchez-Martín J, Menardo F, Xue MF, Yang LJ, Roffler S, Böni R, Herren G, McNally KE, Ben-David R, Parlange F, Oberhaensli S, Flückiger S, Schäfer LK, Wicker T, Yu DZ, Keller B (2017) *AvrPm2* encodes an RNase-like avirulence effector which is conserved in two different specialized forms of wheat and rye powdery mildew fungus. New Phytol 213:1301-1314

Pritchard L, Birch PRJ (2014) The zigzag model of plant-microbe interactions: is it time to move on? Mol Plant Pathol 15:865-870

Putnam NH, O'Connell BL, Stites JC, Rice BJ, Blanchette M, Calef R, Troll CJ, Fields A, Hartley PD, Sugnet CW, Haussler D, Rokhars DS, Green RE (2016) Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. Genome Res 26:342-350

Qi LL, Gill BS (2001) High-density physical maps reveal that the dominant male-sterile gene *Ms3* is located in a genomic region of low recombination in wheat and is not amenable to map-based cloning. Theor Appl Genet 103:998-1006

Qi LL, Echalier B, Chao S, Lazo GR, Butler GE, Anderson OD, Akhunov ED, Dvorak J, Linkiewicz AM, Ratnasiri A *et al.* (2004) A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. Genetics 168: 701-712

Rajamaran J, Douchkov D, Hensel G, Stefanato FL, Gordon A, Ereful N, Caldararu OF, Petrescu AJ, Kumlehn J, Boyd LA, Schweizer P (2016) An LRR/malectin receptor-like kinase mediates resistance to non-adapted and adapted powdery mildew fungi in barley and wheat. Front Plant Sci 7:1836

Randhawa M, Bansal U, Valárik M, Klocová B, Doležel J, Bariana H (2014) Molecular mapping of stripe rust resistance gene *Yr51* in chromosome 4AL of wheat. Theor Appl Genet 127: 317-324

Rasheed A, Hao YF, Xia XC, Khan A, Xu YB, Varshney RK, He ZH (2017) Crop breeding chips and genotyping platforms: progress, challenges and perspectives. Mol Plant 10:1047-1064

Rawat N, Pumphrey MO, Liu SX, Zhang XF, Tiwari VK, Ando K, Trick HN, Bockus WW, Akhunov E, Anderson JA, Gill BS (2016) Wheat *Fhb1* encodes a chimeric lectin with agglutinin domains and a pore-formin toxin-like domain conferring resistance to Fusarium head blight. Nat Genet 48:1576-1580

Reinert S, Kortz A, Leon J, Naz AA (2016) Genome-wide association mapping in the global diversity set reveals new QTL controlling root system and related shoot variation in barley. Front Plant Sci 7:1061

Riley R, Chapman V (1958) Genetic control of the cytologically diploid behavior of hexaploid wheat. Nature 182:713-715

Romero CCT, Aghnoum R, Niks RE (2015) Fine-mapping of a QTL for nonhost resistance to *Blumeria graminis* f. sp. *tritici* in barley. Paper presented at: 14th International Cereal Rusts and Powdery Mildews Conference; Helsingør, Denmark; July 5-8, 2015

Romero CCT, Vermeulen JP, Vels A, Himmelbach A, Mascher M, Niks RE (2018) Mapping resistance to powdery mildew in barley reveals a large-effect nonhost resistance QTL. Theor Appl Genet. DOI: 10.1007/s00122-018-3055-0

Saintenac C, Zhang WJ, Salcedo A, Rouse MN, Trick HN, Akhunov E, Dubcovsky J (2013) Identification of wheat gene *Sr35* that confers resistance to UG99 stem rust race group. Science 341:783-786

Salamini F, Ozkan H, Brandolini A, Schafer-Pregl R, Martin W (2002) Genetics and geography of wild cereal domestication in the Near East. Nat Rev Genet 3:429-441

Sánchez-Martín J, Steuernagel B, Ghosh S, Herren G, Hurni S, Adamski N, Vrána J, Kubaláková M, Krattinger S, Wicker T, Doležel J, Keller B (2016) Genome Biol 17:221

Sears ER (1966) Nullisomic-tetrasomic combinations in hexaploid wheat. In: Riley R, Lewis KR (eds) Chromosome Manipulations and Plant Genetics. Springer, Boston, MA, USA, pp 29-45

Sears ER, Sears LMS (1978) The telocentric chromosomes of common wheat. In: Ramanujam S. (ed) Proceedings of 5th International Wheat Genetic Symposium. New Delhi, India, pp 389-407

Schulze-Lefert P, Panstruga R (2011) A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. Trends Plant Sci 16:117-125

Seeholzer S, Tsuchimatsu T, Jordan T, Bieri S, Pajonk S, Yang W, Jahoor A, Shimizu KK, Keller B, Schulze-Lefert P (2010) Diversity at the *Mla* powdery mildew resistance locus from cultivated barley reveals sites of positive selection. Mol Plant Microbe In 23:497-509

Sehgal D, Singh R, Rajpal VR (2016) Quantitative trait loci mapping in plants: concepts and approaches. In: Rajpal VR, Rao SR, Raina SN (eds) Molecular Breeding for Sustainable Crop Improvement: Volume 2. Springer International Publishing AG, Cham, Switzerland, pp 31-59

Semagn K, Bjørnstad Å, Ndjiondjop MN (2006) An overview of molecular marker methods for plants. Afr J Biotechnol 5:2540-2568

Shen QH, Zhou FS, Brieri S, Haizel T, Shirasu K, Schulze-Lefert P (2003) Recognition specificity and RAR1/SGT1 dependence in barley *Mla* disease resistance genes to the powdery mildew fungus. Plant Cell 15:732-744

Shiferaw B, Smale M, Braun HJ, Duveiller E, Reynolds M, Muricho G (2013) Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. Food Secur 5:291-317

Shirasu K, Lahaye T, Tam MW, Zhou F, Azevedo C, Schulze-Lefert P (1999) A novel class of eukaryotic zinc-binding proteins is required for disease resistance signaling in barley and development in *C. elegans*. Cell 99:355-366

Schoenlechner R, Siebenhandl S, Berghofer E (2008) Pseudocereals. In: Arendt E, Dal Bello F (eds) Gluten-Free Cereal Products and Beverages, Elsevier Academic Press Inc, San Diego, CA, USA, pp 149-190

Sonah H, Deshmukh RK, Bélanger RR (2016) Computational prediction of effector proteins in fungi: opportunities and challenges. Front Plant Sci 7:126

Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, Stuber K, van Themaat EVL, Brown JKM, Butcher SA, Gurr SJ *et al.* (2010) Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. Science 330:1543-1546

Sprink T, Metje J, Hartung F (2015) Plant genome editing by novel tools: TALEN and other sequence specific nucleases. Curr Opin Biotech 32:47-53

Srichumpa P, Brunner S, Keller B, Yahiaoui N (2006) Allelic series of four powdery mildew resistance genes at the *Pm3* locus in hexaploid bread wheat. Plan Physiol 139:885-895

St Clair DA (2010) Quantitative disease resistance and quantitative resistance loci in breeding. Annu Rev Phytopathol 48:247-268

Staňková H, Valárik M, Lapitan NLV, Berkman PJ, Batley J, Edwards D, Luo MC, Tulpová Z, Kubaláková M, Stein N, Doležel J, Šimková H (2015) Chromosomal genomics facilitates fine mapping of a Russian wheat aphid resistance gene. Theor Appl Genet 128:1373-1383

Steuernagel B, Periyannan SK, Hernández-Pinzón I, Witek K, Rouse MN, Yu G, Hatta A, Ayliffe M, Bariana H, Jones JDG, Lagudah ES, Wulff BBH (2016) Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. Nat Biotechnol 34:652-655

Stirnweis D, Milani SD, Brunner S, Herrer G, Buchmann G, Peditto D, Jordan T, Keller B (2014) Suppression among alleles encoding nucleotide-binding-leucine-rich repeat resistance proteins interferes with resistance in F1 hybrid and allele-pyramided wheat plants. Plant J 79: 893-903

Sukumaran S, Lopes MS, Dreisigacker D, Dixon LE, Zikhali M, Griffiths S, Zheng BY, Chapman S, Reynolds MP (2016) Identification of earliness per se flowering time locus in spring wheat through a genome-wide association study. Crop Sci 56:2962-2972

Šafář J, Bartoš J, Janda J, Bellec A, Kubaláková M, Valárik M, Pateyron S, Weiserová J, Tušková R, Číhalíková J, Vrána J, Šimková H, Faivre-Rampant P, Sourdille P, Caboche M, Bernard M, Doležel J, Chalhoub B (2004) Dissecting large and complex genomes: flow sorting and BAC cloning of individual chromosome from bread wheat. Plant J 39:960-968

Takamatsu S, Matsuda S (2003) Evolution of host-parasite relationships of *Golovinomyces* (Ascomycete: Erysiphaceae) inferred from nuclear rDNA sequences. Mol Phylogenet Evol 27:314-327

Tester M, Langridge P (2010) Breeding technologies to increase crop production in a changing world. Science 327:818-822

Thind AK, Wicker T, Šimková H, Fossati D, Moullet O, Brabant C, Vrána J, Doležel J, Krattinger S (2017) Rapid cloning of genes in hexaploid wheat using cultivar-specific long-range chromosome assembly. Nat Biotechnol 35:793-796

Troch V, Audenaert K, Wyand RA, Haesaert G, Höfte M, Brown JKM (2014) *Formae speciales* of cereal powdery mildew: close or distant relatives? Mol Plant Pathol 15:304-314

Tucker EJ, Baumann U, Kouidri A, Suchecki R, Baes M, Garcia M, Okada T, Dong CM, Wu YZ, Sandhu A, Singh M, Langridge P, Wolters P, Albertsen MC, Cigan AM, Whitford R (2017) Molecular identification of the wheat male fertility gene *Ms1* and its prospects for hybrid breeding. Nat Commun 8:869

Uauy C, Distelfeld A, Fahima T, Blechl A, Dubcovsky J (2006) A NAC gene regulating senescence improves grain protein, zinc and iron content in wheat. Science 314:1298-1301

Uitterlinden AG (2016) An introduction to genome-wide association studies: GWAS for dummies. Semin Reprod Med 34:196-204

Valluru R, Reynolds MP, Davies WJ, Sukumaran S (2017) Phenotypic and genome-wide association analysis of spike ethylene in diverse wheat genotypes under heat stress. New Phytol 214:271-283

Vogel JP, Garvin DF, Mockler TC, Schmutz J, Rokhsar D, Bevan MW, Barry K, Lucas S, Harmon-Smith M, Lail K *et al.* (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature 463:763-768

Vu GTH, Wicker T, Buchmann JP, Chandler PM, Matsumoto T, Graner A, Stein N (2010) Fine mapping and syntenic integration of the semi-dwarfing gene *sdw3* of barley. Funct Integr Genomic 10: 509-521

Wang MH, Jiang N, Jia TY, Leach L, Cockram J, Waugh R, Ramsay L, Thomas B, Luo ZW (2012) Genome-wide association mapping of agronomic and morphologic traits in highly structured populations of barley cultivars. Theor Appl Genet 124:233-246

Wang YP, Cheng X, Shan QW, Zhang Y, Liu JX, Gao CX, Qiu JL (2014) Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. Nat Biotechnol 32:947-951

Wei F, Gobelman-Werner K, Morroll AM, Kurth J, Mao L, Wing R, Leister D, Schulze-Lefert P, Wise RP (1999) The *Mla* (powdery mildew) resistance cluster is associated with three NBS-LRR gene families and suppressed recombination within a 240-kb DNA interval on chromosome 5S (1HS) of barley. Genetics 153:1929-1948

Wicker T, Mayer KFX, Gundlach H, Martis M, Steuernagel B, Scholz U, Šimková H, Kubaláková M, Choulet F, Taudien S, Platzer M, Feuillet C, Fahima T, Budak H,

Doležel J, Keller B, Stein N (2011) Frequent gene movement and pseudogene evolution is common to the large and complex genomes of wheat, barley and their relatives. Plant Cell 23:1706-1718

Wicker T, Oberhaensli S, Parlange F, Buchmann JP, Shatalina M, Roffler S, Ben-David R, Doležel J, Šimková H, Schulze-Lefert P, Spanu PD, Bruggmann R, Amselem J, Quesneville H, van Themaat EVL, Paape T, Shimizu KK, Keller B (2013) The wheat powdery mildew genome shows the unique evolution of an obligate biotroph. Nat Genet 45:1092-1096

Wolfe MS (1984) Trying to understand and control powdery mildew. Plant Pathol 33:451-466

Wolfe MS, McDermott JM (1994) Population genetics of plant pathogen interactions: the example of the *Erysiphe graminis-Hordeum vulgare* pathosystem. Annu Rev Phytopathol 32:89-113

Wolt JD, Wang K, Yang B (2016) The regulatory status of genome-edited crops. Plant Biotechnol J 14:510-518

Wulff BBG, Moscou MJ (2014) Strategies for transferring resistance into wheat: from wide crosses to GM cassettes. Front Plant Sci 5:692

Wyand RA, Brown JKM (2003) Genetic and *forma specialis* diversity in *Blumeria graminis* of cereals and its implications for host-pathogen co-evolution. Mol Plant Pathol 4:187-198

Yahiaoui N, Srichumpa P, Dudler R, Keller B (2004) Genome analysis at different ploidy levels allows cloning of the powdery mildew resistance gene *Pm3b* from hexaploid wheat. Plant J 37:528-538

Yahiaoui N, Brunner S, Keller B (2006) Rapid generation of new powdery mildew resistance genes after wheat domestication. Plant J 48:85-98

Yao C, Wu Y, Nie H, Tang D (2012) RPN1a, a 26S proteasome subunit, is required for innate immunity in *Arabidopsis*. Plant J 71:1015-1028

Yeo FKS, Bouchon R, Kuijken R, Loriaux A, Boyd C, Niks RE, Marcel TC (2017) High-resolution mapping of genes involved in plant stage-specific partial resistance of barley to leaf rust. Mol Breeding 37:45

Zhang Z, Henderson C, Perfect E, Carver TLW, Thomas BJ, Skamnioti P, Gurr SJ (2005) Of genes and genomes, needles and haystacks: *Blumeria graminis* and functionality. Mol Plant Pathol 6:561-575

Zhang WJ, Pedersen C, Kwaaitaal M, Gregersen PL, Mørch SM, Hanisch S, Kristensen A, Fuglsang AT, Collinge DB, Thordal-Christensen H (2012) Interaction of barley powdery mildew effector candidate CSEP0055 with the defence protein PR17c. Mol Plant Pathol 13:1110-1119

Zhou FS, Kurth J, Wei FS, Elliot C, Valè G, Yahiaoui N, Keller B, Somerville S, Wise R, Schulze-Lefert P (2001) Cell-autonomous expression of barley *Mla1* confers race-specific resistance to the powdery mildew fungus via a *Rar1*-independent signaling pathway. Plant Cell 13:337-350

Zhou Y, Tang H, Cheng MP, Dankwa KO, Chen ZX, Li ZY, Gao S, Liu YX, Jiang QT, Lan XJ, Pu ZE, Wei YM, Zheng YL, Hickey LT, Wang JR (2017) Genome-wide association study for pre-harvest sprouting resistance in a large germplasm collection of Chinese wheat landraces. Front Plant Sci 8:401

Zou S, Wang H, Li Y, Kong Z, Tang D (2017) The NB-LRR gene *Pm60* confers powdery mildew resistance in wheat. New Phytol, doi: 10.1111/nph.14964. [Epub ahead of print]

8 ABBREVIATIONS

ABC	ATP-Binding Cassette
ADP	Adenosine Diphosphate
ATP	Adenosine Triphosphate
Avr	Avirulence
BAC	Bacterial Artificial Chromosome
Bgh	*Blumeria graminis* f. sp. *hordei*
Bgt	*Blumeria graminis* f. sp. *tritici*
CS	*Triticum aestivum* cultivar Chinese Spring
CSEP	Candidate Secreted Effector Protein
cv.	Cultivar
DNA	Deoxyribonucleic Acid
EMS	Ethyl Methanesulfonate
ETI	Effector-Triggered Immunity
ETS	Effector-Triggered Susceptibility
FAO	Food and Agriculture Organization of the United Nations
ff. spp.	*Formae speciales*
f. sp.	*Forma specialis*
GFG	Gene-For-Gene
GWAS	Genome Wide Association Study
HR	Hypersensitive Cell Death Response
IBSC	International Barley Sequencing Consortium
IWGSC	International Wheat Genome Sequencing Consortium

LRR	Leucine-Rich Repeat
LTR	Long Terminal Repeat
Mlo	Mildew resistance Locus O
MTP	Minimal Tiling Path
NB	Nucleotide-Binding
NHR	Nonhost Resistance
PAMP	Pathogen-Associated Molecular Pattern
PTI	PAMP-triggered immunity
PRR	Pattern Recognition Receptor
RNAi	RNA Interference
QTL	Quantitative Trait Loci
QR	Quantitative Resistance
R genes	Major-effect Resistance genes
rDNA	Ribosomal DNA
Ror1	Required for mlo-specified Resistance1
RNA	Ribonucleic Acid
SNP	Single Nucleotide Polymorphism
Svr	Suppressor of avirulence
TE	Transposable Element
TILLING	Targeted Induced Lesions IN Genomes
TM	Transmembrane
Tm	*Triticum militinae*

9 CURRICULUM VITAE

PERSONAL INFORMATION

Name:	Eva Janáková (née Komínková)
Date of birth:	09.12.1987
Place of birth:	Olomouc, Czech Republic
Residence:	Slatinky 60, 783 42, Czech Republic

AFFILIATION

2013 – 2018	**Centre of Structural and Functional Genomics** **Institute of Experimental Botany AS CR** Šlechtitelů 31, 783 71 Olomouc, Czech Republic

EDUCATION

2013 – 2018	**Postgraduate study** at Palacký University Olomouc, Faculty of Science, Molecular and Cell Biology
2011 – 2013	**M.Sc.** at Palacký University Olomouc, Faculty of Science, Molecular and Cell Biology
2008 – 2011	**B.Sc.** at Palacký University Olomouc, Faculty of Science, Molecular and Cell Biology

RESEARCH FELLOWSHIPS

09/2014 – 11/2014	**Tallinn University of Technology, Department of Chemistry and Biotechnology, Estonia** Research focused on phenotyping of wheat powdery mildew supervised by Dr. Kadri Järve

03/2016 – 05/2016 **Institute of Plant Biology, Department of Plant and Microbial Biology, University of Zurich, Switzerland**

Research focused on barley powdery mildew genomics and bioinformatics supervised by Prof. Beat Keller

TRAINING COURSES

04/2017 **2nd Pathogenomics Training School of the SUSTAIN COST Action (FA1208); Norwich, United Kingdom**

LIST OF PUBLICATIONS

- Cápal P, Endo TR, Vrána J, Kubaláková M, Karafiátová M, Komínková E, Mora-Ramírez I, Weschke W, Doležel J (2016): The utility of flow sorting to identify chromosomes carrying a single copy transgene in wheat. Plant Methods 12: 24

- Komínková E, Dreiseitl A, Malečková E, Doležel J, Valárik M (2016): Genetic diversity of *Blumeria graminis* f. sp. *hordei* in Central Europe and its comparison with Australian population. Plos One 11: e0167099

- Abrouk M, Balcárková B, Šimková H, Komínková E, Martis M, Jakobson I, Timofejeva L, Rey E, Vrána J, Kilian A, Järve K, Doležel J, Valárik M (2017): The *in silico* identification and characterization of a bread wheat/*Triticum militinae* introgression line. Plant Biotechnology Journal 15:249-256

- Janáková E, Jakobson I, Peusha H, Abrouk M, Škopová M, Šimková H, Šafář J, Vrána J, Järve K, Doležel J, Valárik M: Divergence between bread wheat and *Triticum militinae* in the powdery mildew resistance *QPm.tut-4A* locus and its implications for cloning of the resistance gene. Theoretical and Applied Genetics: accepted for publication. https://doi.org/10.1007/s00122-018-3259-3

10 LIST OF APPENDICES

Publications

APPENDIX I: Divergence between bread wheat and *Triticum militinae* in the powdery mildew resistance *QPm.tut-4A* locus and its implications for cloning of the resistance gene

APPENDIX II: Genetic diversity of *Blumeria graminis* f. sp. *hordei* in Central Europe and its comparison with Australian population

APPENDIX III: The *in silico* identification and characterization of a bread wheat/ *Triticum militinae* introgression line

APPENDIX IV: The utility of flow sorting to identify chromosomes carrying a single copy transgene in wheat

Unpublished results

APPENDIX V: Utilizing long-term and extensive virulence monitoring data as a basis for *in silico* identification of candidate avirulence effectors of barley powdery mildew fungus

Conference contributions

APPENDIX VI: Positional cloning of powdery mildew resistance gene introgressed to bread wheat from *Triticum militinae* (2017)

APPENDIX VII: Molecular description of *Blumeria graminis* f. sp. *hordei* isolates (2015)

APPENDIX VIII: Positional cloning of powdery mildew resistance gene introgressed to bread wheat from *Triticum militinae* (2015)

APPENDIX I

Divergence between bread wheat and *Triticum militinae* in the powdery mildew resistance *QPm.tut-4A* locus and its implications for cloning of the resistance gene

Janáková E, Jakobson I, Peusha H, Abrouk M, Škopová M, Šimková H, Šafář J, Vrána J, Järve K, Doležel J, Valárik M

Theoretical and Applied Genetics: accepted for publication

https://doi.org/10.1007/s00122-018-3259-3

IF 3.930 (2017/2018)

Theoretical and Applied Genetics
https://doi.org/10.1007/s00122-018-3259-3

ORIGINAL ARTICLE

Divergence between bread wheat and *Triticum militinae* in the powdery mildew resistance *QPm.tut-4A* locus and its implications for cloning of the resistance gene

Eva Janáková[1] · Irena Jakobson[2] · Hilma Peusha[2] · Michael Abrouk[1,4] · Monika Škopová[1,3] · Hana Šimková[1] · Jan Šafář[1] · Jan Vrána[1] · Jaroslav Doležel[1] · Kadri Järve[2] · Miroslav Valárik[1]

Received: 10 August 2018 / Accepted: 3 December 2018

Abstract

A segment of *Triticum militinae* chromosome 7G harbors a gene(s) conferring powdery mildew resistance which is effective at both the seedling and the adult plant stages when transferred into bread wheat (*T. aestivum*). The introgressed segment replaces a piece of wheat chromosome arm 4AL. An analysis of segregating materials generated to positionally clone the gene highlighted that in a plant heterozygous for the introgression segment, only limited recombination occurs between the introgressed region and bread wheat 4A. Nevertheless, 75 genetic markers were successfully placed within the region, thereby confining the gene to a 0.012 cM window along the 4AL arm. In a background lacking the *Ph1* locus, the localized rate of recombination was raised 33-fold, enabling the reduction in the length of the region containing the resistance gene to a 480 kbp stretch harboring 12 predicted genes. The substituted segment in the reference sequence of bread wheat cv. Chinese Spring is longer (640 kbp) and harbors 16 genes. A comparison of the segments' sequences revealed a high degree of divergence with respect to both their gene content and nucleotide sequence. Of the 12 *T. militinae* genes, only four have a homolog in cv. Chinese Spring. Possible candidate genes for the resistance have been identified based on function predicted from their sequence.

Introduction

Disease imposes an important constraint on crop productivity and is most effectively and sustainably managed by breeding cultivars harboring genetic resistance. The ability of many pathogens to overcome host resistance means that the discovery of additional sources of resistance is a must for any crop improvement program. Since the gene pool of bread wheat (*Triticum aestivum*) has been so extensively narrowed by more than a century of intensive breeding (Feuillet et al. 2008), the search for novel resistance genes needs to be extended to older materials such as landraces and even to related cultivated and wild species (Mondal et al. 2016; Zamir 2001). Over many years, wheat cytogeneticists and breeders have succeeded in developing a diverse collection of germplasm harboring introgression segments of variable length, many of which have targeted the introduction of genes conditioning resistance to leaf pathogens (King et al. 2017; Valkoun 2001). Advances in genomic technologies have enabled a much greater precision than has been possible hitherto in the characterization of these materials (Abrouk et al. 2017; Tiwari et al. 2014; Winfield et al. 2016).

Communicated by Hermann Buerstmayr.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00122-018-3259-3) contains supplementary material, which is available to authorized users.

✉ Miroslav Valárik
valarik@ueb.cas.cz

[1] Institute of Experimental Botany of the Czech Academy of Sciences, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 31, 78371 Olomouc, Czech Republic

[2] Department of Chemistry and Biotechnology, Tallinn University of Technology, Akadeemia tee 15, 19086 Tallinn, Estonia

[3] Present Address: Limagrain Central Europe Cereals, s.r.o., Hrubčice 111, 79821 Bedihošť, Czech Republic

[4] Present Address: Biological and Environmental Science and Engineering Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Kingdom of Saudi Arabia

Published online: 07 December 2018

A number of introgressed genes conferring resistance to a fungal pathogen have been successfully isolated in hexaploid wheat: these currently comprise two conferring resistance to *Puccinia triticina* (Huang et al. 2003; Thind et al. 2017), five to *P. graminis* (Mago et al. 2015; Periyannan et al. 2013; Saintenac et al. 2013; Steuernagel et al. 2016) and one to *Blumeria graminis* (Hurni et al. 2013). *B. graminis* (commonly referred to as powdery mildew) epiphytotics can induce significant yield losses (Conner et al. 2003; Leath and Bowen 1989). A number of major resistance genes (*Pm* genes) have been identified (www.wheat.pw.usda.gov/cgi-bin/GG3/browse.cgi?class=gene); as most of these confer race-specific resistance, they are prone to being overcome by the rapidly evolving pathogen. Often durable form of resistance, referred to as adult plant resistance (APR), is typically conditioned by multiple genes, each conferring a small, but cumulative effect. As a result, uncovering its genetic basis normally requires quantitative trait locus (QTL) analysis. A meta-analysis (Lillemo and Lu 2015) has revealed that such genes are dispersed over 24 QTL-harboring regions, located on 18 of the 21 wheat chromosomes.

The tetraploid bread wheat relative *T. militinae* (genome formula A^tG) is generally considered to be a spontaneous mutant of *T. timopheevii* (Dorofeyev et al. 1976), although it has also been suggested to have arisen from an introgressive hybridization between *T. timopheevii* and the BA^u tetraploid *T. carthlicum* (Järve et al. 2002). Derivatives of a wide cross between the bread wheat cultivar (cv.) Tähti (genome formula BA^uD) and *T. militinae* include a selection (line 8.1) which harbors *T. militinae* segments incorporated within chromosomes 1A, 1B, 2A, 4A, 5A, 5B and 7A (Jakobson et al. 2006). Among other traits introduced from *T. militinae,* line 8.1 exhibits a marked improvement in the level of powdery mildew resistance expressed at both the seedling and adult plant stages. According to a QTL analysis, the genetic basis of this resistance is dominated (respectively, 33% and 54% of the variance shown by seedlings and adult plants) by a genes within a segment located near the distal end of chromosome arm 4AL, referred to as *QPm.tut-4A* (Jakobson et al. 2006). When present in a cv. Tähti background, the resistance decreases the number of secondary haustoria formed by the pathogen and enhances host cell apoptosis (Islamov et al. 2015). The gene (or genes) present within *QPm.tut-4A*, which acts in a race non-specific manner (Jakobson et al. 2012), is located in *T. militinae* itself on chromosome 7G (Abrouk et al. 2017). In a hybrid between wild-type bread wheat and the introgression line, little recombination occurs between the *T. militinae* segment and the segment of chromosome 4AL which has been replaced in line 8.1. However, it has been possible to define the genetic length and position of the segment to a 2.5 cM interval (Jakobson et al. 2012). Here, the focus was to isolate the gene(s) which determine the *QPm.tut-4A* resistance. By deploying a range of genetic and genomic strategies, it has proved possible to identify a small number of candidate genes and to characterize and contrast the sequences which originated from chromosome 7G with those which they replaced on 4A.

Materials and methods

Plant materials and mapping populations

The Jakobson et al. (2012) mapping population comprised 98 F_2 progeny bred from the cross cv. Chinese Spring (CS) × line 8.1; this was extended for the purpose of increasing the level of resolution by self-pollinating F_2 individuals which were heterozygous for *QPm.tut-4A* through to the F_5. In addition, a second mapping population (hereafter referred to as the "*ph1* population") was created by crossing the CS *ph1b* mutant (Sears 1977) with T312.30.38.16, a derivative of introgressive line 8.1 in which the only *T. militinae* segment present was the one harboring *QPm.tut-4A*. In order to derive a parental plant which was both homozygous for the *ph1b* allele and heterozygous for the segment harboring *QPm.tut-4A*, the resulting F_1 hybrid was back-crossed to the *ph1b* mutant. Marker-assisted selected BC_1F_1 individuals were then allowed to self-pollinate. Selection for the *ph1b* allele was performed using a multiplex PCR assay involving the markers AWJL3, PSR128, PSR2120 and PSR574 (Roberts et al. 1999), while the *QPm.tut-4A* harboring segment was marked by *owm76* and *owm96*. The subsequent generations (BC_1F_3 and further) were not employed for this study due to their low viability and fertility presumably caused by extensive chromosomal rearrangements associated with the action of *ph1*. Doubled haploid line DH397 containing the *T. militinae* resistance locus only on 4AL was derived from cross Tähti × 8.1 (Jakobson et al. 2012). A bacterial artificial chromosome (BAC) library was constructed from the DNA of 4AL telosomic chromosome flow sorted from 4AL ditelosomic line carrying the 4A *T. militinae* introgression (Jakobson et al. 2012). Grain of the CS aneuploid stocks nullisomic 4A-tetrasomic 4B and nullisomic 4A-tetrasomic 4D, required to validate the chromosome specificity of newly developed markers, was provided by the National BioResource Centre (Kyoto, Japan). A doubled haploid line (DH81) carrying the same *T. militinae* translocations which determined the powdery mildew resistance as line 8.1 (Jakobson et al. 2012) served as a further control. Finally, the bread wheat cv. Kanzler was used a susceptible host in experiments involving powdery mildew inoculations.

Phenotyping for powdery mildew resistance

Seedling resistance to powdery mildew was scored using an assay based on detached first seedling leaves of 10-day-old plants. Each leaf was cut into four segments, each of which was then laid in a Petri dish containing 0.6% agar supplemented with 0.35% *w/v* benzimidazol. The leaf segments were inoculated with four different isolates (2.1, 9.8, 13 and 14), as described by Jakobson et al. (2012). The Petri dishes were held at 17.5 °C, and the response was evaluated after 10 days using the 0–9 scale devised by Lutz et al. (1992). For each self-pollinated recombinant line selected from F_2-derived F_{3-5} families, up to 30 progeny homozygous in the *QPm.tut-4A* region were phenotyped. The resistance status of each recombinant line was verified by comparing its progeny scores with scores of 12–16 progeny of homozygous nonrecombinant sister line selected from the same self-pollination.

The development of markers used for high-density genetic mapping

The informativeness of a potential marker was first assessed via an *in silico* inspection of the chromosome-specific survey sequences of 4AL-7G (Abrouk et al. 2017) and the chromosome survey sequences (CSS) of $4AL^{CS}$ generated by IWGSC (2014). The choice of sequences was governed by the need to saturate the genetic map of the *QPm.tut-4A* region, so it was based on a virtual gene order of the chromosome 4A represented by GenomeZippers (Abrouk et al. 2017, Hernandez et al. 2012). 4AL CSS scaffolds obtained from the GenomeZipper delimited by the markers flanking the *QPm.tut-4A* segment were aligned with 4AL-7G sequence scaffolds using the BlastN algorithm (Altschul et al. 1990). Only low-copy sequences associated with a nucleotide identity of > 95% were considered for marker development, and those containing short indels were preferred. Where no indel was identifiable, markers were based on single-nucleotide polymorphisms using the cleaved amplified polymorphic sequence approach (Michaels and Amasino 1999). PCR primers were designed using Primer3 software (Untergasser et al. 2012). To ensure specificity for 4AL, given the presence of homoeologous sequence on 7AS and 7DS, primers were positioned to ensure the presence of a variant nucleotide close to 3′ end in the homoeologous sequences. Finally, the specificity of the putative amplicons was verified by a BlastN search against the whole wheat genome CSS (IWGSC 2014). A second marker discovery strategy profited from a BAC library-based, established CS 4AL-specific physical map (IWGSC 2018, URGI; urgi.versailles.inra.fr/). BAC clones making up the set of contigs which covered the *QPm.tut-4A* region were selected from the minimal tiling path and sequenced. Subsequently, sequence scaffolds positioned at the target location were employed for marker development based on the first strategy. A final strategy designed to extend the saturated portion of the *QPm.tut-4A* region beyond what was achievable using the first two strategies was based on the sequence of the *T. dicoccoides* 7AS region (Avni et al. 2017) and on the CS reference sequence WGA v0.4 (www.wheatgenome.org); this also was informative for identifying additional $4AL^{CS}$ scaffolds for targeted marker development. All primer pairs were tested on a template of CS, DH81, a CS/line 8.1 derivative heterozygous for *QPm.tut-4A*, nullisomic 4A-tetrasomic 4B and nullisomic 4A-tetrasomic 4D and DNA from flow-sorted chromosome arms $4AL^{CS}$ and 4AL-7G amplified according to Šimková et al. (2008). The methods used for PCR amplification and electrophoretic separation are given in the following section, and the primer sequences and associated information are summarized in Table S1. Where the physical position of a marker was uncertain, it was determined by screening 62 three-dimensional pools prepared from the minimal tiling path of the CS 4AL-specific physical map.

Genetic mapping

DNA was extracted using Agencourt® Genfind® v2 magnetic beads (Beckman Coulter Life Sciences, Indianapolis, IN, USA) on a Beckman Coulter® Biomek® NX^P workstation, as described by Ivaničová et al. (2016). Markers *owm82* and *Xgwm160* (Röder et al. 1998) were used to select lines in which a recombination event had occurred in the region of the introgression segment in the mapping populations. The full set of markers was applied to genotype those arising from the CS × line 8.1 population, while a subset of 19 selected markers was applied to those arising from the *ph1* population. Each 15 µL PCR contained 0.01% (*w/v*) o-cresolsulphonephtalein, 1.5% (*w/v*) sucrose, 0.2 mM of each dNTP, 0.6 U Taq DNA polymerase, 1 µM of each primer, 10 mM Tris-HCl, 50 mM KCl, 1.5 mM $MgCl_2$ and 0.1% (*v/v*) Triton X-100. The template comprised either 10–20 ng genomic DNA or 5 ng DNA amplified from $4AL^{CS}$ or 4AL-7G. The reaction conditions consisted of an initial denaturation step of 95 °C/5 min, followed by 40 cycles of 95 °C/30 s, an optimized annealing temperature (Table S1) for 30 s and 72 °C for 30 s per 500 bp amplicon length; the reactions were completed with an elongation step of 72 °C/5 min. Cleaved amplified polymorphic sequence assays were completed with a digestion using the appropriate restriction endonuclease. The amplicons were electrophoretically separated through 4% non-denaturing polyacrylamide gels and visualized by ethidium bromide staining.

Construction of the physical map

Sequence scaffolds taken from IWGSC RefSeq v1.0 (IWGSC 2018) were used to span the genomic region of CS 4AL replaced by the *QPm.tut-4A* segment in line 8.1. To obtain the corresponding sequence from *T. militinae*, a chromosome walking approach was initiated. A chromosome-specific BAC library designated TaaP-mt4ALhA (www.olomouc.ueb.cas.cz/dna-libraries/cereals) was constructed from flow-sorted 4AL-7G chromosome arms according to the Šimková et al. (2011) protocol. BAC library plate pools were generated by mixing 40 µL of the bacterial culture from each well of a given plate, centrifuging (10 min at 2700 g), suspending the precipitate in 0.5 mL TE and boiling for 30 min. The suspension was then re-centrifuged (60 min at 2700 g), and a 450-µL aliquot of the supernatant was diluted 100-fold. The subsequent PCRs used as template a 1.5-µL aliquot of the diluted plate pool DNA. The plate pools were screened with the markers *owm169* and *owm228,* together with other codominant markers located within the CS 4AL segment replaced by *QPm.tut-4A* (*owm156*, *owm136*, *owm221*, *owm209*, *owm227*, *owm236*, *owm139* and *owm235*). Row and column pools of positive plates were prepared and used in the same way as plate pools. Selected BAC clones were sequenced on a MiSeq instrument (Illumina Inc., San Diego, CA, USA) using a Nextera DNA Library Prep Kit (Illumina) according to the manufacturer's protocol. Paired-end reads were assembled by Ray software (Boisvert et al. 2010). Insertion site-based polymorphism or site-specific presence/absence variation markers were developed from the ends of sequenced BAC clones and used for a further round of BAC library screening. The procedure was iterated until the physical map had been assembled.

The identification and sequence analysis of candidate genes

The gene content along the *QPm.tut-4A* segment on 4AL-7G was annotated using the TriAnnot pipeline (Leroy et al. 2011, www6.inra.fr/decodage/TriAnnot). Predicted genes were subjected to a BlastP search against the set of non-redundant protein sequences (www.ncbi.nlm.nih.gov/BLAST), and conserved domains of putative proteins were annotated using the NCBI Conserved Domain Database (Marchler-Bauer et al. 2017) and the MOTIF search tool (www.genome.jp/tools/motif/) based on the Pfam database (Finn et al. 2016). Searches for homologs and the comparison of candidate gene sequences were based on the BlastN algorithm, applying a threshold of 90% identity and 60% coverage.

Reverse transcription PCR

Segments of the first leaf of 10-day-old seedlings of the doubled haploid line DH397 and line T312.30.38.16 were collected 0-, 24- and 48-h post-inoculation with powdery mildew and snap frozen in liquid nitrogen. Leaf segments from three independent inoculations were bulked for the purpose of RNA extraction, and negative control samples were formulated from non-inoculated leaf segments. The leaf tissue was homogenized using ball mill MM 301 (Retsch, Haan, Germany) with three 3 mm tungsten beads at 30 Hz for 45 s, and total RNA was extracted using a miRNeasy Mini Kit (Qiagen Inc., Hilden, Germany) according to the manufacturer's protocol. The synthesis of the first cDNA strand was achieved using a Transcriptor High Fidelity cDNA Synthesis Kit (Roche Life Sciences, Indianapolis, IN, USA) based on an anchored-oligo$(dT)_{18}$ primer. Gene-specific primers were designed for all 12 candidate genes such that the amplicons derived from gDNA differed in length from those derived from cDNA (Table S2). A portion of the *Actin* gene (GenBank accession number AB181991.1) was used as the reference sequence. The procedures used for PCR and electrophoretic separation were as described above.

Results

High-resolution genetic mapping around QPm.tut-4A

A set of 102 new genetic markers was developed to saturate the genetic map in the region of *QPm.tut-4A* (Figs. 1, 2, Table S1). Majority of these were developed using GenomeZipper and 4AL-specific BAC clones. Besides markers *owm16* and *owm39*, polymorphism was inferred using the 4AL-7G and $4AL^{CS}$ survey sequences (Abrouk et al. 2017, IWGSC 2014). Segregation patterns observed in the extended CS × 8.1 mapping population were used to establish marker order. The genotyping of 8425 individuals with respect to *owm82* and *Xgwm160* revealed 30 new recombination events. Testing these recombinant individuals for their reaction to the four powdery mildew isolates showed that the resistance phenotype was fully correlated with the presence of the *T. militinae* segment—the disease scores for the four isolates were, respectively, 0.5, 0.4, 0.1, 1.1 in the presence of the segment, and 3.3, 2.5, 2.6, 3.8 in its absence. The net effect of the mapping was reducing the genetic length of the segment harboring *QPm.tut-4A* to 0.012 cM (Fig. 1). The *ph1* population was derived from the self-pollination of 22 (out of 107 screened) BC_1F_1 ([line 8.1 × *ph1b*] × *ph1b*) selections which were simultaneously homozygous for the *ph1b* allele and heterozygous for the *QPm.tut-4A* segment. The genotyping of the resulting 1255

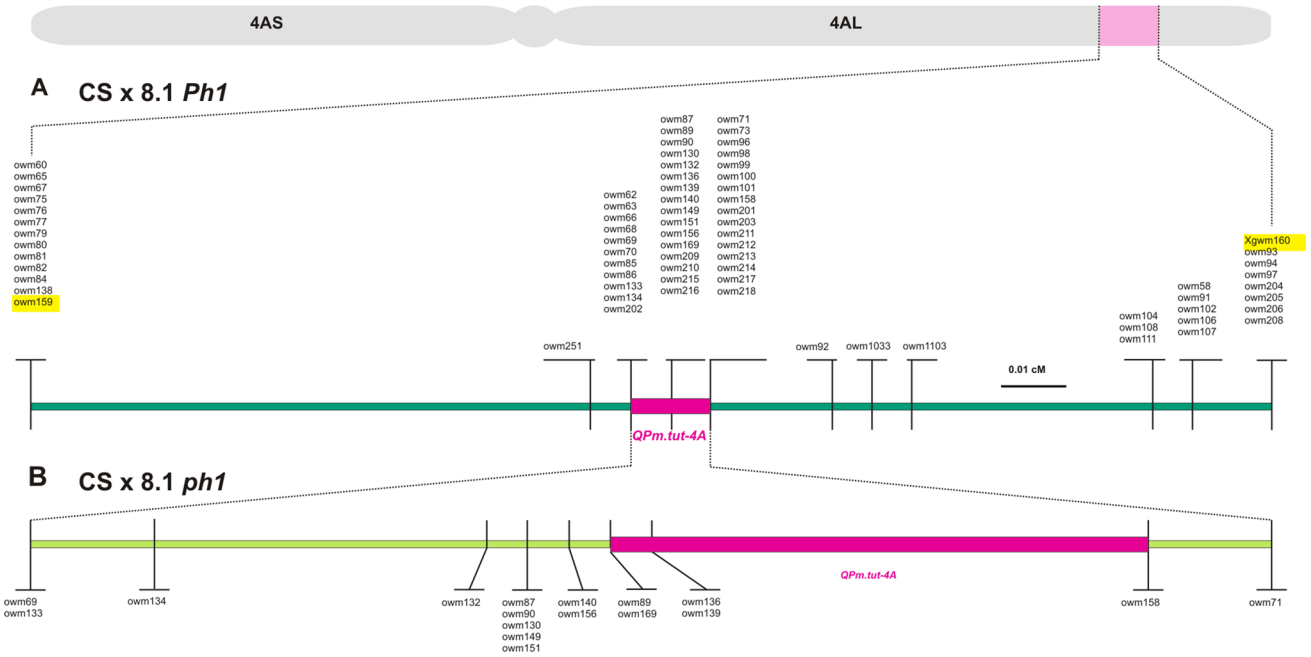


Fig. 1 A high-density genetic map of the region harboring *QPm.tut-4A*. **a** Marker segregation displayed by 8519 progeny bred from the CS × 8.1 mapping population (including the F_2 generation described at Jakobson et al. 2012) defines the genetic length of the *QPm.tut-4A* segment (marked in red) to 0.012 cM, lying within a 0.18 cM window within chromosome 4AL. The flanking markers used for identification of recombination events within the segment are highlighted in yellow. **b** The level of mapping resolution achieved was enhanced through the use of the *ph1* mapping population, which provided an additional 30 new recombination events, thereby splitting the *QPm.tut-4A* region into eight subregions

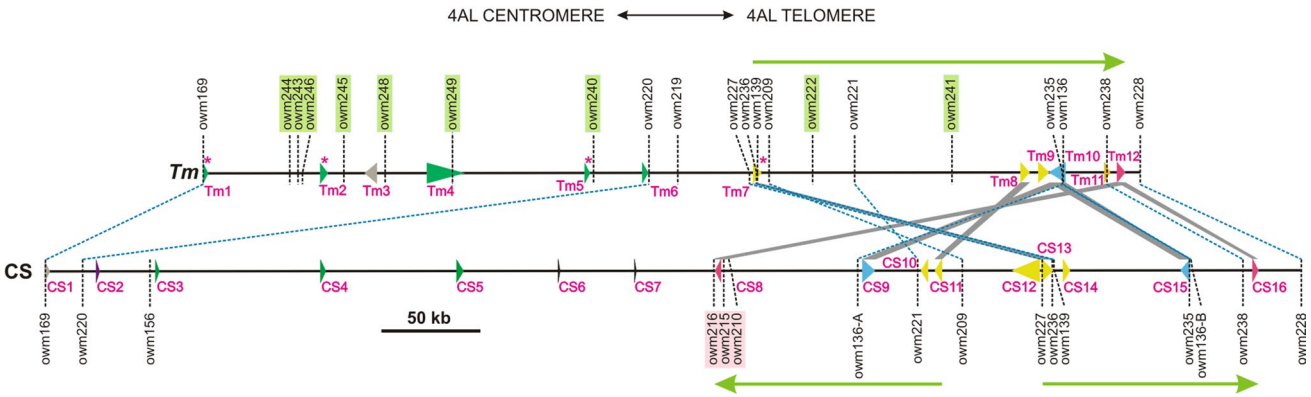


Fig. 2 Physical maps showing the gene content of the 480.2 kbp *QPm.tut-4A* segment inherited from *T. militinae* and the 640.8 kbp segment of chromosome 4AL replaced by the introgression in line 8.1. The predicted genes (*Tm1–Tm12* and *CS1–CS16*) are depicted by arrowheads indicating their orientation. The size of the arrowheads is proportional to the length of the gene, and their color reports the presence of conserved domains: green: CC-(LRR)-NB-ARC; yellow: PGG; orange: ANK and PGG; blue: LRR/MAL/PK; red: patatin-like; black: PMEI-like; and gray: uncharacterized protein. *Tm* and *CS* genes linked by gray lines share sequence identity in their coding region of at least 90% at a coverage of ≥ 60% (a coverage < 80% applied only to the putative pseudogene *CS15*). *Tm* genes which produced a detectable transcript are indicated by a red asterisk. Shared markers are linked by a blue dashed line, those marking *T. militinae* but not CS genomic sequence, or *vice versa,* are highlighted in, respectively, light green and light red. The *T. militinae* region bounded by *owm227* and *owm228* was present in CS as two tandemly arranged, tail-to-tail orientated copies, as indicated by the green arrows

BC_1F_2 progeny revealed 155 recombination events between *owm82* and *Xgwm160*, equivalent to a 33-fold increase in recombination as compared to the rate observed in the presence of *Ph1*. The 0.012 cM *QPm.tut-4A* region was thereby split into eight subregions (Fig. 1b).

The physical map of the QPm.tut-4A segment

Six additional markers for the *QPm.tut-4A* region were developed by inspection of the three CS reference sequence IWGSC WGA v0.4 scaffolds (134864, 108602 and 47761) (www.wheatgenome.org) selected on the basis of markers lying within the *owm169–owm158* interval (Table S1). The *owm228* marker was derived from same scaffold (134864) which contained the *QPm.tut-4A* flanking marker *owm169*. The application of these markers to the set of recombinant segregants defined the physical length of the CS sequence replaced by the *QPm.tut-4A* introgression to 640.8 kbp (Fig. 2). A 4AL-7G-specific BAC library was constructed in an effort to acquire sequences in the *QPm.tut-4A* region of line 8.1; the mean insert size harbored by the 43,008 clones generated provided a 6.8-fold coverage of the 4AL-7G chromosome arm (www.olomouc.ueb.cas.cz/dnalib/taapmt4alha). Pools of the set of 112 plates were assembled to perform a PCR screen based on the flanking markers *owm169* and *owm228*, along with eight codominant markers (*owm156*, *owm136*, *owm221*, *owm209*, *owm227*, *owm236*, *owm139* and *owm235*) mapping within the region (Fig. 2). One to seven positive BAC clones per marker were sequenced so that their end sequences could be used for the next round of chromosome walking-based marker development: 17 new markers were designed in this way (Table S1). In total, 26 BAC clones were sequenced to provide a complete sequence of the *QPm.tut-4A* region in line 8.1: The physical length of the segment bounded by *owm169* and *owm228* was 480.2 kbp.

Annotation and comparative analysis of the QPm.tut-4A segment

The *T. militinae* region was 25% shorter than the CS one (480.2 vs. 640.8 kbp). Gene annotation suggested that the *QPm.tut-4A* segment harbored 12 high-confidence (HC) protein-encoding genes (denominated *Tm1–Tm12*, Fig. 2, Table 1), while the 4AL region of CS harbored 16 HC genes (*CS1–CS16*, Table 2). The intergenic regions displayed no sequence similarity. A comparison of the *T. militinae* and *CS* genes based on a threshold of 90% identity and 60% coverage implied that only four of the *T. militinae* genes (*Tm7*, *Tm8*, *Tm10* and *Tm12*) shared an appreciable level of homology with members of the *CS* set (*CS8*, *CS9*, *CS11*, *CS13*, *CS15* and *CS16*) (Table 1, Fig. 2). The sequence of

Table 1 Predicted candidate genes present in the *QPm.tut-4A* segment introgressed from *T. militinae* into line 8.1

No.	Putative conserved protein domains	No. of exons	Length of protein sequence (aa)	CS homologous genes in the QPm.tut-4A region[a]	Identity (%)	Coverage (%)	RNA expression[b]
Tm1	CC; NB-ARC	2	225	–	–	–	+
Tm2	CC; LRR; NB-ARC	2	753	–	–	–	+
Tm3	–	3	253	–	–	–	–
Tm4	CC; LRR; NB-ARC	7	894	–	–	–	–
Tm5	CC; LRR; NB-ARC	2	557	–	–	–	+
Tm6	CC; LRR; NB-ARC	2	1017	–	–	–	–
Tm7	5× PGG	2	963	TraesCS4A01G450100.2 (CS13)	97	99	+
Tm8	5× PGG	3	764	TraesCS4A01G449900 (CS11)	90	98	–
Tm9	4× PGG	2	849	–	–	–	–
Tm10	LRR; Malectin; Pkinase	23	913	TraesCS4A01G449700 (CS9) TraesCS4A01G450300 (CS15)	97 98	84 60	–
Tm11	Ankyrin repeats; PGG	4	636	–	–	–	–
Tm12	Patatin-like phospholipase	5	424	TraesCS4A01G449600 (CS8) TraesCS4A01G450400 (CS16)	98 97	100 100	–

CC coiled coil, *NB-ARC* nucleotide-binding adaptor shared by APAF-1, certain *R* gene products and CED4, *LRR* leucine-rich repeat, *PGG* proline-glycine-glycine (domain named for this highly conserved sequence motif found at its start), *Pkinase* protein kinase

[a]Acronym for each gene used in the body of this manuscript and Fig. 2 is provided in the parentheses

[b]Detected in first leaves of DH397 and T312.30.38.16 seedlings

Table 2 Gene content of the segment of CS chromosome arm 4AL replaced by the *QPm.tut-4A* segment introgressed from *T. militinae* into line 8.1

No.	Official designation[a]	Putative conserved protein domains	No. of exons	Length of protein sequence (aa)
CS1	TraesCS4A01G448900	CC	3	313
CS2	TraesCS4A01G449000	Transferase	1	472
CS3	TraesCS4A01G449100	CC; NB-ARC	2	353
CS4	TraesCS4A01G449200	NB-ARC; LRR	2	838
CS5	TraesCS4A01G449300	CC; LRR; NB-ARC	3	1008
CS6	TraesCS4A01G449400	Plant invertase/pectin methylesterase inhibitor	1	203
CS7	TraesCS4A01G449500	Plant invertase/pectin methylesterase inhibitor	1	206
CS8	TraesCS4A01G449600	Patatin-like phospholipase	5	424
CS9	TraesCS4A01G449700	LRR; Malectin; Pkinase	21	867
CS10	TraesCS4A01G449800	5× PGG	2	993
CS11	TraesCS4A01G449900	5× PGG	2	1020
CS12	TraesCS4A01G450000	4× PGG	3	761
CS13	TraesCS4A01G450100.2	5× PGG	3	984
CS14	TraesCS4A01G450200	5× PGG	2	1024
CS15	TraesCS4A01G450300	LRR; Malectin	15	549
CS16	TraesCS4A01G450400	Patatin-like phospholipase	5	424

[a]The International Wheat Genome Sequencing Consortium RefSeq v1.0 annotation (IWGSC 2018)

the proximal ends of the CS and line 8.1 regions (beyond *owm169*) were quite distinct with respect to both their sequence and gene content: The segment from *owm169* to *owm220* was 228 kbp in line 8.1, but only 19 kbp in CS, and they contained, respectively, six (*Tm1–Tm6*) and one (*CS1*) genes. The central part of the segment (*owm220–owm227* in line 8.1 and *owm220–owm216* in CS) was longer in CS (322 vs. 51 kbp) and contained no predicted gene in line 8.1 and *CS2–CS7* in CS. The distal region (*owm227–owm228* in line 8.1 and *owm216–owm228* in CS) was, respectively, of length 200 kbp (genes *Tm7–Tm12*) and 299 kbp (*CS8–CS16*) (Fig. 2). The proximal segment harboring *Tm1–Tm6* was not represented on CS chromosome 4A. However, four of these genes (*Tm1*, *Tm2, Tm5* and *Tm6*) each have a strong homoeolog (Table 3) lying in reverse orientation within a ~410 kbp region of chromosome arm 7AS situated about 17.4 Mbp from the telomere. The position of the *QPm.tut-4A* region is 41.05 Mbp distant from the 4AL telomere. The distal part of the region (*owm227–owm228*) has been duplicated in CS, and the gene content has been differentiated. The two copies are arranged tandemly in a tail-to-tail orientation, resulting in the non-syntenic location of the *owm227* marker (Fig. 2).

Functional characterization of the candidate genes for powdery mildew resistance

The five genes *Tm1*, *Tm2* and *Tm4–Tm6* each encoded a coiled-coil (CC) domain and a nucleotide-binding (NB) domain, thereby being members of the NB-ARC family (van der Biezen and Jones 1998). *Tm2*, *Tm4* and *Tm6* also encode a leucine-rich repeat (LRR) domain, indicating them as members of the disease resistance-associated NLR gene family (Ye and Ting 2008). The shorter length of both *Tm1* and *Tm5* (Table 1) implies that both are incomplete

Table 3 Homoeologs on CS chromosome arm 7AS of the NLR family genes present in the *QPm.tut-4A* segment introgressed from *T. militinae* into line 8.1

No.	CS homoeolog on 7AS	Identity (%)	Coverage (%)	Putative conserved protein domains of homoeologs	Length of protein sequence (aa)	Length difference relative to *T. militinae* homoeolog (aa)
Tm1	TraesCS7A01G038800	95	100	CC; NB-ARC	452	+227
Tm2	TraesCS7A01G039300	97	100	CC; LRR; NB-ARC	972	+219
Tm4	–	–	–	–	–	–
Tm5	TraesCS7A01G03910	98	100	CC; LRR; NB-ARC	1074	+517
Tm6	TraesCS7A01G039400	93	79	CC; LRR; NB-ARC	875	−142

genes. *Tm10* is predicted to encode a protein harboring an LRR/malectin/protein kinase (LRR/MAL/PK) domain, also shared by a number of plant disease resistance gene products (Sekhwal et al. 2015); this gene lies in the distal part of the introgression segment and shares homology with both *CS9* and the truncated *CS15* (Fig. 2). The sequences of the other six *Tm* genes have little or no connection with disease resistance: *Tm7–Tm9* each encode multiple PGG domains, *Tm11* features ankyrin repeats, while *Tm12* encodes a protein belonging to the patatin-like phospholipase family (Table 1). Four of the 12 *Tm* genes (*Tm1*, *Tm2*, *Tm5* and *Tm7*) produced a detectable level of mRNA, and three of them (*Tm1*, *Tm2* and *Tm5*) encode NLR proteins: *Tm2* appears to be the only one of these which is intact.

Discussion

The introduction into the bread wheat gene pool of genes harbored by species belonging to its secondary and tertiary gene pools represents an attractive strategy for broadening the genetic diversity of a crop which has been intensively bred for over a century. However, the success of the strategy has been not infrequently limited by the simultaneous introgression of linked genes which are deleterious to either productivity and/or product quality. This phenomenon of linkage drag results from the suppression of recombination around an introgressed segment, a phenomenon which also hampers positional cloning, since it magnifies the ratio between physical (in bp) and genetic (in cM) distance and prevents efficient high-density mapping. The *T. militinae* segment harboring *QPm.tut-4A* suffers from exactly this problem. Recent significant progress in the next-generation sequencing technologies allowed development of two new gene cloning approaches called MutRenSeq (Steuernagel et al. 2016) and MutChromSeq (Sánchez-Martín et al. 2016) which can bypass the high-density map construction in gene cloning process. The MutRenSeq approach employing exome capture and sequencing focusses on identification of simultaneous deleterious mutations in single NB-LRR-like gene within mutants with lost resistance. The MutChromSeq uses similar approach consisting in identification of knockout mutations in single gene within mutant lines, but it requires flow sorting of respective chromosomes from all mutant lines and their sequencing. This means that only single dominant major-effect genes can be identified by these approaches. However, the *QPm.tut-4A* locus is associated with race non-specificity and incomplete resistance which suggests the resistance may be encoded by gene different from major-effect *R* genes (predominantly NB-LRR-like genes) or by more than one gene (Jakobson et al. 2006, 2012). The MutRenSeq and MutChromSeq approaches are therefore not feasible for the *QPm.tut-4A* gene/genes cloning. Fortunately, recent advances in DNA technology and the development of sophisticated genomic resources have substantially eased the processes of constructing a high-density map and of acquiring relevant sequence from both the donor and the recipient genomes.

High-density mapping

According to the Wheat-Composite2004-4A map (wheat.pw.usda.gov/GG3), the markers flanking *QPm.tut-4A* (*Xwmc232* and *Xgwm160*) are separated by ~9 cM, which led Jakobson et al. (2012) to suggest that a plant heterozygous for the introgression experiences suppression of recombination in the region, as has been documented for a number of other bread wheat introgression segments (Bariana et al. 2001; Järve et al. 2000; Jia et al. 1996; Lukaszewski 2015). Several genomic resources developed in CS were exploited here to saturate the region with markers in order to fine map the resistance locus, and these greatly enhanced the efficiency with which such markers could be elaborated, especially compared to conventional approaches to marker discovery such as microsatellite screening (Röder et al. 1998) or searching for sequence polymorphism in expressed sequence (e.g. Valárik et al. 2006). The GenomeZipper and chromosome-specific survey sequences (Abrouk et al. 2017, Hernandez et al. 2012, IWGSC 2014) were particularly effective in this context, but use was also made of sequence scaffolds developed in the tetraploid wheat *T. dicoccoides* (Avni et al. 2017) and the BAC contigs defining the chromosome 4AL physical map (IWGSC 2018). Initially, these resources allowed the genetic length of the target to be narrowed to just 0.012 cM (Fig. 1). In a fully homologous situation, such as occurs on wheat chromosome 3B, the mean ratio between cM and Mbp in distal regions varies from 0.60 to 0.96 (Choulet et al. 2014). If the recombination between the introgressed segment and the unaltered bread wheat region was unimpeded, the physical length of the segment would have been estimated to be no longer than 20 kbp. The fact that the segment's length was measured in hundreds of kbp demonstrated that there was a substantial localized suppression of recombination.

Structural divergence and recombination in the QPm.tut-4A segment

The pairing of homoeologs in hexaploid wheat is strongly restricted by the action of the *Ph1* locus (Riley and Chapman 1958). The intention of creating the *ph1* population was to induce a higher rate of recombination between the introgression containing *QPm.tut-4A* and its presumed homoeologous segment of CS 4AL, since the rate of recombination achieved in the presence of *Ph1* was $< 0.4\%$ (31 out of 8519) within the *owm82-Xgwm160*

region (Fig. 1): Although this represented a very small genetic distance (0.18 cM), the suppression of pairing/recombination meant that the physical distance involved was potentially rather large. The effect was enhancing the rate of recombination by 33-fold. A comparable elevation in the recombination rate induced by removal of the *Ph1* locus has been reported in a variety of interspecific hybrids (Lukaszewski 1995, Luo et al. 2000). The availability of the genomic sequence of CS (IWGSC 2018) made it possible to identify that the length of the 4AL segment (0.012 cM) replaced by the *owm169-owm228 QPm.tut-4A* locus in line 8.1 was 640 kbp (Figs. 1, 2) and that it harbors 16 predicted genes (Table 2). The inferred relationship between genetic and physical distance in the segment was therefore only 0.019 cM per Mbp, which is much lower than the ratio (0.60–0.96 cM per Mbp) obtained in the distal region of chromosome 3B and is even below the ratio associated with the 3B centromeric region (0.05 cM per Mbp) in which recombination is known to be repressed (Choulet et al. 2014). This major suppression of recombination reflects the lack of homology between the native wheat and the introgressed *QPm.tut-4A* locus. The length of the introgressed *QPm.tut-4A* locus (*owm169–owm228*) was 480 kbp (Fig. 2), and it harbored 12 predicted genes (Table 1), eight of which lacked a homolog in the CS segment (Table 1, Fig. 2). The CS homologs of the other four genes (*Tm7*, *Tm8*, *Tm10* and *Tm12*) lay in a duplicated segment, and one of the duplicated copies was present in inverted orientation. The region was further disrupted by a number of indels, and all these changes suggest a high level of evolutionary dynamics of the wheat genome. While loss of synteny and sequence divergence are commonplace between homoeologous genomes (Saintenac et al. 2013; Wicker et al. 2003), they can also feature in comparisons made between homologous genomes of different hexaploid wheat cultivars (Mago et al. 2014; Tsõmbalova et al. 2016).

The introgressed *QPm.tut-4A* locus included three full-length and two truncated NLR family genes (Table 1). This class of genes is frequently arranged in clusters, which is thought to reflect the outcome of duplication events followed by sequence divergence (Michelmore and Meyers 1998). The coding sequences of the *T. militinae* NLR-like genes shared only a moderate to a high (73–89%) level of identity at the nucleotide level. Their putative homoeologs present on chromosome arm 7AS of CS displayed a similar level of sequence relatedness (71–85%). This chromosomal location confirms the conclusion of Abrouk et al. (2017) that in *T. militinae* itself, the *QPm.tut-4A* segment is present on chromosome 7G. The G genome donor is thought to be a member of the *Sitopsis* section of the genus *Aegilops,* as is also the donor of the bread wheat B genome (Gornicki et al. 2014). The bread wheat chromosome 4A itself is a restructured chromosome composed of a mosaic of segments derived from 4AL, 5AL and 7BS (Devos et al. 1995, Hernandez et al. 2012), and the *QPm.tut-4A* introgression appears to lie within a part of this region which originated from 7BS, consistent with its transfer following meiotic pairing.

The potential function of the QPm.tut-4A candidates

The resistance to powdery mildew associated with the presence of the *QPm.tut-4A* segment was race non-specific and in a cv. Tähti background, accounted for 40% of the variation in resistance (Jakobson et al. 2012). To date, only few wheat genes associated with APR have been isolated, so unlike the case for race-specific seedling resistance genes, many of which belong to the NLR family, it is not clear what functionality the product of such a gene might have. The *Lr34/Yr18/Pm38* gene, which provides protection against three distinct foliar pathogens, has been shown to encode an ABC transporter (Krattinger et al. 2009), while *Lr22a*, which confers broad-spectrum APR to leaf rust, encodes an NLR-like protein (Thind et al. 2017), as does the rice NLR family *Pb1* gene against panicle blast (Hayashi et al. 2010). Assuming that one of the three NLR family genes *Tm2*, *Tm4* and *Tm6* represents the most likely candidate for the *QPm.tut-4A* resistance, the possible basis of its resistance being race non-specific needs to be explored. In one scenario, it may be that either two or even all three of the *Tm* genes, which are each individually race-specific, act together to confer apparent race non-specificity. Alternatively, it is possible that one of the three genes is a "defeated" major resistance gene which has retained some residual broad-spectrum effect, as has been demonstrated for some other defeated major resistance genes (Li et al. 1999). The latter hypothesis is probably the more plausible, given that the only gene for which a mRNA was detected was *Tm2,* although surprisingly it was possible to detect transcript of the two NLR-like pseudogenes *Tm1* and *Tm5*. However, the lack of an LRR domain in the *Tm2*-encoded protein implies that it would be difficult for its product to recognize the pathogen's avirulence signal. An additional candidate is represented by *Tm10* which encodes a protein containing an LRR, a MAL and a PK domain. A barley protein of this domain composition (HvLEMK1) is known to mediate non-host resistance to powdery mildew, while its wheat ortholog acts to enhance the level of wheat host resistance to powdery mildew (Rajaraman et al. 2016). A comparison of the TaLEMK1 and the *Tm10* protein sequences showed that they share only a 34% level of identity; this rather low level of homology, combined with the observations that the wheat *LEMK1* homoeologs map to the group 5 chromosomes and that no *Tm10* transcript was detected, rules out the possibility that the *QPm.tut-4A* resistance is conferred by *Tm10*. None of the other *Tm* genes

encode a product which has been directly associated with disease resistance to date. However, *Tm12*—which encodes a protein belonging to the patatin-like phospholipase family—remains a candidate since the patatin-like protein AtPLP2 has been shown to represent a component of the cell machinery delivering apoptosis, and therefore makes a contribution toward resistance against an obligate biotroph (La Camera et al. 2009).

Author contribution statement KJ, JD and MV designed the study; MV, EJ, MŠ, HŠ, JŠ and IJ were responsible for marker development, genotyping, BAC library construction, chromosome sorting, data analysis and the construction of the mapping populations. HP performed the phenotypic evaluation and IJ the statistical analysis. The bioinformatics analyses were conducted by EJ and MA, who also contributed to data interpretation. Other experiments were conducted by EJ. The manuscript was drafted by EJ and MV, and all the authors contributed to its editing and proofreading.

Acknowledgements The technical assistance of Marie Seifertová, Radka Tušková, Eva Jahnová and Helena Tvardíková is highly appreciated. This research was financially supported by the Czech Republic Ministry of Education, Youth and Sports (Award LO1204 from the National Program of Sustainability I), by the Czech Science Foundation (Award 18-11688S), the Czech Republic Ministry of Agriculture (Award QK1710302) and by an IUT 193 Grant from the Estonian Ministry of Education and Research. The authors thank the International Wheat Genome Sequencing Consortium (IWGSC) for pre-publication access to IWGSC RefSeq v1.0, the National BioResource Centre (Kyoto, Japan) for providing the CS aneuploid stocks and S. Reader (John Innes Centre, Norwich, UK) for the gift of grain of the *ph1b* mutant. Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum provided under the program "Projects of Large Research, Development and Innovation Infrastructures" (CESNET LM2015042) is greatly appreciated.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author confirms that no conflict of interest applies.

Data availability The DNA sequence of the *T. militinae*-derived *QPm. tut-4A* introgression segment has been deposited in GenBank (Accession No. MG672525). Supporting data are available from the corresponding author upon reasonable request.

Funding The study was financially supported by the grants listed in the Acknowledgements section.

References

Abrouk M, Balcárková B, Šimková H, Komínková E, Martis M, Jakobson I, Timofejeva L, Rey E, Vrána J, Kilian A, Järve K, Doležel J, Valárik M (2017) The *in silico* identification and characterization of a bread wheat/*Triticum militinae* introgression line. Plant Biotechnol J 15:249–256

Altschul SA, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Avni R, Nave M, Barad O, Baruch K, Twardziok SO, Gundlach H, Hale I, Mascher M, Spannagl M, Wiebe K, Jordan KW, Golan G, Deek J, Ben-Zvi B, Ben-Zvi G, Himmelbach A, MacLachlan RP, Sharpe AG, Fritz A, Ben-Davis R, Budak H, Fahima T, Korol A, Faris JD, Hernandez A, Mikel MA, Levy AA, Steffenson B, Maccaferri M, Tuberosa R, Cattivelli L, Faccioli P, Ceriotti A, Kashkush K, Pourkheirandish M, Komatsuda T, Eilam T, Sela H, Sharon A, Ohad N, Chamovitz DA, Mayer KFX, Stein N, Ronen G, Peleg Z, Pozniak CJ, Akhunov ED, Distelfeld A (2017) Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. Science 357:93–97

Bariana HS, Hayden MJ, Ahmed NU, Bell JA, Sharp PJ, McIntosh RA (2001) Mapping of durable adult plant and seedling resistances to stripe rust and stem rust diseases in wheat. Aust J Agric Res 52:1247–1255

Boisvert S, Laviolette F, Corbeil J (2010) Ray: simultaneous assembly of reads from a mix of high-throughput sequencing technologies. J Comput Biol 17:1519–1533

Choulet F, Alberti A, Theil S, Glover N, Barbe V, Daron J, Pingault L, Sourdille P, Couloux A, Paux E, Leroy P, Mangenot S, Guilhot N, Le Gouis J, Balfourier F, Alaux M, Jamilloux V, Poulain J, Durand C, Bellec A, Gaspin C, Safar J, Dolezel J, Rogers J, Vandepoele K, Aury JM, Mayer K, Berges H, Quesneville H, Wincker P, Feuillet C (2014) Structural and functional partitioning of bread wheat chromosome 3B. Science 345:1249721

Conner RL, Kuzyk AD, Su H (2003) Impact of powdery mildew on the yield of soft white spring wheat cultivars. Can J Plant Sci 83:725–728

Devos KM, Dubcovsky J, Dvorak J, Chinoy CN, Gale MD (1995) Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. Theor Appl Genet 91:282–288

Dorofeyev VF, Jakubtsiner MM, Rudenko MI, Migushova EF, Udachin RA, Merezhko AF, Semenova LV, Novikova MV, Gradchaninova OD, Shitova IP (1976) The wheats of the world. In: Brezhnev DD, Dorofeyev VF (eds), Kolos Publication, Leningrad, 487 pp. (**in Russian**)

Feuillet C, Langridge P, Waugh R (2008) Cereal breeding takes a walk on the wild side. Trends Genet 24:24–32

Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A (2016) The Pfam protein database: towards a more sustainable future. Nucleic Acids Res 44:D279–D285

Gornicki P, Zhu H, Wang J, Challa GS, Zhang Z, Gill BS, Li W (2014) The chloroplast view of the evolution of polyploid wheat. New Phytol 204:704–714

Hayashi N, Inoue H, Kato T, Funao T, Shirota M, Shimizu T, Kanamori H, Yamane H, Hayano-Saito Y, Matsumoto T, Yano M, Takatsuji H (2010) Durable panicle blast-resistance gene *Pb1* encodes an atypical CC-NBS-LRR protein and was generated by acquiring a promoter through local genome duplication. Plant J 64:498–510

Hernandez P, Martis M, Dorado G, Pfeifer M, Gálvez S, Schaaf S, Jouve N, Šimková H, Valárik M, Doležel J, Mayer KFX (2012) Next-generation sequencing and syntenic integration of flow-sorted arms of wheat chromosome 4A exposes the chromosome structure and gene content. Plant J 69:377–386

Huang L, Brooks SA, Li W, Fellers JP, Trick HN, Gill BS (2003) Map-based cloning of leaf rust resistance gene *Lr21* from the large and polyploid genome of bread wheat. Genetics 164:664–665

Hurni S, Brunner S, Buchmann G, Herren G, Jordan T, Krukowski P, Wicker T, Yahiaoui N, Mago R, Keller B (2013) Rye *Pm8* and wheat *Pm3* are orthologous genes and show evolutionary conservation of resistance function against powdery mildew. Plant J 76:957–969

Islamov B, Peusha H, Jakobson I, Järve K (2015) *Triticum militinae* introgressions into bread wheat affect host responses to powdery mildew challenge. In: Poster session presented at 14th international cereal rusts and powdery mildews conference, Helsingør, Denmark, 5–8 July 2015

Ivaničová Z, Jakobson I, Reis D, Šafář J, Milec Z, Abrouk M, Doležel J, Järve K, Valárik M (2016) Characterization of new allele influencing flowering time in bread wheat introgressed from *Triticum militinae*. New Biotechnol 33:718–727

IWGSC (2014) A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science 345:1251788

IWGSC (2018) Shifting the limits in wheat research and breeding through a fully annotated and anchored reference genome sequence. Science 361:eaar7191

Jakobson I, Peusha H, Timofejeva L, Järve K (2006) Adult plant and seedling resistance to powdery mildew in a *Triticum aestivum* × *Triticum militinae* hybrid line. Theor Appl Genet 112:760–769

Jakobson I, Reis D, Tiidema A, Peusha H, Timofejeva L, Valárik M, Kladivová M, Šimková H, Doležel J, Järve K (2012) Fine mapping, phenotypic characterization and validation of non-race-specific resistance to powdery mildew in a wheat-*Triticum militinae* introgression line. Theor Appl Genet 125:609–623

Järve K, Peusha HO, Tsymbalova J, Tamm S, Devos KM, Enno TM (2000) Chromosome location of a *Triticum timopheevi*-derived powdery mildew resistance gene transferred to common wheat. Genome 43:377–381

Järve K, Jakobson I, Enno T (2002) Tetraploid wheat species *Triticum timopheevi* and *Triticum militinae* in common wheat improvement. Acta Agron Hung 50:463–477

Jia J, Devos KM, Chao S, Miller TE, Reader SM, Gale MD (1996) RFLP-based maps of the homoeologous group-6 chromosomes of wheat and their application in the tagging of *Pm12*, a powdery mildew resistance gene transferred from *Aegilops speltoides* to wheat. Theor Appl Genet 92:559–565

King J, Grewal S, Yang C, Hubbart-Edwards S, Scholefield D, Ashling S, Broomfield D, Howells C, King IP (2017) A step change in the transfer of interspecific variation into wheat from its wild relatives. In: Buerstmayr H, Lang-Mladek C, Steiner B, Michel S, Buerstmayr M, Lemmens M, Vollmann J, Grausgruber H (eds) Proceedings of the 13th international wheat genetics symposium, Tulln, Austria, 23–28 April 2017. BOKU-University of Natural Resources and Life Sciences, Vienna, p. 40

Krattinger SG, Lagudah ES, Spielmeyer W, Singh RP, Huerta-Espino J, McFadden H, Bossolini E, Selter LL, Keller B (2009) A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. Science 323:1360–1363

La Camera S, Balagué C, Geoffroy P, Legrand M, Feussner I, Roby D, Heitz T (2009) The Arabidopsis patatin-like protein 2 (PLP2) plays an essential role in cell death execution and differentially affects biosynthesis of oxylipins and resistance to pathogens. Mol Plant Microbe Interaction 22:469–481

Leath S, Bowen KL (1989) Effects of powdery mildew, triadimenol seed treatment and triadimenol foliar sprays on yield of winter wheat in North Carolina. Phytopathology 79:152–155

Leroy P, Guilhot N, Sakai H, Bernard A, Choulet F, Theil S, Reboux S, Amano N, Flutre T, Pelegrin C, Ohyanagi H, Seidel M, Giacomoni F, Reichstadt M, Alaux M, Gicquello E, Legeai F, Cerutti L, Numa H, Tanaka T, Mayer K, Itoh T, Quesneville H, Feuillet C (2011) TriAnnot: a versatile and high performance pipeline for the automated annotation of plant genomes. Front Plant Sci 3:1–14

Li ZK, Luo LJ, Mei HW, Paterson AH, Zhao XH, Zhong DB, Wang YP, Yu XQ, Zhu L, Tabien R, Stansel JW, Ying CS (1999) A "defeated" rice resistance gene acts as a QTL against a virulent strain of *Xanthomonas oryzae* pv. *oryzae*. Mol Gen Genet 261:58–63

Lillemo M, Lu Q (2015) A meta analysis of partial resistance loci to powdery mildew in wheat. In: Paper presented at 14th international cereal rusts and powdery mildews conference, Helsingør, Denmark, 5–8 July 2015

Lukaszewski AJ (1995) Physical distribution of translocation breakpoints in homoeologous recombinants induced by the absence of the *Ph1* gene in wheat and triticale. Theor Appl Genet 90:714–719

Lukaszewski AJ (2015) Introgressions between wheat and rye. In: Molnár-Láng M, Ceoloni C, Doležel J (eds) Alien introgression in wheat. Springer, New York, pp 163–190

Luo MC, Yang ZL, Kota RS, Dvořák J (2000) Recombination of chromosome $3A^m$ and $5A^m$ of *Triticum monococcum* with homeologous chromosomes *3A* and *5A* of wheat: the distribution of recombination across chromosomes. Genetics 154:1301–1308

Lutz J, Limpert E, Bartos P, Zeller FJ (1992) Identification of powdery mildew resistance genes in common wheat (*Triticum aestivum* L.). Plant Breed 108:33–39

Mago R, Tabe L, Vautrin S, Šimková H, Kubaláková M, Upadhyaya N, Berges H, Kong X, Breen J, Doležel J, Appels R, Ellis JG, Spielmeyer W (2014) Major haplotype divergence including multiple germin-like protein genes, at the wheat *Sr2* adult plant stem rust resistance locus. BMC Plant Biol 14:379

Mago R, Vautrin S, Šimková H, Bansal U, Luo MC, Rouse M, Karaoglu H, Periyannan S, Kolmer J, Jin Y, Ayliffe MA, Bariana H, Park RF, McIntosh R, Doležel J, Bergès H, Spielmeyer W, Lagudah ES, Ellis JG, Dodds PN (2015) The wheat Sr50 gene reveals rich diversity at a cereal disease resistance locus. Nat Plants 1:15186

Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Geer LY, Bryant SH (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. Nucleic Acids Res 45:D200–D203

Michaels SD, Amasino RM (1999) A robust method for detecting single-nucleotide changes as polymorphic markers by PCR. Plant J 14:381–385

Michelmore RW, Meyers BC (1998) Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. Genome Res 8:1113–1130

Mondal S, Rutkoski JE, Velu G, Singh PK, Crespo-Herrera LA, Guzmán C, Bhavani S, Lan C, He X, Singh RP (2016) Harnessing diversity in wheat to enhance grain yield, climate resilience, disease and insect pest resistance and nutrition through conventional and modern breeding approaches. Front Plant Sci 7:991

Periyannan S, Moore J, Ayliffe M, Bansal U, Wang X, Huang L, Deal K, Luo M, Kong X, Bariana H, Mago R, McIntosh R, Dodds P, Dvorak J, Lagudah E (2013) The gene *Sr33*, an ortholog of barley Mla genes, encodes resistance to wheat stem rust race Ug99. Science 341:786–788

Rajaraman J, Douchkov D, Hensel G, Stefanoto FL, Gordon A, Ereful N, Caldararu OF, Petrescu AJ, Kumlehn J, Boyd LA, Schweizer P (2016) An LRR/malectin receptor-like kinase mediates resistance to non-adapted and adapted powdery mildew fungi in barley and wheat. Front Plant Sci 7:1836

Riley R, Chapman V (1958) Genetic control of the cytologically diploid behavior of hexaploid wheat. Nature 182:713–715

Roberts MA, Reader SM, Dalgliesh C, Miller TE, Foote TN, Fish LJ, Snape JW, Moore G (1999) Induction and characterization of *Ph1* wheat mutants. Genetics 153:1909–1918

Röder MS, Korzun V, Wendehake K, Plaschke J, Tixier MH, Leroy P, Ganal MW (1998) A microsatellite map of wheat. Genetics 149:2007–2023

Saintenac C, Zhang W, Salcedo A, Rouse MN, Trick HN, Akhunov E, Dubcovsky J (2013) Identification of wheat gene *Sr35* that confers resistance to Ug99 stem rust race group. Science 341:783–786

Sánchez-Martín J, Steuernagel B, Ghosh S, Herren G, Hurni S, Adamski N, Vrána J, Kubaláková M, Krattinger SG, Wicker T, Doležel J, Keller B, Wulff BBH (2016) Rapid gene isolation in barley and wheat by mutant chromosome sequencing. Genome Biol 17:221

Sears ER (1977) Induced mutant with homoeologous pairing in common wheat. Can J Genet Cytol 19:585–593

Sekhwal MK, Li PC, Lam I, Wang XE, Cloutier S, You FM (2015) Disease resistance gene analogs in plants. Int J Mol Sci 16:19248–19290

Šimková H, Svensson JT, Condamine P, Hřibová E, Suchánková P, Bhat PR, Bartoš J, Šafář J, Close TJ, Doležel J (2008) Coupling amplified DNA from flow-sorted chromosomes to high-density SNP mapping in barley. BMC Genom 9:294

Šimková H, Šafář J, Kubaláková M, Suchánková P, Číhalíková J, Robert-Quatre H, Azhaguvel P, Weng Y, Peng J, Lapitan NLV, Ma Y, You FM, Luo M, Bartoš J, Doležel J (2011) BAC libraries from wheat chromosome 7D: efficient tool for positional cloning of aphid resistance genes. J Biomed Biotechnol 2011:302543

Steuernagel B, Periyannan SK, Hernández-Pinzón I, Witek K, Rouse MN, Yu G, Hatta A, Ayliffe M, Bariana H, Jones JDG, Lagudah ES, Wulff BBH (2016) Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. Nat Biotechnol 34:652–655

Thind AK, Wicker T, Šimková H, Fossati D, Moullet O, Brabant C, Vrána J, Doležel J, Krattinger S (2017) Rapid cloning of genes in hexaploid wheat using cultivar-specific long-range chromosome assembly. Nat Biotechnol 35:793–796

Tiwari VK, Wang S, Sehgal S, Vrána J, Friebe B, Kubaláková M, Chhuneja P, Doležel J, Akhunov E, Kalia B, Sabir J, Gill BS (2014) SNP discovery for mapping alien introgressions in wheat. BMC Genom 15:273

Tsõmbalova J, Karafiátová M, Vrána J, Kubaláková M, Peusha H, Jakobson I, Järve M, Valárik M, Doležel J, Järve K (2016) A haplotype specific to North European wheat (*Triticum aestivum* L.). Genet Resour Crop Evol 64:653–664

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3: new capabilities and interfaces. Nucleic Acids Res 40:e115

Valárik M, Linkiewicz AM, Dubcovsky J (2006) A microcolinearity study at the earliness per se gene Eps-Am 1 region reveals an ancient duplication that preceded the wheat-rice divergence. Theor Appl Genet 112:945–957

Valkoun JJ (2001) Wheat pre-breeding using wild progenitors. Euphytica 119:17–23

van der Biezen EA, Jones JDG (1998) The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. Curr Biol 8:R226–R227

Wicker T, Yahiaoui N, Guyot R, Schlagenhauf E, Liu ZD, Dubcovsky J, Keller B (2003) Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and A^m genomes of wheat. Plant Cell 15:1186–1197

Winfield MO, Allen AM, Burridge AJ, Barker GLA, Benbow HR, Wilkinson PA, Coghill J, Waterfall C, Davassi A, Scopes G, Pirani A, Webster T, Brew F, Bloor C, King J, West C, Griffiths S, King I, Bentley AR, Edwards KJ (2016) High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. Plant Biotechnol J 14:1195–1206

Ye ZM, Ting JPY (2008) NLR, the nucleotide-binding domain leucine-rich repeat containing gene family. Curr Opin Immunol 20:3–9

Zamir D (2001) Improving plant breeding with exotic genetic libraries. Nat Rev Genet 2:983–989

APPENDIX II

Genetic diversity of *Blumeria graminis* f. sp. *hordei* in Central Europe and its comparison with Australian population

Komínková E, Dreiseitl A, Malečková E, Doležel J, Valárik M

Plos One (2016) 11: e016709

DOI: 10.1371/journal.pone.0167099

IF 2.806

PLOS ONE

RESEARCH ARTICLE

Genetic Diversity of *Blumeria graminis* f. sp. *hordei* in Central Europe and Its Comparison with Australian Population

Eva Komínková[1], Antonín Dreiseitl[2], Eva Malečková[1], Jaroslav Doležel[1], Miroslav Valárik[1]*

1 Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Olomouc, Czech Republic, **2** Department of Integrated Plant Protection, Agrotest Fyto Ltd., Kroměříž, Czech Republic

* valarik@ueb.cas.cz

Citation: Komínková E, Dreiseitl A, Malečková E, Doležel J, Valárik M (2016) Genetic Diversity of *Blumeria graminis* f. sp. *hordei* in Central Europe and Its Comparison with Australian Population. PLoS ONE 11(11): e0167099. doi:10.1371/journal.pone.0167099

Editor: Dragan Perovic, GERMANY

Received: July 11, 2016

Accepted: November 7, 2016

Published: November 22, 2016

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work and JD, MV and EK have been supported by the Czech Ministry of Education, Youth and Sports (grant awards, LO1204 and LD14105) and AD by the Ministry of Agriculture of the Czech Republic (project no. RO0211) and Agrotest Fyto Ltd. The funder Agrotest Fyto Ltd provided support in the form of salaries for author AD but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The

Abstract

Population surveys of *Blumeria graminis* f. sp. *hordei* (*Bgh*), a causal agent of more than 50% of barley fungal infections in the Czech Republic, have been traditionally based on virulence tests, at times supplemented with non-specific Restriction fragment length polymorphism or Random amplified polymorphic DNA markers. A genomic sequence of *Bgh*, which has become available recently, enables identification of potential markers suitable for population genetics studies. Two major strategies relying on transposable elements and microsatellites were employed in this work to develop a set of Repeat junction markers, Single sequence repeat and Single nucleotide polymorphism markers. A resolution power of the new panel of markers comprising 33 polymorphisms was demonstrated by a phylogenetic analysis of 158 *Bgh* isolates. A core set of 97 Czech isolates was compared to a set 50 Australian isolates on the background of 11 diverse isolates collected throughout the world. 73.2% of Czech isolates were found to be genetically unique. An extreme diversity of this collection was in strong contrast with the uniformity of the Australian one. This work paves the way for studies of population structure and dynamics based on genetic variability among different *Bgh* isolates originating from geographically limited regions.

Introduction

Since the onset of agriculture, cereals have played a crucial role in human nutrition. However, the needs of the growing human population, which is projected to reach 9.7 billion by 2050 [1], together with the changing climate represent a serious challenge for breeders and researchers to meet the growing demand for food. Despite the fact that barley represents only 5.2% of the total world production of cereals with 144 mil. tons harvested per year [2], this cereal ranks among major crops in many, especially European countries responsible for 59.7% of total world production. The Czech Republic has a long tradition in barley production mainly due to the brewery industry with 14.9% of sowing areas occupied by this crop and its yearly harvest representing as much as 22.4% of total cereal production [3].

specific role of this author is articulated in the 'author contributions' section.

Competing Interests: AD is employed by Agrotest Fyto Ltd. There are no patents, products in development or marketed products to declare. This does not alter our adherence to all the PLOS ONE policies on sharing data and materials, as detailed online in the guide for authors.

The yield is constantly exposed to the risk of adverse effects due to different abiotic and biotic factors. An obligate biotrophic fungus *Blumeria graminis*, a causal agent of powdery mildew disease, is currently ranked among the Top 10 most important fungal plant pathogens [4]. *B. graminis* f. sp. *hordei* (*Bgh*) is responsible for 57% and 51% of leaf disease epidemics occurring on spring and winter barley, respectively, in local conditions [5–6]. To prevent yield losses, breeding resistant cultivars offers an effective crop protection strategy without the need for fungicides. However, interaction between host and pathogen is a highly intricate and dynamic process and our understanding of its complexity is a necessary prerequisite to reach the goal.

To study the pathogen diversity and populations, phenotypic traits including virulence and sensitivity to different fungicides used to be the only approach available before the rise of molecular biology [7–19]. However, in spite of providing valuable data, such an analysis itself can be limiting and its combination with genotype data is therefore highly desirable.

Initially, markers based on isozymes were used in addition to phenotype data. While they proved to be sufficient to characterize genetic variation between different *formae speciales* (ff. spp.) of *B. graminis*, they showed very low level of polymorphism within individual ff. spp. [20,21]. Later on, the first generation of DNA markers provided a widely-used and relatively efficient tool for characterizing *Bgh* populations in addition to virulence gene studies [22–25]. Considerably low genetic variation resulting in majority of isolates sharing the same Restriction Fragment Length Polymorphism (RFLP) pattern in some cases was explained by their clonal origin [22] or geographical isolation together with a lack of selection pressure [25]. In contrast, a number of studies [26–28] reporting application of non-specific Random Amplified Polymorphic DNA (RAPD) markers on *Bgh* isolates collected either across Europe or on single localities indicated a large genetic variation, even within common pathotypes. More recently, Single Nucleotide Polymorphism (SNP) markers were applied to study evolutionary relationships between different ff. spp. of *Bgh*. Wyand and Brown [29] investigated polymorphism of β-tubulin gene and ITS regions both of which did not show any variation within any f. sp., whereas Inuma *et al.* [30] relied additionally on chitin synthase 1 gene and 28S rDNA. Oberhaensli *et al.* [31] were the first to use polymorphism in non-coding intergenic sequence including transposable elements (TEs). However, variation among isolates of single f. sp. has not been reported.

Until recently, Single Sequence Repeat (SSR) markers were rarely used in *B. graminis* due to time-consuming and laborious development. Wang *et al.* [32] described development of SSR markers suitable for analysis of genetic diversity between isolates of *B. graminis* f. sp. *tritici* (*Bgt*) from microsatellite-enriched genomic libraries. A crucial milestone in *Bgh* research have been achieved by whole genome sequencing of strain DH14 performed by Spanu *et al.* [33]. The analysis of genome sequence revealed massive colonization by TEs, predominantly retrotransposons, accounting for 64% of the genome size. Whole genome sequencing of two additional *Bgh* isolates and their comparative analysis with the reference genome DH14 revealed highly polymorphic isolate-specific DNA blocks indicating large genetic variation in the *Bgh* population [34]. Tucker *et al.* [35] were the first to employ the genome sequence data for developing SSR markers to characterize a set of Australian *Bgh* isolates. Together with the possibility of developing markers based on TEs or SNPs, the genomic sequence provides a rich source of polymorphism for diversity and population studies.

In our preliminary experiments, we examined the possibility to use ITS sequences supplemented with markers derived from glyceraldehyde-3-phosphate dehydrogenase gene for assessment of *Bgh* genetic diversity. However, no polymorphism was detected among 14 tested *Bgh* isolates.

A study comparing phenotypes of Central European and Australian isolates reported large difference in virulence complexity between the two populations [36]. This finding raises a

question whether the observed diversity in phenotype corresponds to genetic variability. Considering the geographically limited area of the Czech Republic, there are two possible population structure hypotheses. The first one suggests high variability of *Bgh* isolates with different genotypes, the second one anticipates rather small number of widespread genotypes. Clonal propagation and airborne spread of the pathogen during growing season together with low variability in sequences of "housekeeping" genes and ITS support the hypothesis of low number of genotypes. In contrast, presence of isolates with significantly higher virulence complexity in comparison to Australian *Bgh* population prefers the hypothesis of high genetic variability among *Bgh* isolates facilitated by evolutionary forces favored in Central European conditions (high gene flow, population size and selection pressure exerted by deployment of different R genes in grown barley cultivars). To resolve the question, we mined the DH14 genomic sequence [33] to find microsatellites and retrotransposons suitable for designing new markers. The main objective of this study was i) development of a marker panel with sufficient resolution within the population of Czech *Bgh* isolates and ii) comparison of genetic diversity between Czech and Australian populations.

Results

Marker development

Repeat Junction Markers (RJM) were designed manually from retrotransposons of superfamilies *Copia* and *Gypsy*. 10 contigs containing complete *Copia* or *Gypsy* element including Long Terminal Repeats (LTRs) and Target Site Duplication (TSD) delimiting the retrotransposon insertion site were selected. Out of 20 RJM-derived primer pairs (termed *obm1-obm20*) (S1 Table), five revealed presence/absence variation (PAV) (*obm13*, *obm14*, *obm15*, *obm16*, *obm18*). However, they belonged only to three different insertion sites and as a result, two pairs of markers provided redundant information (*obm13*/*obm14* and *obm15*/*obm16*). Additionally, *obm15*/*obm16* failed to amplify reproducibly and thus, only markers *obm14* and *obm18* were considered.

Eleven primer pairs (*obm2*, *obm3*, *obm4*, *obm6*, *obm7*, *obm8*, *obm9*, *obm10*, *obm17*, *obm19*, *obm20*) provided single fragment of expected size and were selected as candidates for sequencing and SNP discovery. Remaining primer pairs produced multiple PCR products (*obm1*) or no product at all (*obm5*, *obm11*, *obm12*). Six of the sequenced amplicons (*obm2*, *obm3*, *obm4*, *obm7*, *obm8*, *obm19*) showed 100% identity across all testing isolates and were not used in further work. *Obm17* was discarded as well because it yielded a mixture of amplicons of the same length. Primer pairs *obm9* and *obm10* provided SNPs with redundant genotypes and thus *obm9* was considered only.

In summary, this strategy resulted in identification of two reliable PAV markers (*obm14*, *obm18*) and three PCR products (*obm6*, *obm9*, *obm20*) suitable for SNP development (S1 Table). Each identified SNP was scored as individual polymorphism. The *obm6* amplicon yielded four SNPs which were marked as *obm6.1—obm6.4*. The *obm9* yielded two SNP markers designated as *obm9.1* and *obm9.3*. (marker *obm9.2* showed 63 bp indel detected only in isolates Y-069 and H-148 originating from Israel). Finally, *obm20* yielded seven SNP markers, *obm20.1 –obm20.7* (S2 Table). SNP marker designated *obm20.6* was omitted from further analysis since it provided identical genotypes as *obm20.2*.

Out of 10 SSR-based primer pairs (*obm21-obm30*) derived from random DH14 sequence scaffolds, four (*obm24*, *obm27*, *obm28*, *obm29*) produced five, three, three and seven polymorphic bands, respectively. Two primer pairs, *obm22* and *obm26*, produced too complex patterns (*obm22* yielded as much as 48 length variants) for reliable scoring and the markers were not used further. Remaining four primer pairs provided monomorphic amplicons (S3 Table).

Different length variants of *obm24*, *obm27*, *obm28* and *obm29* were scored separately and provided sixteen reliable polymorphisms labeled in the same manner as the SNP markers.

Genotyping of *Blumeria graminis* isolates

50 isolates collected in Australia (AUS), a subset of 11 isolates (W) originating from the genebank maintained at Agrofest Fyto Ltd. and 97 isolates (CZE) collected in the main barley growing regions of the Czech Republic were genotyped using the panel of markers described above. Altogether, 33 polymorphisms were scored and 5,214 data points were obtained. Among all tested isolates, each SNP marker provided two variants. All detected polymorphisms including PAV in RJMs and SSR markers or different variants of nucleotides in SNP markers could be therefore converted to binary data 0/1 or "?" in case of missing or unreliable genotypes. The missing data were observed in 17 cases and represented only 0.33% of the whole dataset. The highest number of polymorphisms (28) was observed among Czech isolates. Only five markers (*obm9.2*, *obm24.1*, *obm28.3*, *obm29.3*, *obm29.5*) were monomorphic. On the other hand, polymorphism was detected only by seven markers for Australian isolates (*obm6.1*, *obm6.2*, *obm9.3*, *obm24.3*, *obm24.4*, *obm29.2*, *obm29.6*, *obm29.7*). Finally, 24 polymorphic markers were observed within the subset of world collection of *Bgh* and nine markers were monomorphic (*obm6.3*, *obm6.4*, *obm20.1*, *obm20.3*, *obm20.5*, *obm27.1*, *obm27.3*, *obm29.1*, *obm29.7)*. The W set comprised five private polymorphisms while the CZE eight and the AUS did not show any private polymorphism.

Pathogen diversity

Molecular variance among and within isolate sets and its polymorphism patterns are summarized in the Table 1 and Fig 1.

The W set has the highest diversity (0.224) comparable with the CZE set (0.220) and the lowest diversity (0,038) was observed in the A set (Fig 1). Results of the AMOVA (Table 1) indicated that most (75%) of the molecular variation is present among individuals within populations whereas only 25% was detected among populations. Permutation tests (based on 999 permutations) suggest that the overall ΦPT was significant (ΦPT = 0.255, P = 0.001; Table 1) which indicates that the differences among the sets are significant. Pairwise population ΦPT analysis identified relatively high ΦPT 0.341 and 0.349 from comparison of CZE and W with AUS set, respectively. On the other hand, the ΦPT from comparison of the CZE and W set was only 0.084.

A phylogenetic analysis of the Czech, Australian and worldwide collection of *Bgh* isolates was used to assess resolution of the marker panel and resulted in a tree (Fig 2) where isolate of *B. graminis* f. sp. *tritici* (*Bgt*258) was used as an outgroup. As expected, Australian isolates showed low level of diversity and majority of them (30 out of 50) shared identical genotype profiles (Fig 2, Section A). This cluster represents one isolate or highly similar family collected from all Australian barley growing regions. Another section (Fig 2, Section B) comprised 14% of isolates highly similar to those in Section A but distributed on seven localities of the southeastern Australian territories including Tasmania. The remaining, genetically more diverse isolates (Fig 2, Section C) are genetically close to the Czech and Uruguayan *Bgh* isolates.

Table 1. Analysis of molecular variance (AMOVA) for 172 individuals from 4 *Bgh* isolate sets based on 33 polymorphisms derived from SSR, RJM, indel and SNP markers.

Source of variation	*Degrees of freedom*	*Sum of squares*	*Means square*	*Est. Var.*	*% variation*	*P value*
Among Pops	3	104.684	34.895	0.954	25%	0.001
Within Pops	168	469.401	2.794	2.794	75%	0.001
Total	171	574.085		3.748	100%	0.001
ΦPT = 0.255						0.001

doi:10.1371/journal.pone.0167099.t001

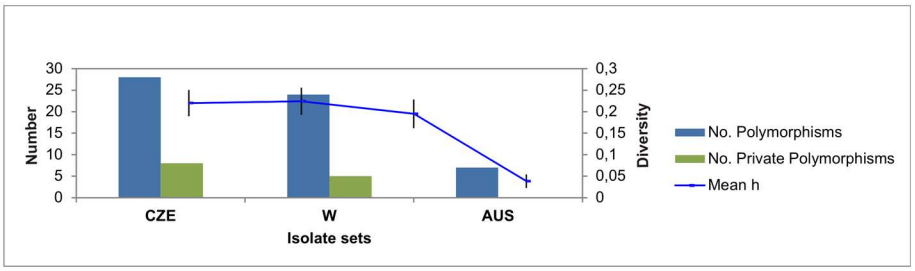


Fig 1. Polymorphism patterns across sets of *Bgh* isolates inferred from genotyping by SSR, RJM, indel and SNP markers. The number of polymorphisms observed in the population of isolates collected in the Czech Republic in season 2012 (CZE), collection of isolates from around the world (W) and population of isolates collected around Australia in season 2011 (AUS) is represented by the blue column. The green column illustrates amount of private polymorphisms. The mean diversity (Mean h) among the sets of isolates is marked by blue line and vertical bars represent standard error.

doi:10.1371/journal.pone.0167099.g001

On the other hand, majority of isolates collected within the Czech Republic (71 out of 97) constituted a separate branch providing an evidence for high resolution of developed marker panel. The exceptions were one group of five and one group of three undistinguished isolates (E2, E3, E37, I32, N3 and A6, I2, N2, respectively) together with nine pairs of isolates with identical genotype profiles. Nevertheless, no correlation between phylogenetic relationships and geographical origin of the isolates was apparent.

Among the *Bgh* isolates of the worldwide collection, the most distinct one was isolate Race I of Japanese origin collected in 1953 together with both Israeli isolates (H-148, Y-069). The remaining isolates are relatively evenly distributed across the tree with no distinct pattern (Fig 2).

Discussion

B. graminis is able to reproduce both sexually and clonally which allows evolution of new pathotypes followed by their fast spread across host. Clarifying the host-pathogen interactions and monitoring of fungicide resistance could significantly enhance the efficiency of disease control. This task requires knowledge of the pathogen population structure and its changes in time and space including its proper molecular characterization. So far, a majority of studies on characterization of *Bgh* populations were based on RFLP and RAPD markers [22–28]. Specific, gene-based SNP markers were only applied to distinguish different *formae speciales* of *B. graminis* [29,30]. However, as mentioned above, our pilot study with this type of markers did not reveal polymorphism among Czech *Bgh* isolates (unpublished data). Since the gene-coding sequences did not yield any polymorphism, we focused on repetitive sequences which accumulate mutations at higher rate. While a successful development of SSR markers for Chinese *Bgt* and Australian *Bgh* isolates has been reported [32,35], markers based on TEs have not been employed so far. Our study is the first to propose them as suitable candidates for population genetic studies focused on this pathogen.

We obtained two unique and reliable PAV markers out of 20 primer pairs designed for 10 different TE insertion sites (S1 Table). Low polymorphism could be explained by recent differentiation of *Bgh* isolates resulting in low frequency of novel retrotransposon insertions. This explanation is supported by finding of Oberhaensli *et al.* [31] who reported large difference in TE content between two different *formae speciales* of *B. graminis* suggesting numerous TE-insertion events which occurred after their divergence dated about 10 million years ago. Nevertheless, high quality reference sequences of both *Bgh* and *Bgt* and higher number of TE-based markers would be required to make a reliable conclusion on this matter. The PAV

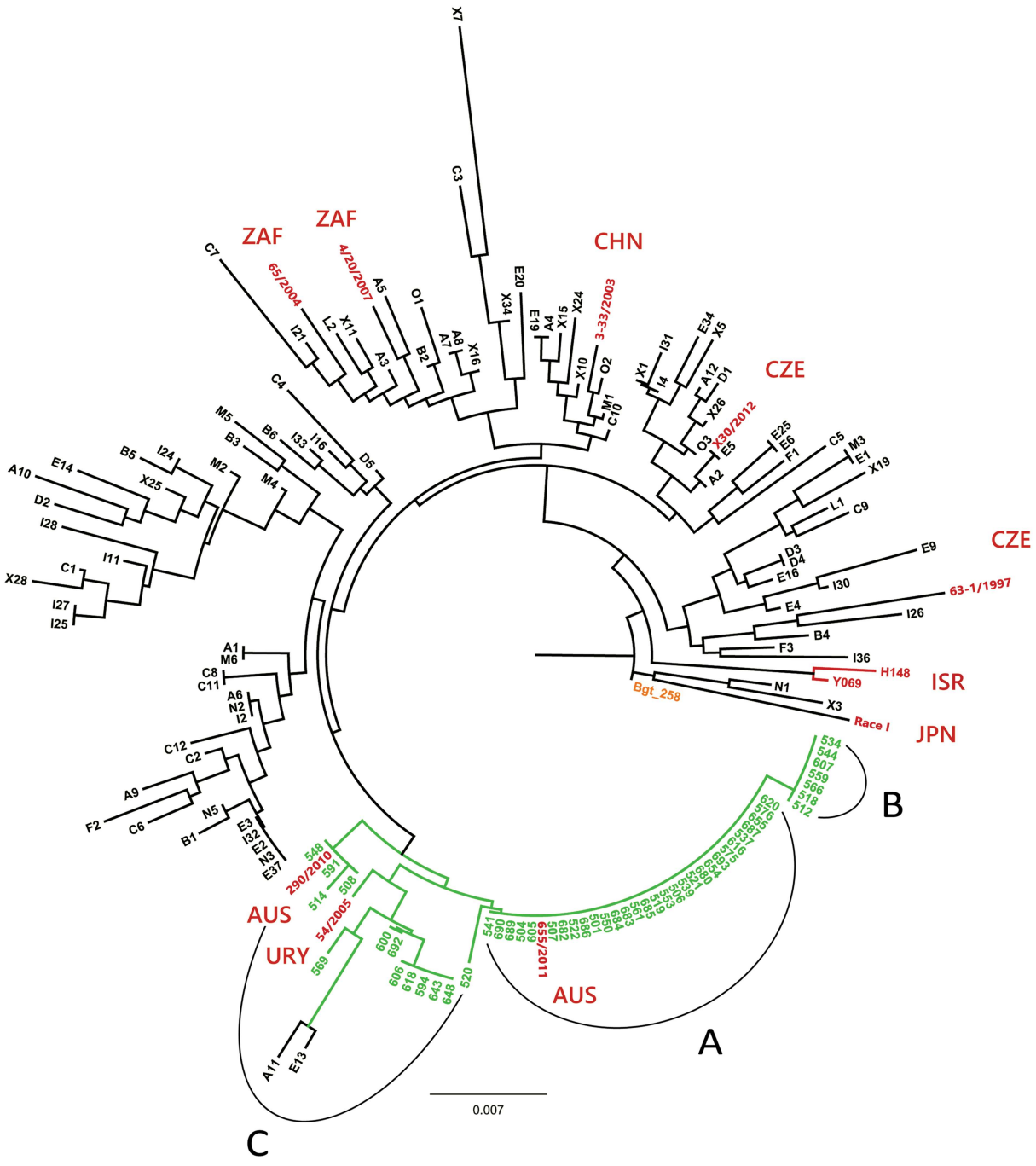


Fig 2. Resolution of the marker panel demonstrated using phylogenetic tree. Phylogenetic tree of *Blumeria graminis* f. sp. *hordei* (*Bgh*) isolates based on 33 polymorphisms (RJMs, SSRs, SNPs). The outgroup (*Bgt*) isolate *Bgt258* is highlighted in orange, Australian isolates collected in 2011 are marked green, isolates of the worldwide collection (CHN = China, CZE = Czech Republic, ISR = Israel, JPN = Japan, URY = Uruguay, ZAF = South African Republic) are highlighted in red and the remaining isolates in black originate from collection made in the Czech Republic in 2012. Letters A, B and C delimit three sections of highly similar Australian isolates. Please note that only few Czech isolated remained undistinguished proving high resolution power of the marker panel. Interestingly, the Czech isolates show no association with area of collection (Fig 3). Additionally, samples representing the worldwide *Bgh* diversity are evenly distributed within the Czech isolates. The exception is isolate Race I originating from Japan and collected more than 60 years ago and two isolates (H148, Y069) collected in Israel. On the other hand, the Australian isolates show low genetic variability and wide spreading of the observed haplotypes. The A) section represents *Bgh* haplotype spread through whole Australian barley growing areas (Fig 4). The B) section shows closely related isolates to A and found only in south east territories. The C) section represents isolates collected mostly around costal territories with high diversity level suggesting recent import from abroad. This hypothesis is supported by haplotype similarity with isolate from Uruguay.

doi:10.1371/journal.pone.0167099.g002

markers were supplemented with TE-based SNPs identified by sequencing monomorphic amplicons. This strategy yielded 13 informative SNPs and one indel (S2 Table) and indicated that described approach provides a rich source of polymorphisms in intergenic regions of *Bgh* genome. Recently, SNP markers has become very popular due to their abundance and availability of high-throughput genotyping platforms [37–39]. However, such platforms require large arrays of SNP markers which are not available for *B. graminis* yet.

The second approach presented here benefits from abundant microsatellite sequences. Out of 10 primer pairs, six resulted in polymorphic bands representing 60% success rate. However, only four of them (*obm24*, *obm27*, *obm28*, *obm29)* could be reliably scored and provided a total of 18 alleles (S3 Table). The remaining two markers produced too many amplicons with unreproducible sizes even using fluorescent labeling and capillary fragment analysis (data not shown). A similar approach of SSR marker design and capillary fragment analysis was employed by Tucker *et al.* [35] who tested eight microsatellite loci out of 30 SSR amplicons. All of them were reported to show polymorphism and an average of seven alleles per locus was detected when genotyping a set of 111 *Bgh* isolates. In contrast, Wang *et al.* [32] did not have genome reference sequence at hand and relied on *de novo* microsatellite isolation by cloning-based approach. Analysis of 90 *Bgt* isolates resulted in five polymorphic microsatellites out of 31 tested with the mean number of observed alleles reaching 5.8. The current work together with studies mentioned above demonstrates that microsatellites offer a rich and not fully explored source of polymorphism for marker development in *B. graminis*.

Genotyping of analyzed samples resulted in detection of 33 polymorphisms. The dataset was subjected to AMOVA analysis which allows for a partitioning of molecular variance within and among populations and tests variance components significance using permutation test [40]. The ΦPT values, which are analogous to the fixation index Fst and are more suitable for binary-haploid dataset used in this study [41], were calculated. The ΦPT = 0.255 (P = 0.0001) for whole dataset indicated that majority (75%) of the molecular variation in the *Bgh* sets occurs among individuals within populations, with only 25% of total variation found among populations. Such high variability within populations compared to variability among populations was observed before in fungi with mixed reproductive system [42] or plants [43]. Assessment of the pairwise population ΦPT values indicates relatively high variability (34.1%) between Australian and Czech populations. Similarly, variability between the AUS population and the set of worldwide isolates was 34.9%. In contrast, comparison between CZE and W indicates only 8.4% variability between the collections. This suggests restriction in gene flow between the Australian *Bgh* population and the rest of the world. On the contrary, gene flow between the Czech *Bgh* pathotypes and pathotypes of surrounding world seems to be unobstructed. However, the W collection do not represents distinct population but rather representation of worldwide diversity among the *Bgh* pathotypes. Moreover, its small size (11 isolates) brings high level of bias and thus the results from comparisons with the W set should be considered only as suggestive. Even though the W collection is small, it provided proportionally high number of private polymorphisms and with increased number of tested pathotypes, higher diversity can be expected even compared to the CZE population.

To assess and graphically visualize the resolution power of the marker panel, the obtained data were used to construct a phylogenetic tree. The resolution power was demonstrated on a set of isolates originating from the Czech Republic among which 73.2% could be unambiguously identified (Fig 2). The level of diversity within the Czech *Bgh* population (84.5% of unique genotypes) is in agreement with phenotype survey [44] which revealed 95% of isolates with distinct virulence spectra out of 521 isolates collected between years 2011–2014. However, nine pairs of isolates from our dataset shared identical genotype profiles (Fig 2). Out of them, five pairs were collected in the same region but their phenotype profiles [44] were different

with the exception of isolates I25 and I27 which might be considered redundant. Similarly, one group of three and one group of five genotypically identical isolates were identified. In the latter, another pair of isolates sharing the locality of origin together with genotype and phenotype profile was detected (E2 and E3). No additional redundancies were observed. The resulting tree topology indicates that the major characteristic of Czech isolates is a lack of correlation between their genotype and geographical origin. This phenomenon can be explained by an ability of long-distance migration of this airborne pathogen [45].

High pathotype diversity of Czech *Bgh* population contrasts with the uniformity of Australian pathotypes [36], which were also characterized by low DNA polymorphism, exhibited low resolution in the phylogenetic tree and formed relatively compact clusters (Fig 2). The low diversity of Australian population is supported by pathogenicity survey of 362 Australian isolates including those analyzed in the present study [46]. Virulence assessment using 32 differential barley varieties resulted in detection of mere 27 different pathotypes with 92% of all isolates belonging to 15 of them. Low level of both genetic and phenotypic variation within Australian *Bgh* population was described earlier. Among 57 isolates collected in different regions of Australia, Whisson [25] obtained identical pathogenicity profiles using 22 differential barley lines and no polymorphism was detected by RFLP analysis using a probe that had previously provided variable fingerprints within British *Bgh* population. Limited virulence and genotype variability of the Australian *Bgh* population is in concordance with its isolation from the rest of the world which restricts gene flow [11,36]. Moreover, the breeding practices in Australia and Europe have been in strong contrast. Europe has long history of barley production with wide range of host varieties (of both spring and winter barley) containing many specific resistance genes. Such long-term directional selection posed upon the pathogen is the major factor increasing virulence complexity of Central European pathotypes. On the other hand, employment of specific resistances has been so far negligible in barley cultivars grown in Australia. This substantial difference makes the two studied populations real extremes.

On the contrary, Tucker *et al.* [35] reported high level of genetic diversity using eight SSR markers within a collection of 111 Australian *Bgh* isolates which constituted 97 unique haplotypes. To cast more light on the discrepancy, we tested all eight SSR markers of Tucker *et al.* [35] on our set of Australian isolates. In one case, no PCR product was obtained. Another two markers were monomorphic after separation by polyacrylamide gel electrophoresis. One marker yielded too complex pattern to be reproducibly scored while the remaining four markers showed polymorphism and provided new data to extend our results.

This work describes an effective strategy for development of markers for population genetics-based studies of cereal powdery mildew pathogen. The markers can be reliably scored using gel electrophoresis and easily converted for use by high-throughput genotyping systems. With only nine primer pairs we identified 33 polymorphisms useful as markers suitable for molecular characterization of *Bgh* isolates from the Czech Republic. 5,214 acquired data points and comparison of the Czech and Australian *Bgh* populations suggests that a higher genetic variability is linked with a higher relative virulence complexity. Additional knowledge on population structure and eventually dynamics might contribute to effective control measures and pathogen-informed strategy for sustainable and broad-spectrum crop resistance.

Materials and Methods

Blumeria graminis isolates

Four sets of *Blumeria graminis* (DC.) Golovin ex Speer f. sp. *hordei* Em. Marchal isolates were used for marker development, validation and diversity assessment. The first one

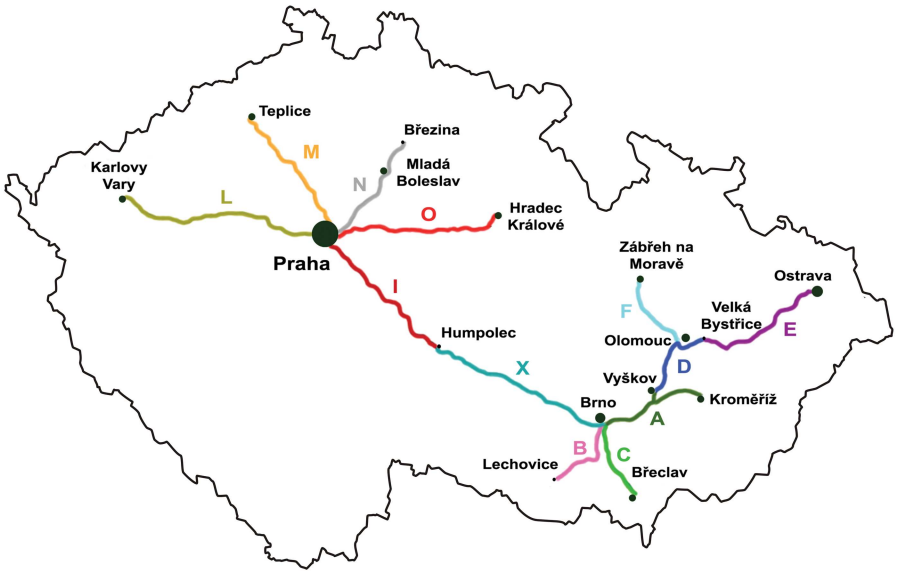


Fig 3. Localities of *Blumeria graminis* f. sp. *hordei* collection in the Czech Republic. The country is divided into sub-regions corresponding to highways passing through main barley growing areas. Each sub-region is labeled as color line and marked with different letter. The letters corresponds to isolate designations in S5 Table. Numbers of isolates collected in individual sub-regions are as follows: A (12), B (6), C (12), D (5), E (15), F (3), I (15), L (2), M (6), N (4), O (3), X (14).

doi:10.1371/journal.pone.0167099.g003

represented fifteen isolates selected for marker testing. This group consisted of twelve isolates from 2009 collection around the Czech Republic [47], two isolates from the South African Republic [48] and one isolate of *Bgt* (*Bgt*258) collected in Olomouc, Czech Republic in 2010 which was used as an outgroup (S7 Table). The second set comprised a collection of 97 isolates originating from the main barley growing areas of the Czech Republic collected in season 2012 [44] (S5 Table, Fig 3). The third set included 50 isolates from barley growing areas of Australia collected in 2011 [46] (S6 Table, Fig 4). The fourth set was a selection of 11 reference isolates (S7 Table) of the pathogen genebank built as a core collection at the Agrotest Fyto Ltd. The genebank comprises isolates collected around the world during the past six decades and includes also Israeli isolates collected on wild barley (*Hordeum vulgare* subsp. *spontaneum*) [49].

DNA extraction

About 10–20 mg of spores were enzymatically treated for 2 hours at 37˚C in 150 μl solution of 10 mg·ml^{-1} Lysing Enzymes from *Trichoderma harzianum* (syn. Glucanex®, Sigma-Aldrich, USA), 1% Triton X-100 (Sigma-Aldrich, USA), 4% 2-mercaptoethanol (Sigma-Aldrich, USA), and 50 mmol·l^{-1} EDTA (Sigma-Aldrich, USA), pH 5.6. Samples were stirred several times during the treatment. Subsequently, 100 μl of lysis solution consisting of 500 mmol·l^{-1} NaCl, 100 mmol l^{-1} Tris-HCl, 50 mmol·l^{-1} EDTA (pH 8.0), 0.02% Sodium Dodecyl Sulphate, (Serva, Germany), 0.5% w/v L-ascorbic acid (Sigma-Aldrich, USA), 0.03% w/v proteinase K (Roche Diagnostics, Switzerland) and 4% 2-mercaptoethanol was added to the suspension. The lysis was carried out by incubation 45 min at 65˚C. Isolated DNA was then purified by phenol-chloroform extraction. Finally, the DNA was precipitated with isopropanol and dissolved in sterile deionized water. RNA was eliminated from the samples by incubation 20 min at 37˚C with 10 mg·l^{-1} ribonuclease A (Sigma-Aldrich, USA).

Fig 4. Origin of Australian *Blumeria graminis* f. sp. *hordei* isolates. Number of isolates collected on individual localities is as follows: Perth (5), Rosedale (6), Horsham (5), Tasmania (7), Wagga Wagga (9), Tamworth (10), Toowoomba (6), Gatton (2). Since the Australian isolates exhibit low level of genetic variability, only three distinct groups were identified (Fig 2). The major haplotype A) is spread through whole Australia and may represent the oldest *Bgh* introduction. The haplotype B) which is highly similar to A) and found only in the south east territories. C) The group with relatively high diversity and similarity to abroad isolate from Uruguay was found only in the costal territories and may represent the latest influx of *Bgh* isolates into Australia.

doi:10.1371/journal.pone.0167099.g004

Marker development and genotyping

The sequence assembly of *Bgh* isolate DH14 (available in the GenBank database under WGS project accession CAUH01000000; [33]) was used as template for marker development. Randomly selected sequence contigs exceeding 10 kb were considered for the analysis. First, the sequence contigs were aligned to a set of repetitive sequences specific for *B. graminis* available at database Repbase (http://www.girinst.org) [33,50] using BLASTN algorithm [51]. For one insertion site of each identified TE, two primer pairs were designed using software Primer3 (http://primer3.ut.ee) [52]. One pair was directed towards the identified LTR and the second was directed out of the LTR. In both cases, one primer of each pair was spanning the insertion site. The amplicon size was expected to range between 450–650 bp (S1 Table). RJMs were genotyped as presence/absence variation. An internal standard was used to verify that observed absence of PCR amplicon is not caused by a problem with PCR amplification. For this purpose, a 100 bp PCR fragment (Forward primer: `ACGCACCCATGTTTGTCAT`, Reverse primer: `CCAATGGGGCAAGACAGTTA`) of glyceraldehyde-3-phosphate dehydrogenase gene (GeneBank: X99732.1) was employed. Subsequently, SNP markers were derived from monomorphic

RJM PCR products. RJMs which gave single PCR fragment for all tested samples were Sanger sequenced from both primers using BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, USA) and DNA Analyzer ABI 3730xl (Applied Biosystems, USA). Finally, SSR motifs were identified using a WebSat software (http://wsmartins.net/websat; [53]). To ensure sufficient resolution in native polyacrylamide gel electrophoresis and to increase the probability of finding polymorphism, only SSRs with unit length of 2–6 bp and comprising at least eight units were considered. Primers were designed in SSR flanking regions using the Primer3 software with expected amplicon size ranging from 70 to 150 bp (S3 Table).

Genotyping of *Blumeria graminis* isolates

PCR amplification was carried out using C1000 Touch™ Thermal Cycler (Bio-Rad, USA) in 15 μl of reaction mix containing 10 $mmol \cdot l^{-1}$ Tris-HCl, 50 $mmol \cdot l^{-1}$ KCl, 1.5 $mmol \cdot l^{-1}$ $MgCl_2$, 0.1% Triton X-100, 0.01% o-cresolsulfonephthalein, 1.5% sacharose, 0.2 $mmol \cdot l^{-1}$ of each dNTP (Fermentas, Lithuania), 0.3 U of Taq Polymerase (Fermentas, Lithuania), 1 $\mu mol \cdot l^{-1}$ of each forward and reverse primer and 500 pg of genomic DNA. The PCR conditions were as follows: 1) Initial denaturation at 95˚C for 5 min; 2) 40 cycles comprising 30 s at 95˚C, 30 s at 50˚-60˚C depending on particular primer pair, 1 min at 72˚C; 3) Final extension at 72˚C for 10 min. PCR products were separated using high-throughput Mega-Gel Vertical Electrophoresis system (C.B.S. Scientific, USA) on 4% (RJMs) or 6% (SSRs) non-denaturing polyacrylamide gels stained with ethidium bromide (Sigma-Aldrich, USA). All designed markers were tested using a selection of fifteen *B. graminis* isolates of the Czech Republic and the South African Republic (S7 Table). Markers polymorphic in this set were used for subsequent genotyping of all studied *Bgh* isolates.

Data analysis

Polymorphic bands of SSR markers were scored separately as individual PAV markers. The same approach was applied for RJMs. MEGA5 software [54] was used for multiple alignment of DNA sequence data and identification of SNPs which were subsequently manually converted to binary data to match the output of PAV markers. Each SNP position was scored as individual marker. For SNPs with redundant genotypes (originating from different primer pairs designed for the same retrotransposon), only one marker was considered for further analysis. To demonstrate the power of the marker panel to discriminate individual *Bgh* isolates, a phylogenetic analysis was performed using software package PHYLIP 3.69 [55]. First, a set of binary data in the PHYLIP format was converted into distance matrix by Restdist tool. In the next step, the matrix was used for construction of an unrooted tree based on the neighbor-joining algorithm [56] by Neighbor tool selected due to its suitability for analyzing large datasets. Results were visualized by software Geneious 9.0.5 (Biomatters Ltd, Auckland, New Zealand). Population analysis of the isolate sets was carried out using software GenAlEx v6.502 [41]. All the analyses were performed with data loaded as binary (haploid). Polymorphism frequencies, diversity and genetic distance estimations were calculated using the tool "Frequency" with option Set-by-step. Binary genetic distance matrix was used for AMOVA analysis and for calculation of ΦPT (the analog of fixation index Fst when data are haploid or binary) with 999 permutations.

Supporting Information

S1 Table. Primer sequences and final markers based on retrotransposon insertion sites.
(DOCX)

S2 Table. Sequence-based markers derived from RJM amplicons.
(DOCX)

S3 Table. Primer sequences and final markers based on microsatellite loci.
(DOCX)

S4 Table. *Blumeria graminis* f. sp. *hordei* isolates used for marker validation.
(DOCX)

S5 Table. *Blumeria graminis* f. sp. *hordei* isolates collected across the Czech Republic in 2012.
(DOCX)

S6 Table. *Blumeria graminis* f. sp. *hordei* isolates collected in Australia in 2011.
(DOCX)

S7 Table. Worldwide collection of *Blumeria graminis* f. sp. *hordei* isolates provided by Agrotest Fyto Ltd.
(DOCX)

Acknowledgments

This study has been performed with support and in the framework of FA COST Action FA1208.

Author Contributions

Conceptualization: MV AD JD.

Formal analysis: MV EK.

Funding acquisition: JD MV AD.

Investigation: EK EM.

Methodology: MV EK.

Project administration: MV EK.

Resources: AD MV JD.

Supervision: MV AD.

Visualization: EK MV.

Writing – original draft: EK MV AD JD EM.

Writing – review & editing: MV EK AD.

References

1. United Nations Department of Economic and Social Affairs/Population Division World Population Prospects: The 2015 Revision, Key Findings and Advance Tables. 2015; Available: https://www.un.org/. Accessed 22 October 2015
2. FAOSTAT—Food and Agriculture Organization of the United Nations, FAO Statistical Databases Agricultural production data. 2015; Available: http://faostat3.fao.org/. Accessed 22 October 2015
3. Czech Statistical Office Agriculture data: Time series 2015. Available: https://www.czso.cz/. Accessed 22 October 2015

4. Dean R, van Kan JAL, Pretorius ZA, Hammond-Kosack KE, di Pietro A, Spanu PD, Rudd JJ, et al. The Top 10 fungal pathogens in molecular plant pathology. Mol Plant Pathol. 2012; 13:414–430. doi: 10.1111/j.1364-3703.2011.00783.x PMID: 22471698
5. Dreiseitl A, Jurečka D. Disease occurrence on spring barley in the Czech Republic in 1989–1995. Ochrana Rostlin. 1996; 32:221–229.
6. Dreiseitl A, Jurečka D. Disease occurrence on winter barley in the Czech Republic in 1989–1996. Ochrana Rostlin. 1997; 33:177–186.
7. Andrivon D, de Vallavieille-Pope C. Racial diversity and complexity in regional populations of *Erysiphe graminis* f. sp. *hordei* in France over a 5-year period. Plant Pathol. 1993; 42:443–464.
8. Brown JKM, Slater SE, See KA. Sensitivity of *Erysiphe graminis* f. sp. *hordei* to morpholine and piperidine fungicides. Crop Prot. 1991; 10:445–454.
9. Bousset L, Hovmøller MS, Caffier V, de Vallavieille-Pope C, Østergård H. Observed and predicted changes over eight years in frequency of barley powdery mildew avirulent to spring barely in France and Denmark. Plant Pathol. 2002; 51:33–44.
10. Caffier V, Hoffstadt T, Leconte M, de Vallavieille-Pope C. Seasonal changes in pathotype complexity in French populations of barley powdery mildew. Plant Pathol. 1996; 45:454–468
11. Dreiseitl A, Pickering RA. Pathogenicity of *Blumeria graminis* f. sp. *hordei* in New Zealand in 1997. New Zeal J Crop Hort. 1999; 27:273–280.
12. Dreiseitl A, Wang J. Virulence and diversity of *Blumeria gramimis* f. sp. *hordei* in East China. Eur J Plant Pathol. 2007; 117:357–368.
13. Hossain MA, Rahman MS. Pathogenic variability of *Erysiphe graminis* f. sp. *hordei* in South Australia, 1981–1985. Aust J Agr Res. 1993; 44:1931–1945.
14. Hovmøller MS, Caffier V, Jalli M, Andersen O, Besenhofer G, Czembor JH, Dreiseitl A, et al. The European barley powdery mildew virulence survey and disease nursery 1993–1999. Agronomie. 2000; 20:729–743.
15. Hovmøller MS, Munk L, Østergård H. Observed and predicted changes in virulence gene frequencies at 11 loci in a local barley powdery mildew population. Phytopathology. 1993; 83:253–260.
16. Limpert E. Frequencies of virulence and fungicide resistance in the European barley mildew population in 1985. J Phytopathol. 1987; 119:298–311.
17. Limpert E, Andrivon D, Fischbeck G. Virulence patterns in population of *Erysiphe graminis* f. sp. *hordei* in Europe in 1986. Plant Pathol. 1990; 39:402–415.
18. Müller K, McDermott JM, Wolfe MS, Limpert E. Analysis of diversity in populations of plant pathogens: The barley powdery mildew pathogen across Europe. Eur J Plant Pathol. 1996; 102:385–395.
19. Yahyaoui AH, Reinhold M, Scharen AL. Virulence spectrum in populations of the barley powdery mildew pathogen, *Erysiphe graminis* f. sp. *hordei* in Tunisia and Morocco in 1992. Plant Pathol. 1997; 46:139–146.
20. Koch G, Köhler W. Isozyme variation and genetic distances of *Erysiphe graminis* DC. *formae speciales*. J Phytopathol. 1990; 129:89–101.
21. Koch G, Köhler W. Isozyme variation and genetic diversity of the European barley powdery mildew population. J Phytopathol. 1991; 131:333–344.
22. Brown JKM, O'Dell M, Simpson CG, Wolfe MS. The use of DNA polymorphisms to test hypotheses about a population of *Erysiphe graminis* f. sp. *hordei*. Plant Pathol. 1990; 39:391–401.
23. Damgaard C, Giese H. Genetic variation in Danish populations of *Erysiphe graminis* f. sp. *hordei*. Plant Pathol. 1996; 45:691–696.
24. O'Dell M, Wolfe MS, Flavell RB, Simpson CG, Summers RW. Molecular variation in populations of *Erysiphe graminis* on barley, oats and rye. Plant Pathol. 1989; 39:340–351.
25. Whisson DL. Molecular characterization and pathogenicity of *Erysiphe graminis* f. sp. *hordei* in Australia. Australas Plant Path. 1996; 25:55–63.
26. Caffier V, Brändle UE, Wolfe MS. Genotypic diversity in barley powdery mildew populations in northern France. Plant Pathol; 1999. 48:582–587.
27. McDermott JM, Brändle U, Dutly F, Haemmerli UA, Keller S, Müller KE, et al. Genetic variation in powdery mildew of barley: Development of RAPD, SCAR and VNTR markers. Phytopathology. 1994; 84:1316–1321.
28. Newton AC, Hackett CA, Guy DC. Diversity and complexity of *Erysiphe graminis* f. sp. *hordei* collected from barley cultivar mixtures or barley plots treated with a resistance elicitor. Eur J Plant Pathol. 1998; 104:925–931.

29. Wyand RA, Brown JKM. Genetic and *forma specialis* diversity in *Blumeria graminis* of cereals and its implications for host-pathogen co-evolution. Mol Plant Pathol. 2003; 4:187–198. doi: 10.1046/j.1364-3703.2003.00167.x PMID: 20569378

30. Inuma T, Khodaparast SA, Takamatsu S. Multilocus phylogenetic analysis within *Blumeria graminis*, a powdery mildew fungi of cereals. Mol Phylogenet Evol. 2007; 44:741–751. doi: 10.1016/j.ympev.2007.01.007 PMID: 17346991

31. Oberhaensli S, Parlange F, Buchmann JP, Jenny FH, Abbott JC, Burgis TA. Comparative sequence analysis of wheat and barley powdery mildew fungi reveals collinearity, dates divergence and indicates host-pathogen co-evolution. Fungal Genet Biol. 2011; 48: 327–334. doi: 10.1016/j.fgb.2010.10.003 PMID: 20955813

32. Wang M, Xue F, Yang P, Duan XY, Zhou YL, Shen CY, et al. Development of SSR markers for a phytopathogenic fungus, *Blumeria graminis* f. sp. *tritici*, using a FIASCO protocol. J Integr Agr. 2014; 13:100–104.

33. Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, Stüber K, et al. Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. Science. 2010; 330:1543–1546. doi: 10.1126/science.1194573 PMID: 21148392

34. Hacquard S, Kracher B, Maekawa T, Vernaldi S, Schulze-Lefert P, van Themaat EVL. Mosaic genome structure of the barley powdery mildew pathogen and conservation of transcriptional programs in divergent hosts. P Natl Acad Sci USA. 2013; 110: E2219–E2228.

35. Tucker MA, Moffat CS, Ellwood SR, Tan KC, Jayasena K, Oliver RP. Development of genetic SSR markers in *Blumeria graminis* f. sp. *hordei* and application to isolates from Australia. Plant Pathol. 2015; 64:337–343.

36. Dreiseitl A. Pathogenic divergence of Central European and Australian populations of *Blumeria graminis* f. sp. *hordei*. Ann Appl Biol. 2014; 165:364–372.

37. Chen H, Xie W, He H, Yu H, Chen W, Li J, et al. A high-density SNP genotyping array for rice biology and molecular breeding. Mol Plant. 2014; 7:541–553. doi: 10.1093/mp/sst135 PMID: 24121292

38. Close T, Bhat PR, Lonardi S, Wu Y, Rostoks N, Ramsay L, Druka A. Development and implementation of high-throughput SNP genotyping in barley. BMC Genomics. 2009; 10:582. doi: 10.1186/1471-2164-10-582 PMID: 19961604

39. Sim S, Durstewitz G, Plieske J, Wieseke R, Gana MW, Van Deynze A. Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. Plos One. 2012; 7:e40563. doi: 10.1371/journal.pone.0040563 PMID: 22802968

40. Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics. 1992; 131:479–491. PMID: 1644282

41. Peakall R, Smouse PE. GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes. 2006; 6:288–295.

42. Narasimhamoorthy B, Saha MC, Swaller T, Bouton JH. Genetic diversity in switchgrass collections assessed by EST-SSR markers. Bioenerg Res. 2008; 1:136–146.

43. Ferrucho RL, Ceresini PC, Ramirez-Escobar UM, McDonald BA, Cubeta MA, García-Domínguez C. The population genetic structure of Rhizoctonia solani AG-3PT from potato in the Colombian Andes. Phytopathology. 2013; 103:862–869. doi: 10.1094/PHYTO-11-12-0278-R PMID: 23464900

44. Dreiseitl A. Changes in virulence frequencies and higher fitness of simple pathotypes in the Czech population of *Blumeria graminis* f. sp. *hordei*. Plant Protect Sci. 2015; 51:67–73.

45. McDermott JM, McDonald BA. Gene flow in plant pathosystems. Annu Rev Phytopathol. 1993; 31:353–373.

46. Dreiseitl A, Fowler RA, Platz GJ. Pathogenicity of *Blumeria graminis* f. sp. *hordei* in Australia in 2010 and 2011. Australas Plant Path. 2013; 42:713–721.

47. Dreiseitl A. Changes in the population of *Blumeria graminis* f. sp. *hordei* in the Czech Republic from 2009 to 2010. Plant Protect Sci. 2011; 47:43–51.

48. Dreiseitl A, Kosman E. Virulence phenotypes of *Blumeria gramimis* f. sp. *hordei* in South Africa. Eur J Plant Pathol. 2013; 136:113–121.

49. Dreiseitl A, Dinoor A, Kosman E. Virulence and diversity of *Blumeria graminis* f. sp. *hordei* in Israel and in the Czech Republic. Plant Dis. 2006; 90:1031–1038.

50. Jurka J, Kohany O. LTR retrotransposons from barley powdery mildew. Repbase Reports. 2011; 11:2333–2333.

51. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990; 215:403–410. doi: 10.1016/S0022-2836(05)80360-2 PMID: 2231712

52. Rozen S, Skaletsky HJ. Primer3 on the www for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, USA, 2000; pp 365–386

53. Martins WS, Lucas DCS, Neves KFS, Bertioli DJ. WebSat—a web software for microsatellite marker development. Bioinformation. 2009; 3:282–283. PMID: 19255650

54. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011; 28:2731–2739. doi: 10.1093/molbev/msr121 PMID: 21546353

55. Felsenstein J. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle. 2005. Available: http://evolution.genetics.washington.edu/phylip.html.

56. Saitou N, Nei M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987; 4:406–425. PMID: 3447015

APPENDIX III

The *in silico* identification and characterization of a bread wheat/*Triticum militinae* introgression line

Abrouk M, Balcárková B, Šimková H, Komínková E, Martis M, Jakobson I, Timofejeva L, Rey E, Vrána J, Kilian A, Järve K, Doležel J, Valárik M

Plant Biotechnology Journal (2017) 15:249-256

DOI: 10.1111/pbi.12610

IF 6.305

Plant Biotechnology Journal

Plant Biotechnology Journal (2017) **15**, pp. 249–256 doi: 10.1111/pbi.12610

The *in silico* identification and characterization of a bread wheat/*Triticum militinae* introgression line

Michael Abrouk[1], Barbora Balcárková[1], Hana Šimková[1], Eva Komínkova[1], Mihaela M. Martis[2,3], Irena Jakobson[4], Ljudmilla Timofejeva[4], Elodie Rey[1], Jan Vrána[1], Andrzej Kilian[5], Kadri Järve[4], Jaroslav Doležel[1] and Miroslav Valárik[1,*]

[1]*Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Olomouc, Czech Republic*

[2]*Munich Information Center for Protein Sequences/Institute of Bioinformatics and Systems Biology, Institute for Bioinformatics and Systems Biology, Helmholtz Center Munich, Neuherberg, Germany*

[3]*Division of Cell Biology, Department of Clinical and Experimental Medicine, Bioinformatics Infrastructure for Life Sciences, Linköping University, Linköping, Sweden*

[4]*Department of Gene Technology, Tallinn University of Technology, Tallinn, Estonia*

[5]*Diversity Arrays Technology Pty Ltd, Canberra, ACT, Australia*

Received 29 November 2015;
revised 21 July 2016;
accepted 8 August 2016.
**Correspondence* (Tel +420 585 238 714; fax +420 585 238 704; email valarik@ueb.cas.cz)

Keywords: GenomeZipper, alien introgression, comparative analysis, chromosome rearrangement, chromosome translocation, linkage drag.

Summary

The capacity of the bread wheat (*Triticum aestivum*) genome to tolerate introgression from related genomes can be exploited for wheat improvement. A resistance to powdery mildew expressed by a derivative of the cross-bread wheat cv. Tähti × *T. militinae* (*Tm*) is known to be due to the incorporation of a *Tm* segment into the long arm of chromosome 4A. Here, a newly developed *in silico* method termed rearrangement identification and characterization (RICh) has been applied to characterize the introgression. A virtual gene order, assembled using the GenomeZipper approach, was obtained for the native copy of chromosome 4A; it incorporated 570 4A DArTseq markers to produce a zipper comprising 2132 loci. A comparison between the native and introgressed forms of the 4AL chromosome arm showed that the introgressed region is located at the distal part of the arm. The *Tm* segment, derived from chromosome 7G, harbours 131 homoeologs of the 357 genes present on the corresponding region of Chinese Spring 4AL. The estimated number of *Tm* genes transferred along with the disease resistance gene was 169. Characterizing the introgression's position, gene content and internal gene order should not only facilitate gene isolation, but may also be informative with respect to chromatin structure and behaviour studies.

Introduction

Using interspecific hybridization to widen a crop's gene pool is an attractive strategy for reversing the genetic bottleneck imposed by domestication and for compensating the genetic erosion, which has resulted from intensive selection (Feuillet *et al.*, 2008). Much of the pioneering research in this area has focused on bread wheat (*Triticum aestivum*), in which over 50 related species have been exploited as donors thanks to the plasticity of the recipient's genome (Jiang *et al.*, 1993; Wulff and Moscou, 2014). Typically, introgression events have involved the transfer of a substantially sized donor chromosome segment, which, along with the target, probably bears gene(s), which impact negatively on the host's fitness (a phenomenon also called 'linkage drag') (Gill *et al.*, 2011; Qi *et al.*, 2007; Zamir, 2001). For this reason, very few introgression lines are represented in commercial cultivars (Rey *et al.*, 2015). The prime means of reducing the length of an introgressed segment is to induce recombination with its homoeologous region (Niu *et al.*, 2011). The success of this strategy is highly dependent on the conservation of gene content and order between the donor segment and its wheat equivalent.

The level of resolution with which introgression segments can be characterized has developed over the years along with advances in DNA technology. Large numbers of genetic markers have been identified in many crop species, including wheat (Bellucci *et al.*, 2015; Chapman *et al.*, 2015; Sorrells *et al.*, 2011; Wang *et al.*, 2014). In a recent example, a wheat mapping population has been genotyped with respect to >100 000 markers, but the mapping resolution achieved has only enabled the definition of around 90 mapping bins per chromosome (Chapman *et al.*, 2015). Given that the genomes of most donor species are poorly characterized, marker data at best allow only the position of an introgressed segment to be defined on the basis of the loss of wheat markers; they cannot determine either the size of the introduced segment or analyse its genetic content. The recently developed 'Introgression Browser' (Aflitos *et al.*, 2015) combines genotypic data with phylogenetic inferences to identify the origin of an introgressed segment, but to do so, a high-quality reference sequence of the host genome is needed, along with a large set of donor sequence data. The first of these requirements is being addressed by a concerted effort to acquire a reference sequence for bread wheat (www.wheatgenome.org). So far, only chromosome (3B) has been fully sequenced, and the gene content of each wheat chromosome has been obtained (Choulet *et al.*, 2014; IWGSC, 2014). The so-called GenomeZipper method (Mayer *et al.*, 2011), based on a variety of resources, has been used to predict gene order along each of the 21 bread wheat chromosomes (IWGSC, 2014).

The improved resistance to powdery mildew of an introgressive line 8.1 derived from the cross of bread wheat cv. Tähti (genome formula ABD) and tetraploid *T. militinae* (*Tm*; genome formula A^tG) is known to be mainly due to the incorporation of a segment of *Tm* chromatin containing the resistance gene *QPm-tut-4A* into the long arm of chromosome 4A (Jakobson *et al.*, 2006, 2012). Here, a

novel *in silico*-based method, termed rearrangement identification and characterization (RICh), has been developed to identify the sequences suitable for generating markers targeting an introgression segment such as the one from *Tm*. The method integrates the GenomeZipper approach with shotgun sequences of chromosome with the introgression. The RICh method was also effective in confirming the identity of the chromosomal rearrangements, which occurred during the evolution of modern wheat.

Results

Chromosome sorting, sequencing and assembly

The flow karyotype derived from the DAPI-stained chromosomes of the DT4AL-TM line included a distinct peak (Figure S1) corresponding to the 4AL telosome (4AL-TM), which enabled it to be sorted to an average purity of 86.2%. The contaminants in the sorted peak comprised a mixture of fragments of various chromosomes and chromatids. DNA of all 45 000 sorted 4AL-TM telosomes was amplified by DNA multiple displacement amplification (MDA). To minimize the risk of representation bias, the products from three independent amplification reactions were pooled. From the resulting 4.5 μg DNA, a total of ~6.2 Gb of sequence was obtained, which was subsequently assembled into 279 077 contigs of individual length >200 bp, with an N50 of 2068 bp (Table 1). When the assembly was aligned with the reference genome sequences of *Brachypodium distachyon* (Vogel *et al.*, 2010), rice (IRGSP, 2005) and sorghum (Paterson *et al.*, 2009), it was apparent that the 4AL-TM telosome shares synteny with segments of *B. distachyon* chromosomes Bd1 and Bd4, rice chromosomes Os3, Os6 and Os11 and sorghum chromosomes Sb1, Sb5 and Sb10 (Figure S2).

Origin of the introgression segment

The chromosomal origin of the *Tm* introgression segment was established by initially flow sorting the *Tm* chromosome complement. This was achieved by pretreating the chromosomes with fluorescence *in situ* hybridization in suspension (FISHIS) (Giorgi *et al.*, 2013) in which GAA microsatellites were fluorescently labelled by FITC. The resulting DAPI *vs* GAA bivariate flow karyotype succeeded in defining 13 distinct clusters (Figure 1). As the haploid chromosome number of *Tm* is 14, one of the clusters was therefore deemed likely to harbour a mixture of two distinct chromosomes. Two of the clusters (#4 and #8) contained sequences that were amplified by the *Xgwm160* (Roder *et al.*, 1998) and *owm82* primers (these two markers are linked to the *QPm-tut-4A* gene from *Tm* introgression). The dispersed profile of cluster #4 (Figure 1) suggested that it was composed of two different A^t genome chromosomes, because all G chromosomes were identified due to a higher GAA content (Badaeva *et al.*, 2010). The *owm72* marker, also linked to the *QPm-tut-4A* gene, amplified two fragments in *Tm*, one of size 205 bp and the other of size 250 bp; only the former was amplified from 4AL-TM telosome or of cluster #8. The fluorescence *in situ* hybridization (FISH) profile of the chromosomes present in cluster #8 unambiguously identified the introgressed segment as deriving from chromosome 7G.

Table 1 Assembly statistics of chromosome arms 4AL-TM, 4AS-CS and 4AL-CS

	4AS-CS	4AL-CS	4AL-TM
Sequencing read depth	241x	116x	23x
Total contigs	301 954	362 851	279 077
Total bases (bp)	282 335 959	361 971 522	266 737 930
Assembly coverage*	0.89x	0.67x	0.49x
Min contig length (bp)	200	200	200
Max contig length (bp)	70 057	129 043	28 604
Average contig length (bp)	935	998	956
N50 length (bp)	2782	3053	2068

The data for 4AS-CS and 4AL-CS arms are taken from IWGSC (2014) and data for 4AL-TM were acquired in this study.

*The size of chromosome arms 4AS-CS (318 Mbp) and 4AL-CS (540 Mbp) were taken from Šafář *et al.* (2010). To estimate the assembly coverage of the 4AL-TM arm, the 4AL-CS size was used.

GenomeZipper improvement

A chromosome 4A zipper was constructed based on Chinese Spring (CS) chromosome specific survey sequences (CSSs) using 1780 specific DArTseq markers ordered in consensus genetic map (Table S1). As DArTseq marker sequences are short (69 nt) and generally nongenic, they were initially anchored to the CSS assembly; this step reduced the number of useful markers to 632 (CSS-DArTseq markers), of which 102 mapped to the short arm and 530 to the long arm. The first version of the zipper comprised a total of 2398 loci. The resulting model for 4AS was collinear with Bd1, Os3 and Sb01, as reported previously (Hernandez *et al.*, 2012). However, the one for 4AL was a mosaic of 15 orthologous blocks (based on the rice genome as the reference), derived from Os11/Bd4/Sb5, Os3/Bd1/Sb1 and Os6/Bd1/Sb10 (Figure S3a). Validation for this complex structure was sought from analysis of the subset of 2638 SNP loci (Wang *et al.*, 2014), which had been assigned a bin locations based on an analysis of a panel of established 4A deletion lines (Endo and Gill, 1996): of these, 750 mapped to five deletion bins on 4AS

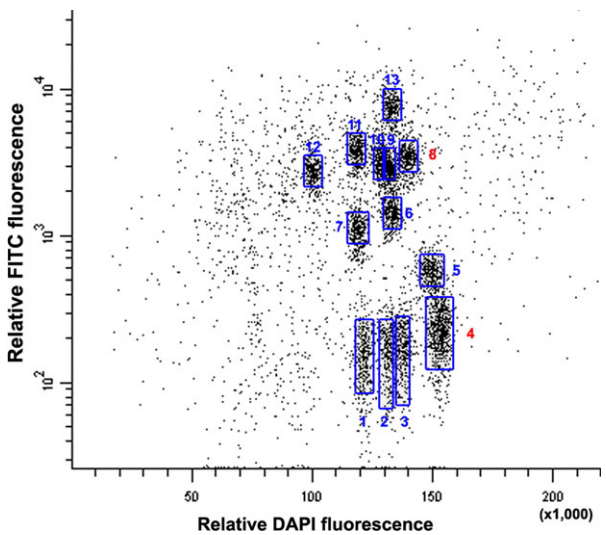


Figure 1 The bivariate flow karyotype of *T. militinae*. Mitotic chromosomes at metaphase were stained with DAPI and GAA microsatellites were labelled with FITC. A set of 13 distinct clusters were obtained (shown boxed). Cluster #8 harbours the *Tm* chromosome (7G) which was the origin of the introgression segment present in line 8.1. Cluster #4 harbours a putative homoeolog of 7G and based on its width and shape most likely comprises a mixture of two distinct chromosomes.

and 1888 to 13 deletion bins on 4AL (Figure S3, Balcárková *et al.*, unpublished). The analysis allowed 329 SNP loci (113 on 4AS, 216 on 4AL) to be integrated into the new 4A zipper. Of the 113 4AS SNP loci, just four mapped to an inconsistent locations, demonstrating the model's accuracy; however, six (#3, #6, #8, #12, #14 and #16) of the 15 4AL blocks were inconsistent with respect to the multiple SNP loci allocations. For example, block #12—positioned in the subtelomeric region according to the zipper—included 18 SNP loci assigned to the pericentromeric region. The GenomeZipper was therefore rerun after first removing the 62 CSS-DArTseq markers associated with the misassignment of the blocks (Table S2); of the 570 CSS-DArTseq markers retained (Table S3), 79 were anchored to at least one of the *B. distachyon*, rice or sorghum scaffolds. The set of 2132 loci (745 on 4AS and 1387 on 4AL) revealed just six (rather than 16) blocks (Figure S3b, Table S2). The final structure resembles that described by Hernandez *et al.* (2012). When the model was retested with SNP markers, no further discrepancies were flagged along distal part of chromosome arm 4AL (Figure S3b).

The *in silico* characterization of the evolutionary chromosome rearrangements on 4AL

The RICh method is based on a stringent identification and density estimation of homoeologs and is validated using a segmentation analysis. To test the approach, the CSS-based scaffolds of chromosome arms 4BS, 4BL, 4DS, 4DL, 5BL, 5DL, 7AS and 7DS (IWGSC, 2014) were compared with that of chromosome 4A, applying as the criteria a 90% level of identity and a minimum alignment length of 100 bp. The numbers of homoeologous loci obtained were, respectively, 719, 762, 636, 877, 850, 673, 602 and 627 (mean 718), but no common distinct blocks allowing for the definition of evolutionary translocations could be identified. A window size of eleven genes was then selected from the 4A zipper for the subsequent segmentation analysis. The ancestral 4AS and 4AL arms had an average density of 0.83, while the remainder of 4AL had a density of only 0.41 (Figure 2a). 4BL and 4DL sequences were homologous to 4AS, and 4BS and 4DS ones to 4AL, confirming the pericentromeric inversion event uncovered before (Devos *et al.*, 1995; Hernandez *et al.*, 2012; Ma *et al.*, 2013; Miftahudin *et al.*, 2004). Immediately following the ancestral 4AL region, the density of homoeologs associated with chromosome group 5 increased (5BL and 5DL: 147 genes, density 0.73), identifying the presence of ancestral 5AL chromatin on this arm (Figure 1b). Finally, the most distal segment of 4AL was associated with an increased density of chromosome group 7 (7AS and 7DS: 557 genes, density 0.45), confirming the ancestral translocation event involving 7BS (Figure 1C).

Characterization of the *Tm* introgression segment

The RICh approach was then used to characterize the 4AL introgressed *Tm* segment. A direct comparison between the 4AL-TM sequence assembly and the 4A-CS zipper (95% identity, 100 bp minimum alignment length) was then made. For the long arm, the segmentation analysis revealed two distinct regions (Figure 3): the more proximal one had a high density of homologous genes (~0.84, 863 loci), so likely corresponds to a region of the 4AL telosome inherited from bread wheat (Figure 3). However, in the distal part of the arm, the homologous gene density fell to ~0.37, suggesting this as the site of the translocation event (Figure 3). Considering the same number of genes in the homologous regions of CS DT4AL chromosome arm (4AL-CS) and 4AL-TM, the comparison between these proximal segments revealed that 16% of homologous genes (167 of 1030) in the 4AL-TM assembly were not identified and may be accounted to the sequencing and assembly imperfection. If this rate of imperfection is applied to the regions including the introgressed segment (357 CS genes *vs* 131 *Tm* homologous genes), the presence of 169 CS nonhomologous genes in the introgression segment could be estimated. The number of such genes represents the size of linkage drag (neglecting allelic variation of the homologous genes).

Discussion

Introgression from related species provides many opportunities to broaden the genetic base of wheat, but its impact on wheat improvement has been limited by a combination of imperfect homology between donor and recipient chromatin, the loss of key recipient genes, the suppression of recombination and linkage drag effect. Thus, obtaining an accurate understanding of the size, homology, orientation and position of an introgressed segment could help to determine which introgression events are more likely to avoid incurring a performance penalty. Such knowledge would also be informative in the context of isolating a valuable gene introduced via an introgression event. Gaining this information requires saturating the target region with molecular markers. In an effort to clone of *QPm-tut-4A* gene introgressed to the wheat 4A chromosome from *T. militinae,* we developed new method for chromosome rearrangements and introgressions identification and characterization.

The presence of ancient intra- and interchromosomal rearrangements is a known complicating issue in the polyploid wheat genome, and the 4AL chromosome arm, which is one of the site of the introgression event selected in line 8.1, has a particularly complex structure. The composition of the proximal segment of the 4AL telosomes carried by DT4AL-TM and the standard CS DT4AL stock was largely identical, as expected. However, distal part of the telosomes differs in presence of *Tm* introgressive segment (Jakobson *et al.*, 2012), but no difference by synteny blocks could be detected. In hybrids between the tetraploid forms *T. turgidum* and *T. timopheevi*, Gill and Chen (1987) noted that while the latter's G genome chromosomes paired most frequently with those from the B genome, chromosome 4A was occasionally involved in pairing with chromosome 7G, presumably as a result of the presence of the 7BS segment on the *T. turgidum* 4AL arm. The likelihood is therefore that the *Tm* chromosome 7G segment, which has contributed the 4A-based powdery mildew resistance of line 8.1, was introduced via homologous recombination with the segment of 4AL carrying 7BS chromatin.

To increase resolution of the analysis, the GenomeZipper method (Mayer *et al.*, 2011), combining genetic maps, data from chromosome shotgun sequencing, and synteny information with sequenced model genomes has been adopted. The method has been useful for developing virtual gene orders in both wheat and barley chromosomes (IWGSC, 2014; Mayer *et al.*, 2011). The most crucial data set is a reliable genetic map, which serves as backbone to integrate and orient the identified syntenic blocks. Two zippers for chromosome 4A have been published to date. The first was based on relatively low coverage sequencing of the chromosome, employing as its backbone a barley linkage map formed from expressed sequence tags distributed over the

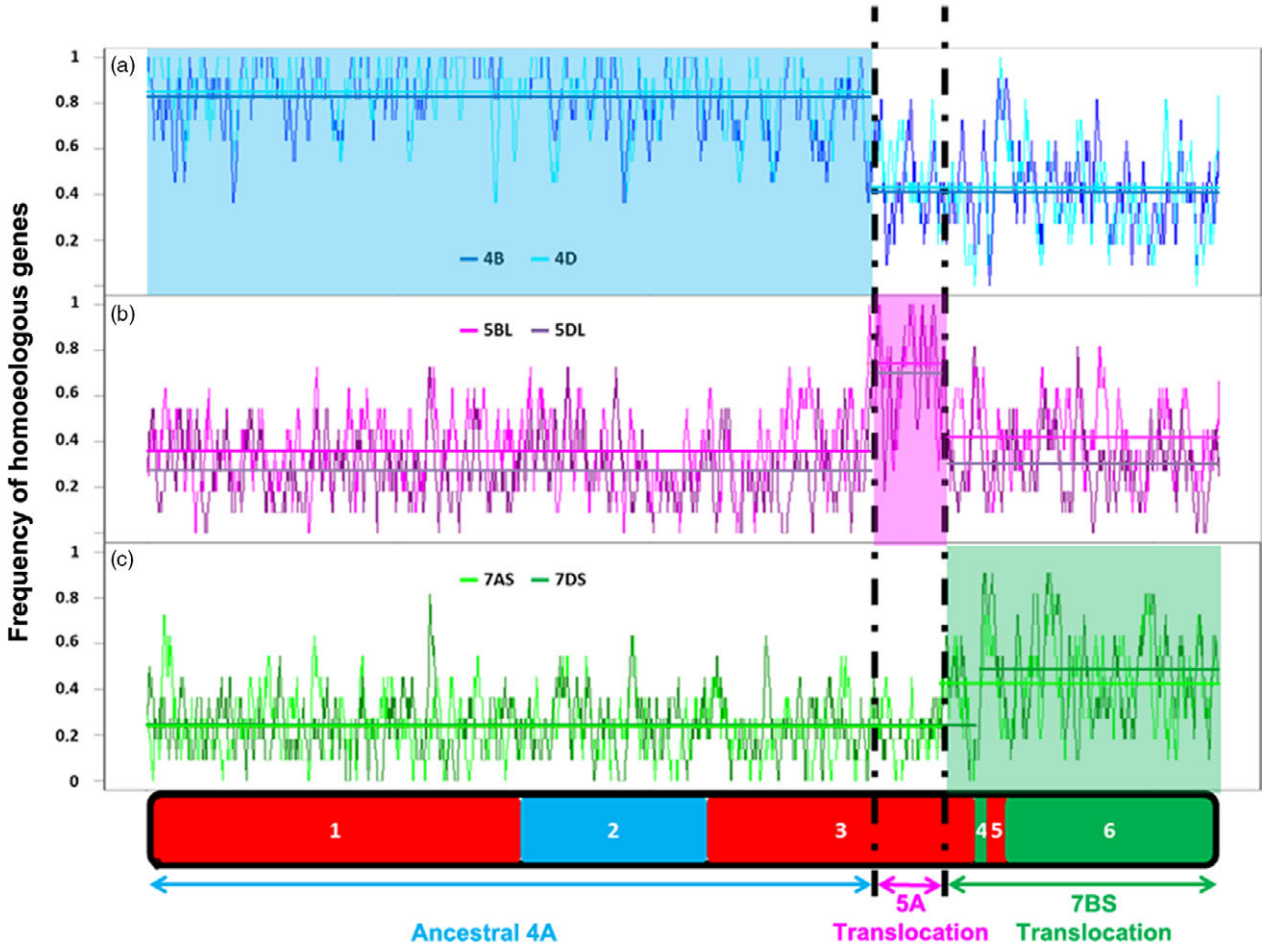


Figure 2 Variation in homoeologous gene density along the various 4A-CS chromosome segments compared to their homoeologous chromosomes. The structure of native 4A-CS chromosome is represented at the bottom with syntenic blocks with rice genome shown in different colours (red = Os3; blue = OS11; green = Os6). (a) The 4A homoeologous gene density compared to 4B and 4D chromosomes, (b) comparison with the 5BL and 5DL chromosome arms and (c) comparison with chromosome arms 7AS and 7DS is shown as homoeologous genes frequency histogram. Homoeologous regions are characterized by a high average frequency (denoted by the horizontal lines). The lower average frequency shown by the group 7 chromosomes reflects a significantly lower sequencing coverage.

chromosome arms 4HS (117 loci), 4HL (16 loci), 5HL (36 loci) and 7HS (36 loci) (Hernandez *et al.*, 2012). The second was based on the 4A CSS and wheat SNP map and consisted of 167 markers on 4AS ordered into 56 mapping bins and 200 (92 mapping bins) markers on 4AL; these were combined with a linkage map developed from a mapping population bred from a cross between bread wheat cv. Opata and a synthetic wheat (Sorrells *et al.*, 2011). Neither of these two zippers was able to provide a sufficient level of resolution to identify the *Tm* introgression into 4AL chromosome arm. The present new zipper was based on consensus DArT map derived from crosses with CS and comprised 55% more markers and 25% more mapping bins than the latter one, which approximately doubled the number of ordered genes/loci (2132 *vs* 1004), and was informative with respect to the *Tm* introgression. When this improved zipper was used in conjunction with the RICh method, it was also possible to recognize the three evolutionary rearrangements, which have long been known to have generated the structure of the modern chromosome arm 4AL (Figure 2) (Devos *et al.*, 1995; Hernandez *et al.*, 2012; Ma *et al.*, 2013; Miftahudin *et al.*, 2004). Similarly, it was able to identify that a lower density of homologous genes obtained at the distal end of the 4AL-TM telosome (Figure 3) is representing the region harbouring the segment introgressed from *Tm*. The *Tm* introgression overlaps with almost the entire chromosome 7BS segment now present on 4AL (Figure 3, Table S2), while the proximal region of the 4AL-CS and 4AL-TM telosomes is essentially of bread wheat origin. The number of wheat loci retained in this latter region did, however, differ by 16% in gene content (4AL-CS—1030 and 4AL-TM—863 genes). This difference may be result of lower sequencing coverage of the 4AL-TM (30x compared to 116x of the 4AL-CS (IWGSC, 2014)) and thus lower representation of the 4AL-TM sequence assembly. If we assume the similar gene density in homologous chromosomes of relative species, as reported before by Tiwari *et al.* (2015), and if the same rate of missing genes as above due to sequencing and assembly imperfections is assumed, estimated 169 CS nonhomologous genes were carried by the introgression in linkage drag. Knowledgeable selection of parental lines that have relatively high frequency of homologous genes in the region of interest (e.g. QTL for resistance in the *Tm* introgression, Figure 3) may

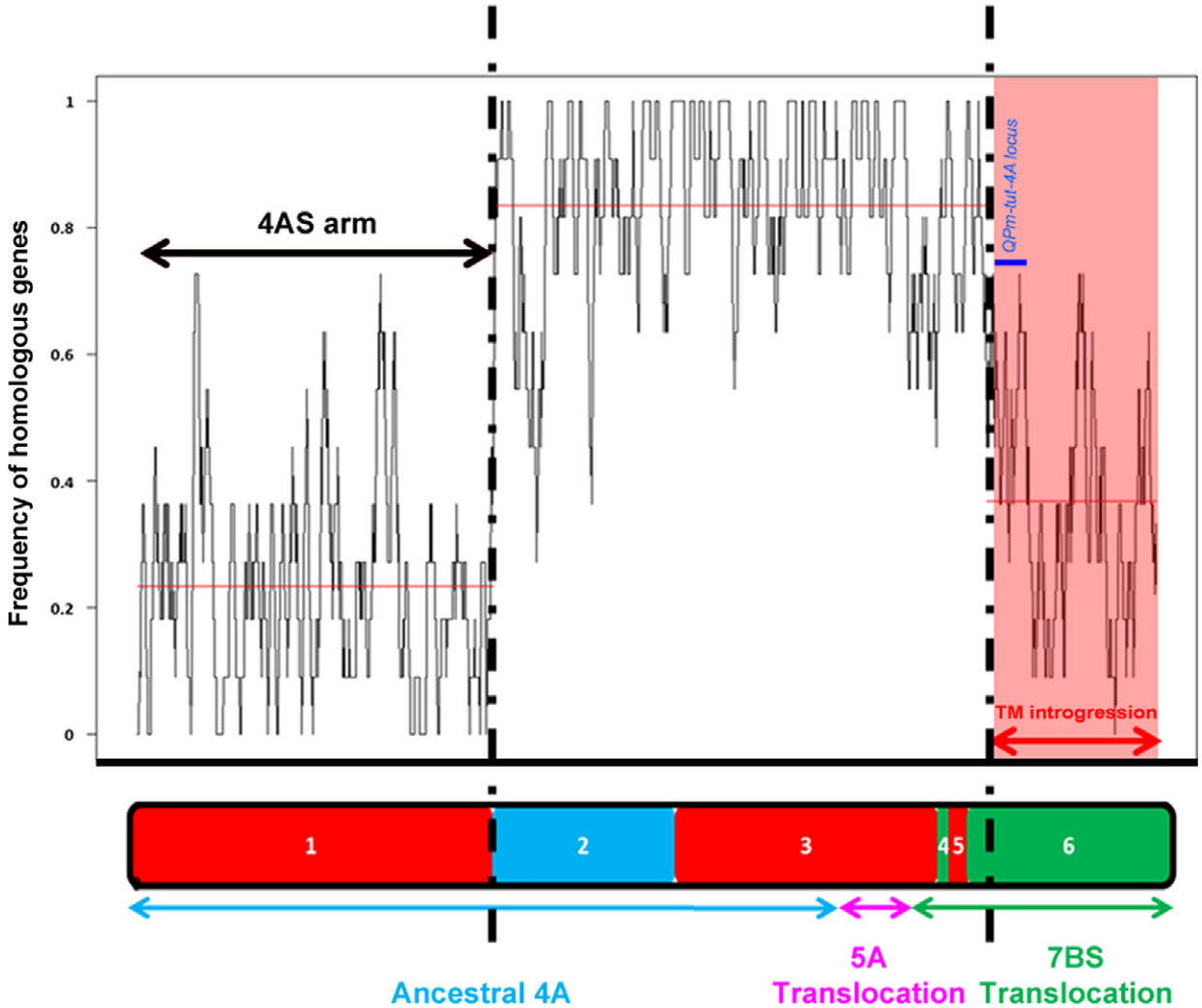


Figure 3 Variation in homologous gene density between 4A-CS chromosome and 4AL-TM telosome. The structure of native 4A-CS chromosome is represented at the bottom with syntenic blocks with rice genome shown in different colours (red = Os3; blue = OS11; green = Os6). The homologous gene density along the 4A-CS zipper compare to 4AL-TM assembly is shown with the black line. The segment of the *Tm* introgression overlaps the 7BS translocation in 4AL (red highlight). The equivalent region on the 4AL-CS telosome harbours 357 genes, only 131 have homologous genes on the *Tm* segment. The dark blue bar represents approximate localization of the *QPm-tut-4A* locus.

increase chances of unobstructed recombination as was observed in the *QPm-tut-4A* locus (Jakobson *et al.*, 2012). So, reducing the length of the introgression segment by inducing further rounds of recombination can lessen (or even eliminate) any negative effects of linkage drag. Application of the RICh approach should prove informative regarding the order or frequency of homologous genes of any such selections. Overall, the RICh method offers a robust means of both characterizing chromosome rearrangements and of predicting the gene content of a specific chromosomal region. Recent advances in high-throughput genotyping permits the elaboration of ever higher density linkage maps (Bellucci *et al.*, 2015; Chapman *et al.*, 2015; Sorrells *et al.*, 2011; Wang *et al.*, 2014). The status of chromosome flow sorting is such that almost any wheat chromosome (Tsõmbalova *et al.*, 2016) and also chromosomes in many crops (Doležel *et al.*, 2014) can now be isolated to a reasonable purity, while the advances in NGS sequencing make RICh widely affordable. These developments should facilitate the preparation of materials needed for applying the RICh approach, thereby offering novel opportunities for a wide range of prebreeding activities, positional cloning, chromatin hybridization and structural studies.

Experimental procedures

Plant materials

Grains of the bread wheat ditelosomic CS DT4AL line were provided by Dr. Bikram Gill (KSU, Manhattan, KS), those of the two nullisomic–tetrasomic lines N4AT4B and N4AT4D (Sears and Sears, 1978) by the National BioResource Centre (Kyoto, Japan), those of *Tm* ($2n = 4x = 28$, genome formula A^tA^tGG) accession K-46007 by the N.I. Vavilov Institute of Plant Industry (St. Petersburg, Russia). The line denoted DT4AL-TM was generated from the cross CS DT4AL × 8.1: the line carries 40 bread wheat chromosomes and a pair of 4AL telosomes with the *Tm* introgression (4AL-TM) and is resistant to powdery mildew (Jakobson *et al.*, 2012).

Flow sorting and amplification of the 4AL telosome carried by 4AL-TM

Liquid suspensions of mitotic chromosomes were prepared from root tips of 4AL-TM seedlings as described by (Vrána *et al.*, 2000). The telosomes were separated from the rest of the genome by flow sorting, using a FACSAria II SORP flow cytometer and sorter (BD Biosciences, San Jose, CA). The level of contamination within a sorted peak was determined using FISH, based on probes detecting telomeric repeats, the Afa repeat and $(GAA)_n$, following the methods described by Kubaláková *et al.* (2003). The flow-sorted 4AL-TM telosomes were treated with proteinase, after which DNA was extracted using a Millipore Microcon YM-100 column (www.millipore.com). Chromosomal DNA was MDA amplified using the Illustra GenomiPhi V2 DNA amplification kit (GE Healthcare) as described by Šimková *et al.* (2008).

Identifying the origin of the introgression segment on the 4AL-TM telosome

Chromosomes of *T. militinae* were flow sorted as described above. However, prior to flow cytometry, GAA microsatellites on chromosomes were labelled by FITC using FISHIS protocol (Giorgi *et al.*, 2013). Bivariate analysis (DNA content/DAPI *vs* GAA/FITC) enabled discrimination of 13 of 14 chromosomes of *T. militinae*. Individual chromosome fractions were sorted into tubes for PCR amplification and onto microscopic slides for identification of sorted chromosomes by FISH. Three markers linked to the *Tm* powdery mildew resistance gene *QPm-tut-4A* were used for the selection of the critical cluster: these were the microsatellite *Xgwm160* (Roder *et al.*, 1998) and two unpublished, one (*owm72*) amplified by the primer pair 5′-TGCTTGCTTGTA GATTGTGCA/5′-CCAGTAAGCTTTGCCGTGTG) and the other (*owm82*) by 5′-GGGAGAGACGAAAGCAGGTA/5′-CTTGCATG CACGCCAGAATA. Each 20 μL PCR contained 0.01% (w/v) o-cresol sulphonephtalein, 1.5% (w/v) sucrose, 0.2 mM dNTP, 0.6 U Taq DNA polymerase and 1 μM of each primer in 10 mM Tris-HCl/50 mM KCl/1.5 mM $MgCl_2$/0.1% (v/v) Triton X-100. The template comprised about 500 sorted chromosomes. Test reactions were seeded with either 20 ng genomic DNA extracted from CS, *Tm*, N4AT4B or N4AT4D, or with 50 pg of MDA amplified DNA from 4AL-CS and 4AL-TM telosomes. The reactions were subjected to an initial denaturation (95 °C/5 min), followed by 40 cycles of 95 °C/30 s, 55 °C/30 s and 72 °C/30 s, and completed with an elongation of 72 °C/10 min. The products were electrophoretically separated through 4% nondenaturing polyacrylamide gels and visualized by EtBr staining. The markers were mapped using a F_2 population bred from the cross CS × 8.1 (Jakobson *et al.*, 2012).

Sequencing of the 4AL telosome

A CSS assembly of CS chromosome arms 4AS (4AS-CS) and 4AL (4AL-CS) were acquired from Internation Wheat Genome Sequencing Consortium (IWGSC, 2014). Two sequencing libraries of DNA amplified from the 4AL-TM telosome were constructed using a Nextera kit (Illumina, San Diego, CA) with the insert size adjusted to 500 and 1000 bp. The resulting clones were sequenced as paired-end reads by IGA (Udine, Italy) using a HiSeq 2000 device (Illumina). The 4AL-TM reads were assembled with SOAPdenovo2 software, applying a range of k-mers (54–99, with a step size of 3) to select the assembly with the highest coverage and the largest N50. Assembled scaffolds (k-mer of 69, minimum length 200 bp) were chosen for further analysis (Table 1).

DArTseq and SNP maps for GenomeZipper construction and validation

A DArTseq consensus map, based on four crosses involving cv. Chinese Spring as a parent has been provided by DArT PL (www.diversityarrays.com). Individual maps were created using DArT PL's OCD MAPPING program (Petroli *et al.*, 2012) to order DArTseq and array-based DArTs. DArT PL's consensus mapping software (Raman *et al.*, 2014) was applied to create a consensus map using similar strategy as described in Li *et al.* (2015). Version 3.0 of consensus map with approximately 70 000 markers was used in this study.

A SNP deletion map (Balcárková *et al.*, unpublished) was used for validation. Genomic DNAs of a set of 15 chromosome 4A deletion lines (Endo and Gill, 1996) and DNAs amplified from 4AL-CS and 4AS-CS chromosome arms as controls were genotyped at USDA-ARS (Fargo, ND) using a iSelect 90k SNP array (Wang *et al.*, 2014) on Infinium platform (Illumina). The raw genotypic data were manually analysed using GenomeStudio V2011.1 software (Illumina).

Comparative analysis and GenomeZipper analysis

Synteny between related genomic segments was assessed using ChromoWIZ software (Nussbaumer *et al.*, 2014). The number of conserved genes present within a series of 0.5-Mbp genomic windows (window shift 0.1 Mbp) was determined. The consensus chromosome 4A linkage map used as the backbone for the GenomeZipper analysis comprised 1780 DArTseq markers (Table S1). As these sequences are mostly short (69 nt) and few identify coding sequence, they were first aligned to the set of 4A CSS contigs, preserving only those contigs that matched the entire DArTseq marker sequence at a level of at least 98% identity. The retained CSS contigs ('CSS-DArTseq markers') were used for the construction of the zipper, which was subsequently validated against the SNP deletion map (2706 SNPs). Similarly as above, only those 4A CSS contigs that aligned with SNP loci along their entire length (98% identity threshold) were retained. Ordering of the CSS-DArTseq markers was compared with that ordered by SNPs from the deletion bin map and CSS-DArTseq markers which do not follow the SNP order were eliminated, and a second version of the zipper was generated using the remaining markers (Table S3). This version was revalidated against the SNP deletion map.

The RICh approach

To identify introgressed/translocated regions, the final 4A zipper was compared to the complete set of CSS sequences obtained from chromosome arms 4BS, 4BL, 4DS, 4DL, 5BL, 5DL, 7AS and 7DS (IWGSC, 2014). Alignments were performed using the BLAST algorithm (Altschul *et al.*, 1990). The BLAST outputs were filtered by applying the following criteria: a minimum identity of either 90% (translocation analysis) or 95% (introgression analysis) and a minimum alignment length of 100 bp. For each comparison, the density of homologous genes was evaluated using a sliding window of eleven genes (five upstream and five downstream), and a segmentation analysis was performed using the R package changepoint v1.1 (Killick and Eckley, 2014), applying the parameter segment neighbourhoods method with a BIC penalty on the mean change. The method allows a statistical detection of gene density changes along the chromosome, corresponding to an

increase or decrease in the level of synteny. For translocation events, an increase in synteny level with one group of homoeologs is required, while for an introgression, a loss of orthology is anticipated.

Acknowledgements

We thank Z. Dubská and R. Šperková for their technical assistance with flow sorting and M. Kubaláková for her FISH-based identification of sorted chromosomes. L. Pingault and E. Paux are acknowledged for their contribution to the statistical analyses, S. Chao for the deletion lines genotyping and J. Bartoš for his critical reading of the manuscript and for his helpful suggestions. This research was supported by grants awarded by the National Program of Sustainability I (No. LO1204), by the Czech Science Foundation (No. 14-07164S) and by the Estonian Ministry of Agriculture.

Conflict of interests

Dr. A Kilian is head of Diversity Arrays Technology Pty Ltd.

References

Aflitos, S.A., Sanchez-Perez, G., de Ridder, D., Fransz, P., Schranz, M.E., de Jong, H. and Peters, S.A. (2015) Introgression browser: high-throughput whole-genome SNP visualization. *Plant J.* **82**, 174–182.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.

Badaeva, E.D., Budashkina, E.B., Bilinskaya, E.N. and Pukhalskiy, V.A. (2010) Intergenomic chromosome substitutions in wheat interspecific hybrids and their use in the development of a genetic nomenclature of Triticum timopheevii chromosomes. *Russian Journal of Genetics*, **46**, 769–785.

Bellucci, A., Torp, A.M., Bruun, S., Magid, J., Andersen, S.B. and Rasmussen, S.K. (2015) Association mapping in Scandinavian winter wheat for yield, plant height, and traits important for second-generation bioethanol production. *Front. Plant Sci.* **6**, 1046.

Chapman, J.A., Mascher, M., Buluc, A., Barry, K., Georganas, E., Session, A., Strnadova, V. *et al.* (2015) A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome. *Genome Biol.* **16**, 26.

Choulet, F., Alberti, A., Theil, S., Glover, N., Barbe, V., Daron, J., Pingault, L. *et al.* (2014) Structural and functional partitioning of bread wheat chromosome 3B. *Science*, **345**, 1249721.

Devos, K.M., Dubcovsky, J., Dvořák, J., Chinoy, C.N. and Gale, M.D. (1995) Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. *Theor. Appl. Genet.* **91**, 282–288.

Doležel, J., Vrána, J., Cápal, P., Kubaláková, M., Burešová, V. and Simková, H. (2014) Advances in plant chromosome genomics. *Biotechnol. Adv.* **32**, 122–136.

Endo, T. and Gill, B. (1996) The deletion stocks of common wheat. *J. Hered.* **87**, 295–307.

Feuillet, C., Langridge, P. and Waugh, R. (2008) Cereal breeding takes a walk on the wild side. *Trends Genet.* **24**, 24–32.

Gill, B.S. and Chen, P.D. (1987) Role of cytoplasm-specific introgression in the evolution of the polyploid wheats. *Proc. Natl Acad. Sci. USA*, **84**, 6800–6804.

Gill, B.S., Friebe, B.R. and White, F.F. (2011) Alien introgressions represent a rich source of genes for crop improvement. *Proc. Natl Acad. Sci. USA*, **108**, 7657–7658.

Giorgi, D., Farina, A., Grosso, V., Gennaro, A., Ceoloni, C. and Lucretti, S. (2013) FISHIS: fluorescence in situ hybridization in suspension and chromosome flow sorting made easy. *PLoS ONE*, **8**, e57994.

Hernandez, P., Martis, M., Dorado, G., Pfeifer, M., Gálvez, S., Schaaf, S., Jouve, N. *et al.* (2012) Next-generation sequencing and syntenic integration of flow-sorted arms of wheat chromosome 4A exposes the chromosome structure and gene content. *Plant J.* **69**, 377–386.

IRGSP. (2005) The map-based sequence of the rice genome. *Nature*, **436**, 793–800.

IWGSC. (2014) A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science*, **345**, 1251788.

Jakobson, I., Peusha, H., Timofejeva, L. and Järve, K. (2006) Adult plant and seedling resistance to powdery mildew in a *Triticum aestivum* x *Triticum militinae* hybrid line. *Theor. Appl. Genet.* **112**, 760–769.

Jakobson, I., Reis, D., Tiidema, A., Peusha, H., Timofejeva, L., Valárik, M., Kladivová, M. *et al.* (2012) Fine mapping, phenotypic characterization and validation of non-race-specific resistance to powdery mildew in a wheat-*Triticum militinae* introgression line. *Theor. Appl. Genet.* **125**, 609–623.

Jiang, J., Friebe, B. and Gill, B. (1993) Recent advances in alien gene transfer in wheat. *Euphytica*, **73**, 199–212.

Killick, R. and Eckley, I.A. (2014) changepoint: an R package for changepoint analysis. *J. Stat. Softw.* **58**, 1–19.

Kubaláková, M., Valárik, M., Bartoš, J., Vrána, J., Číhalíková, J., Molnár-Láng, M. and Doležel, J. (2003) Analysis and sorting of rye (*Secale cereale* L.) chromosomes using flow cytometry. *Genome*, **46**, 893–905.

Li, H., Vikram, P., Singh, R., Kilian, A., Carling, J., Song, J., Burgueno-Ferreira, J. *et al.* (2015) A high density GBS map of bread wheat and its application for dissecting complex disease resistance traits. *BMC Genom.* **16**, 216.

Ma, J., Stiller, J., Berkman, P.J., Wei, Y., Rogers, J., Feuillet, C., Doležel, J. *et al.* (2013) Sequence-based analysis of translocations and inversions in bread wheat (*Triticum aestivum* L.). *PLoS ONE*, **8**, e79329.

Mayer, K., Martis, M., Hedley, P., Simkova, H., Liu, H. and Morris, J. (2011) Unlocking the barley genome by chromosomal and comparative genomics. *Plant Cell*, **23**, 1249–1263.

Miftahudin, Ross, K., Ma, X.-F., Mahmoud, A.A., Layton, J., Milla, M.A.R., Chikmawati, T. *et al.* (2004) Analysis of expressed sequence tag loci on wheat chromosome group 4. *Genetics*, **168**, 651–663.

Niu, Z., Klindworth, D.L., Friesen, T.L., Chao, S., Jin, Y., Cai, X. and Xu, S.S. (2011) Targeted introgression of a wheat stem rust resistance gene by DNA marker-assisted chromosome engineering. *Genetics*, **187**, 1011–1021.

Nussbaumer, T., Kugler, K.G., Schweiger, W., Bader, K.C., Gundlach, H., Spannagl, M., Poursarebani, N. *et al.* (2014) chromoWIZ: a web tool to query and visualize chromosome-anchored genes from cereal and model genomes. *BMC Plant Biol.* **14**, 348.

Paterson, A.H., Bowers, J.E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., Haberer, G. *et al.* (2009) The Sorghum bicolor genome and the diversification of grasses. *Nature*, **457**, 551–556.

Petroli, C., Sansaloni, C., Carling, J., Steane, D., Vaillancourt, R. and Myburg, A. (2012) Genomic characterization of DArT markers based on high-density linkage analysis and physical mapping to the eucalyptus genome. *PLoS ONE*, **7**, e44684.

Qi, L., Friebe, B., Zhang, P. and Gill, B.S. (2007) Homoeologous recombination, chromosome engineering and crop improvement. *Chromosome Res.* **15**, 3–19.

Raman, H., Raman, R., Kilian, A., Detering, F., Carling, J. and Coombes, N. (2014) Genome-wide delineation of natural variation for pod shatter resistance in *Brassica napus*. *PLoS ONE*, **9**, e101673.

Rey, E., Molnár, I. and Doležel, J. (2015) Genomics of wild relatives and alien introgressions. In *Alien Introgression in Wheat: Cytogenetics, Molecular Biology, and Genomics* (Molnár-Láng, M., Ceoloni, C. and Doležel, J., eds), pp. 347–381. Cham: Springer International Publishing.

Roder, M.S., Korzun, V., Wendehake, K., Plaschke, J., Tixier, M.H., Leroy, P. and Ganal, M.W. (1998) A microsatellite map of wheat. *Genetics*, **149**, 2007–2023.

Šafář, J., Šimková, H., Kubaláková, M., Číhalíková, J., Suchánková, P., Bartoš, J. and Doležel, J. (2010) Development of chromosome-specific BAC resources for genomics of bread wheat. *Cytogenet Genome Res.* **129**, 211–223.

Sears, E.R. and Sears, L.M.S. (1978) The telocentric chromosomes of common wheat. In *Proc 5th Int Wheat Genet Symp* (Ramanujam, S., ed), pp. 389–407. New Delhi: Indian Soc of Genet Plant Breed.

Šimková, H., Svensson, J.T., Condamine, P., Hřibová, E., Suchánková, P., Bhat, P.R., Bartoš, J. *et al.* (2008) Coupling amplified DNA from flow-sorted chromosomes to high-density SNP mapping in barley. *BMC Genom.* **9**, 294.

Sorrells, M.E., Gustafson, J.P., Somers, D., Chao, S.M., Benscher, D., Guedira-Brown, G., Huttner, E. *et al.* (2011) Reconstruction of the synthetic W7984 x Opata M85 wheat reference population. *Genome*, **54**, 875–882.

Tiwari, V.K., Wang, S., Danilova, T., Koo, D.H., Vrana, J., Kubalakova, M., Hribova, E. *et al.* (2015) Exploring the tertiary gene pool of bread wheat: sequence assembly and analysis of chromosome 5M(g) of *Aegilops geniculata*. *Plant J.* **84**, 733–746.

Tsõmbalova, J., Karafiátová, M., Vrána, J., Kubaláková, M., Peuša, H., Jakobson, I., Järve, M. *et al.* (2016) A haplotype specific to North European wheat (*Triticum aestivum* L.). *Genet. Resour. Crop Evol.* DOI 10.1007/s10722-016-0389-9

Vogel, J.P., Garvin, D.F., Mockler, T.C., Schmutz, J., Rokhsar, D., Bevan, M.W., Barry, K. *et al.* (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature*, **463**, 763–768.

Vrána, J., Kubaláková, M., Simková, H., Číhalíková, J., Lysák, M.A. and Doležel, J. (2000) Flow sorting of mitotic chromosomes in common wheat (*Triticum aestivum* L.). *Genetics*, **156**, 2033–2041.

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B.E., Maccaferri, M. *et al.* (2014) Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnol. J.* **12**, 787–796.

Wulff, B.B.H. and Moscou, M.J. (2014) Strategies for transferring resistance into wheat: from wide crosses to GM cassettes. *Front. Plant Sci.* **5**, 692. doi: 10.3389/fpls.2014.00692

Zamir, D. (2001) Improving plant breeding with exotic genetic libraries. *Nat. Rev. Genet.* **2**, 983–989.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Figure S1 The flow karyotype of DT4AL-TM, a bread wheat line ditelosomic for 4AL, the distal portion of which includes a segment translocated from *T. militinae*.
Figure S2 A comparative analysis of the telosomes 4AL-CS and 4AL-TM with the *B. distachyon*, rice and sorghum genomes.
Figure S3 Refining the robustness of the 4A zipper.
Table S1 Consensus 4A DArTseq map, based on four independent populations, each involving CS as one parent.
Table S2 The new 4A zipper, composed of 2132 loci, constructed using the CS-based 4A specific DArTseq map and validated by reference to SNPs mapped using a panel of deletion lines.
Table S3 The set of CSS-DArTseq markers used to construct the new zipper.

APPENDIX IV

The utility of flow sorting to identify chromosomes carrying a single copy transgene in wheat

Cápal P, Endo TR, Vrána J, Kubaláková M, Karafiátová M, Komínková E, Mora-Ramírez I, Weschke W, Doležel J

Plant Methods (2016) 12: 24

DOI: 10.1186/s13007-016-0124-8

IF 3.51

Cápal *et al. Plant Methods (2016) 12:24*
DOI 10.1186/s13007-016-0124-8

Plant Methods

METHODOLOGY

Open Access

The utility of flow sorting to identify chromosomes carrying a single copy transgene in wheat

Petr Cápal[1], Takashi R. Endo[1,2], Jan Vrána[1], Marie Kubaláková[1], Miroslava Karafiátová[1], Eva Komínková[1], Isabel Mora-Ramírez[3], Winfriede Weschke[3] and Jaroslav Doležel[1*]

Abstract

Background: Identification of transgene insertion sites in plant genomes has practical implications for crop breeding and is a stepping stone to analyze transgene function. However, single copy sequences are not always easy to localize in large plant genomes by standard approaches.

Results: We employed flow cytometric chromosome sorting to determine chromosomal location of barley sucrose transporter construct in three transgenic lines of common wheat. Flow-sorted chromosomes were used as template for PCR and fluorescence in situ hybridization to identify chromosomes with transgenes. The chromosomes carrying the transgenes were then confirmed by PCR using DNA amplified from single flow-sorted chromosomes as template.

Conclusions: Insertion sites of the transgene were unambiguously localized to chromosomes 4A, 7A and 5D in three wheat transgenic lines. The procedure presented in this study is applicable for localization of any single-copy sequence not only in wheat, but in any plant species where suspension of intact mitotic chromosomes suitable for flow cytometric sorting can be prepared.

Keywords: Transgene localization, Flow cytometric sorting, Single chromosome amplification, *Triticum aestivum*, *Hordeum vulgare*, HvSUT1

Background

During the past 30 years, many cultivars of agricultural crops beneficial to humankind have been developed by means of genetic engineering, including plants resistant to herbicides, pests or viruses, bearing fruits with prolonged shelf life and products more suited for industrial processing [for review see 1]. Wheat ranks 5th in the commodities produced worldwide and is the second most-produced food crop occupying more than 50 % of the world crop area (http://faostat3.fao.org/). In the light of climate change and world population growth, future challenges for the increase of crop production have constantly been discussed. However, FAO statistics show that the wheat production is reaching a plateau and is severely affected by climate change. This is a consequence of a slowdown in wheat yield increase, accounting for only 0.5 % per year in the last decade [2].

Breeding improved cultivars with increased tolerance to adverse climatic conditions and with increased yield and quality could be facilitated by genetic engineering and introduction of beneficial genes from other organisms. The insertion site of a transgene is of great importance for the transgene function [3, 4] which is also influenced by its position on the chromosome, including the flanking DNA sequences [5]. However, transgene localization is not easy by routine approaches, like fluorescence in situ hybridization (FISH), or Southern blotting. A prevalent method for detection of transgenes in animals and plants is FISH, which has its pros and cons [6]. In barley and common wheat, FISH enables cytological localization of cDNAs, as short as 1.5 kb, on a

*Correspondence: dolezel@ueb.cas.cz
[1] Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 31, 78371 Olomouc, Czech Republic
Full list of author information is available at the end of the article

chromosome or chromosomes that had already been known to carry the cDNAs [7, 8]. Although some authors succeeded in localizing transgenes on plant chromosomes using FISH [9–11], this approach has not become a routine application.

Weichert et al. [12] obtained transgenic lines (HOSUT) of hexaploid wheat carrying barley (*Hordeum vulgare*) sucrose transporter HvSUT1 (SUT) gene that is overexpressed under the control of the endosperm-specific Hordein B1 promoter (HO). The HOSUT lines were found to increase grain yield significantly as compared to control non-transformed plants [13]. However, the genomic location of the transgene in these lines was not known. In the present work we employed a novel approach for unambiguous identification of chromosomes carrying the transgene in three HOSUT lines. The protocol takes the advantage of the availability of a procedure for flow cytometric chromosome sorting in wheat and the fact that flow-sorted chromosomes are suitable as templates for PCR and FISH [14]. Moreover, a protocol has been developed recently for representative DNA amplification from single copies of chromosomes [15]. By combining these approaches we could assign the transgene to particular chromosomes in three HOSUT lines of wheat.

Results and discussion

The experimental workflow is shown on Fig. 1. As the first step, we prepared liquid suspensions of intact mitotic chromosomes from all five lines of wheat (see "Methods" section) and analyzed them by flow cytometry. Monovariate flow karyotypes (histograms of relative fluorescence intensity) were obtained after the analysis of DAPI-stained chromosomes, and bivariate flow karyotypes obtained after the analysis of DAPI-stained chromosomes with FITC-labelled GAA microsatellites. We observed differences between flow karyotypes of the HOSUT lines and the model hexaploid wheat cultivar Chinese Spring. The alterations concerned the profiles of major composite peaks on monovariate flow karyotypes (Additional file 1: Figure S1) and the distribution of chromosome populations on bivariate flow karyotypes (Fig. 2). This observation reflected the differences in karyotypes (chromosome polymorphism) between the cultivar Certo, used to produce the HOSUT lines (data not shown), and Chinese Spring. On the other hand, flow karyotypes of the three HOSUT lines were indistinguishable from each other.

In order to identify chromosomes carrying the transgenes, we first used the approach described by Vrána et al. [16]. Fractions of 200 chromosomes were sorted from different regions of monovariate flow karyotypes as shown on Additional file 1: Figure S1, and DNA of the sorted chromosomes was used as template for PCR. This analysis identified one region (sort gate) in each line as representing chromosomes bearing a transgene. As each region (sort gate) on a monovariate flow karyotype may represent more than one chromosome type, in the next step we sorted chromosomes from regions delineated on bivariate flow karyotypes (Fig. 2). The sort gates were designed to include chromosome populations corresponding to the positive sort gates on monovariate flow karyotypes. From these regions, and also from nearby regions, chromosomes were sorted into PCR tubes (100 chromosomes per tube) and immediately afterwards also onto microscopic slides (ca. 1000 chromosomes per slide). The results obtained by PCR with primers amplifying HvSUT-RT sequence (Fig. 3) and identification of chromosomes from sort gates for each transgenic line by FISH with probes targeting Afa-repeat family and GAA-microsatellites (Fig. 4) are summarized in Table 1.

FISH analysis showed that more than 90 % of chromosomes flow-sorted from the region defined by the green rectangle consisted of one type of chromosome in each of the HOSUT lines. This fact together with the results of PCR suggested that the transgene was located on chromosome 7A in HOSUT 12/44, on chromosome 5D in HOSUT 20/6 and on chromosome 4A in HOSUT 24/31. In the former two lines, the critical type of chromosome was not found among the chromosomes flow-sorted from the region defined by red rectangles. However, chromosome 4A was found to represent 12.39 % of chromosomes flow-sorted from the red region in HOSUT 24/31. This was probably due to the similarities in size and the amount of GAA-FITC fluorescence of chromosomes 4A and 7A. Due to this similarity, mixture of the two chromosomes 4A and 7A was also observed in the chromosome fraction sorted from the green region in HOSUT 12/44.

To confirm chromosomal locations of the transgene and avoid ambiguous results due to possible contamination of flow-sorted fraction by other chromosomes, PCR was done on DNA amplified from single flow-sorted chromosomes. As each time only one copy of chromosome is sorted, the DNA cannot be contaminated by other chromosomes. Five single chromosomes were sorted from the green sort regions of the HOSUT lines and their DNA was separately amplified using multiple displacement amplification (MDA). Out of the five sorted chromosomes, whole genome amplification was successful with three chromosomes in HOSUT 12/44, two chromosomes in HOSUT 20/6 and four chromosomes in HOSUT 24/31. The successful amplification was defined by the production of measurable amount of DNA after MDA and by the presence of at least one marker for the

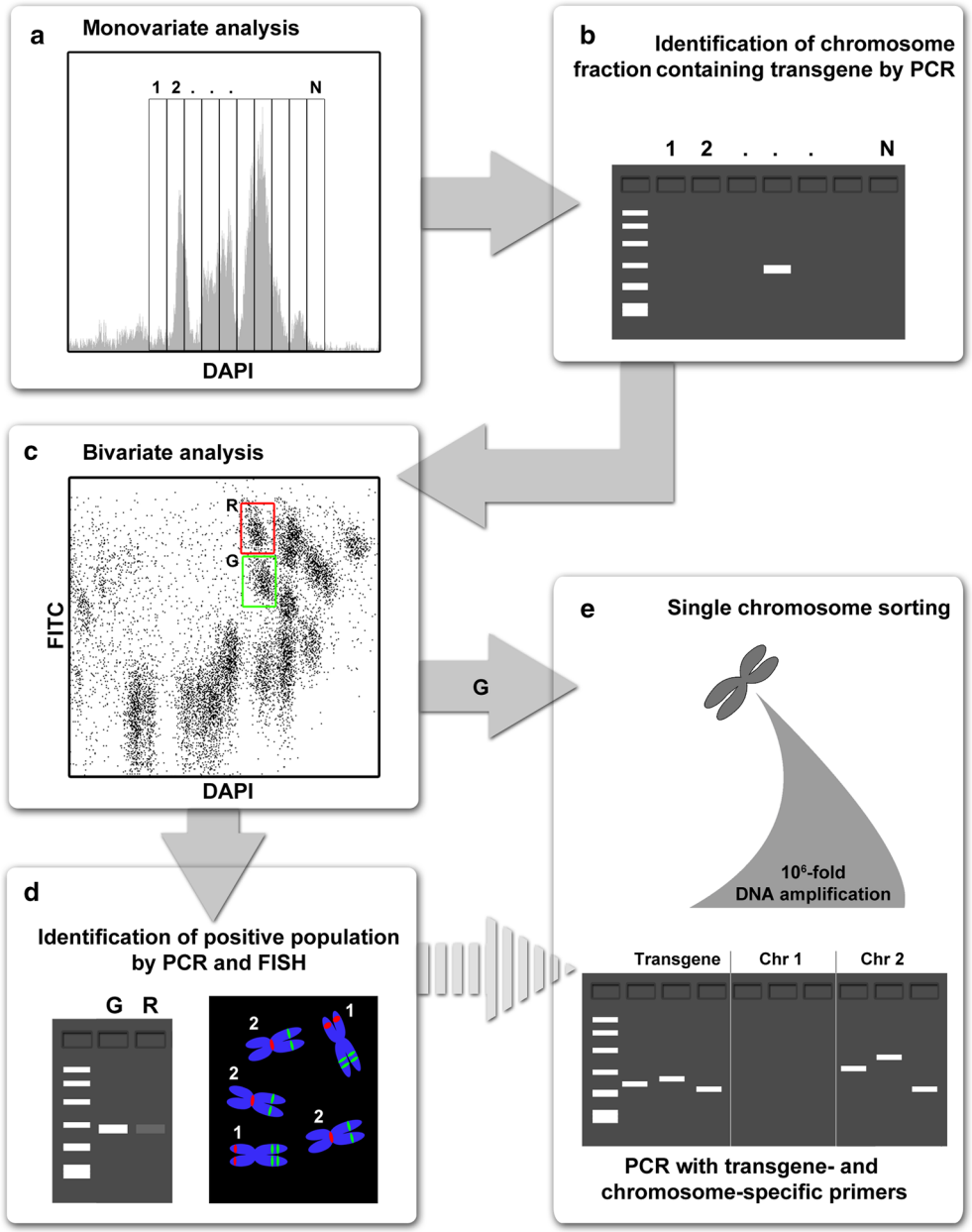


Fig. 1 Experimental workflow. **a** Monovariate flow karyotype is dissected into small regions. From each region, 200 chromosomes are sorted. **b** Flow-sorted chromosomes are used as template for PCR with transgene-specific marker. **c** The region representing chromosome with the transgene identified on monovariate flow karyotype is dissected by sorting chromosomes from sort regions on bivariate flow karyotype. **d** From the sort gates on the bivariate flow karyotype, 100 chromosomes are sorted for PCR with transgene-specific marker. Moreover, 1000 chromosomes are sorted immediately afterwards onto the microscopic slides to identify flow-sorted chromosomes by FISH. **e** From the sort gate most enriched for the transgene-bearing chromosome (region *G* in this example), single chromosomes are sorted individually into PCR tubes. DNA of single-flow sorted chromosomes is amplified and resulting DNA is used as template for PCR to identify the presence of multiple transgene- and chromosome-specific sequences. This step unambiguously confirms the chromosome with integrated transgene

transgene and one marker for the wheat chromosome. The reason for occasional failure to amplify DNA from single chromosomes, which was observed previously [15] is not clear. One explanation is that a droplet with sorted chromosome lands on side wall of PCR tube and the chromosome is excluded from the MDA reaction. The amount of chromosomal DNA in successfully amplified samples ranged from 0.3 to 1.7 µg DNA.

Chromosome specificity of sequence tagged site (STS) markers used in this work to identify individual

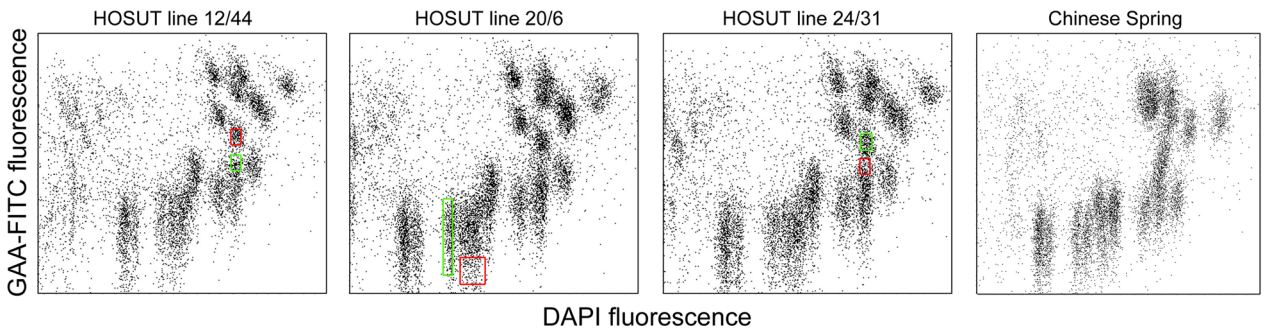


Fig. 2 Bivariate flow karyotypes of three transgenic HOSUT lines of wheat obtained after the analysis of chromosomes with FITC-labelled $(GAA)_n$ microsatellites and stained by DAPI. The position of *red* and *green* regions used to sort particular chromosomes is indicated. The *green* sort gate was found to represent chromosomes carrying transgene. Chromosomes were flow-sorted also from the neighboring population delineated by *red* gate and were used as a control. Although the transgene-bearing chromosome should not be included in this region, the sorted population could potentially be contaminated with transgene-bearing chromosomes due to similarity in chromosome size and DNA content

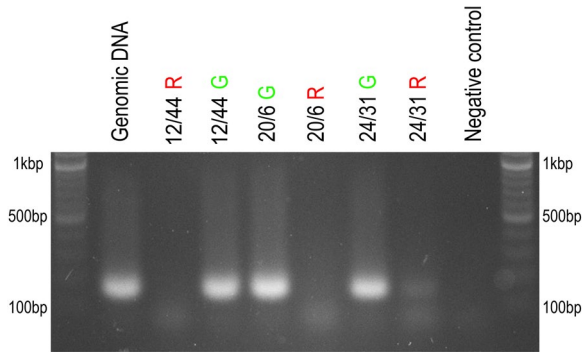


Fig. 3 Agarose gel electrophoresis of PCR products obtained with primers for the transgene and DNA of chromosomes flow-sorted from three HOSUT lines using the *green* and *red* sort regions as shown in Fig. 1. The amplicon of HvSUT-RT (169 bp) was obtained with chromosomes sorted from the *green* sort region in all three HOSUT lines. When chromosomes were sorted from the *red* sort regions, no PCR amplification occurred for HOSUT 12/44 and HOSUT 20/6. However, a weak band was observed for HOSUT 24/31. Genomic DNA of the transgenic lines served as positive control

chromosomes was first tested using the euploid and corresponding nulli-tetrasomic lines of Chinese Spring (Additional file 2: Figure S2). The results confirmed that the markers were suitable for unambiguous identification of wheat chromosomes 1A, 4A, 5D and 7A. PCR analysis using both transgene- and chromosome-specific markers clearly confirmed chromosome location of transgenes as determined in the first part of this study. In case of HOSUT 24/31, where the location of the transgene was ambiguous, all four transgene markers were detected in DNA amplified from single chromosomes sorted from the green region (Fig. 5), and all four 4A-specific markers were also amplified in the same amplicons. None of the four 7A-specific markers was found in the same amplicons.

Conclusions

Coupling PCR and FISH mapping using flow-sorted mitotic chromosomes as templates narrowed down the list of candidate chromosomes harboring the transgene to one or two chromosomes. PCR on DNA amplified

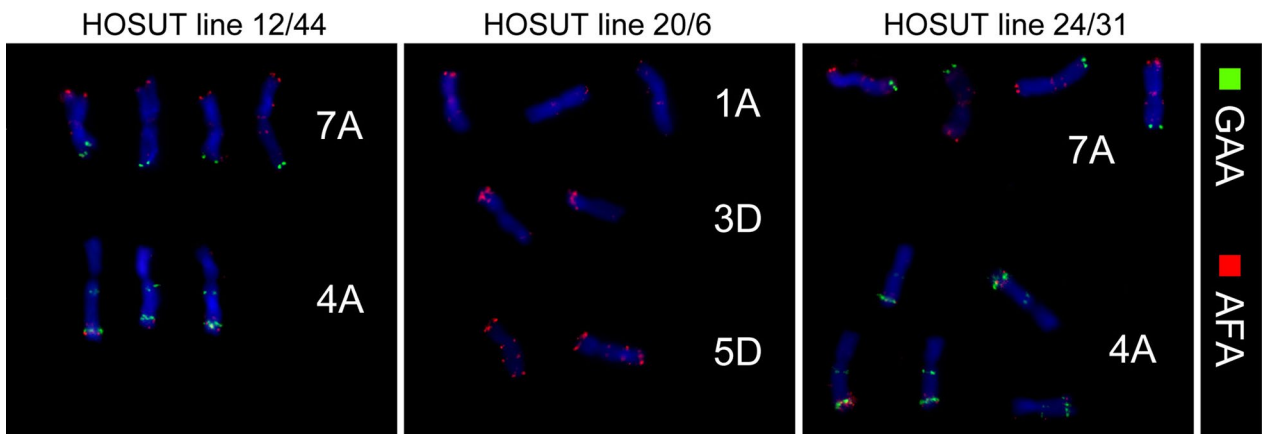


Fig. 4 Representative images of chromosomes flow-sorted from three HOSUT lines using the *green* and *red* sort regions on bivariate flow karyotypes as shown in Fig. 2. FISH was done using probes for Afa-family (*red* signals) and GAA microsatellites (*green* signals). Chromosomes were counterstained with DAPI (*blue*)

Table 1 PCR and FISH analysis of chromosomes sorted from each of the sort gates in three HOSUT lines

Transgenic line	Sort gate[a]	PCR result	Chromosomes identified FISH[c]
HOSUT 12/44	Red	Negative	4A (92.65 %)
	Green	Positive	7A (90.90 %)
			4A (4.45 %)
HOSUT 20/6	Red	Negative	1A (63.75 %)
			3D (36.25 %)
	Green	Positive	5D (94.66 %)
			1A (5.33 %)
HOSUT 24/31	Red	Semi-positive[b]	7A (83.19 %)
			4A (12.39 %)
			2A (4.42 %)
	Green	Positive	4A (97.30 %)

[a] Sort gates delineated with green and red rectangles in Fig. 2

[b] A faint band was visible after agarose gel electrophoresis of PCR product

[c] More than 1000 chromosomes were examined in each sorted fraction in each line

from single flow-sorted chromosomes then unambiguously identified the chromosomes with the integrated transgene. If chromosome-specific PCR-based markers are available, mapping on single copy chromosomes could be an ultimate approach to assign single copy DNA sequences, including transgenes, to particular chromosomes. Moreover, the sequence assembly of amplicons from the chromosome could allow detecting the position of transgene insertion, if enough sequence information on the chromosome is available. However the main purpose of this work was to assign a transgene to particular chromosomes. The approach presented here is currently applicable to more than 25 plant species, which include important cereals and legumes [14] where liquid suspensions of mitotic chromosomes suitable for flow cytometric sorting can be prepared.

Methods

Plant material

We used German winter wheat cultivar Certo (*Triticum aestivum* L., 2n = 6x = 42, genome formula AABBDD) and its three transgenic lines, HOSUT 12/44, HOSUT 20/6 and HOSUT 24/31. The transgenic lines contain a single copy of the HvSUT1-cDNA (1894 bp) fused to the barley HorB1 promoter (550 bp) and the barley HorB1 terminator (1663 bp) [12]. We also used euploid and nullisomic–tetrasomic (Nt1A1B, Nt7A7B, Nt5D5B) lines of hexaploid wheat cultivar Chinese Spring (obtained from NBRP-wheat) to confirm the specificity of PCR markers to particular wheat chromosomes.

Flow cytometric chromosome sorting

Cell cycle synchronization and metaphase accumulation of root tip meristem cells was performed as described

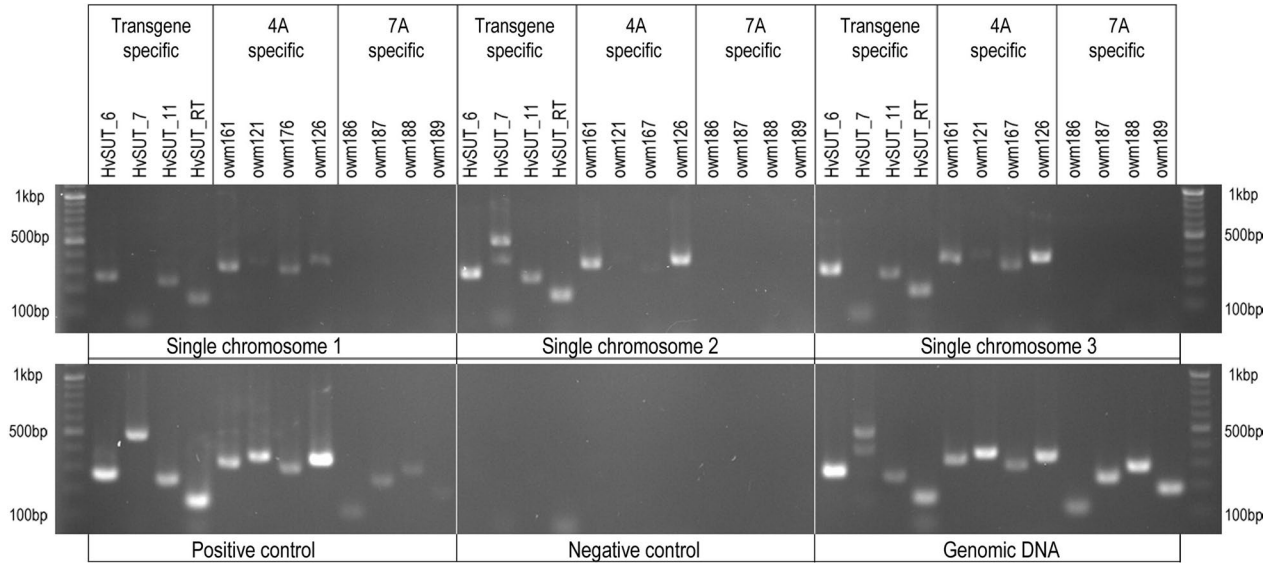


Fig. 5 Agarose gel electrophoresis of PCR products obtained using DNA produced by multiple displacement amplification of three single chromosomes flow-sorted from the sort region representing chromosome 4A in HOSUT 24/31 line. PCR with primers for the four transgenes resulted in products of expected length. The same was true for the chromosome 4A-specific STS markers. Note that none of the chromosome 7A-specific markers was detected in the samples of single chromosome DNA. PCR with genomic DNA of HOSUT 24/31 as template detected both 4A and 7A chromosome-specific markers. PCR with the positive control (represented by 1000 chromosomes sorted from *green* sorting region and amplified) showed slight PCR bands of chromosome 7A, which reflects a minor contamination of the sorted chromosome 4A by chromosome 7A

previously [17], except for the formaldehyde fixation, which was shortened to 15 min. Isolated chromosomes were labelled by FISHIS (fluorescence in situ hybridization in suspension) using FITC-labeled GAA probe following the protocol of Giorgi et al. [18]. Flow cytometric analysis and sorting was done on BD FACSAria II high speed flow sorter equipped with 390 nm laser for DAPI excitation and 488 nm laser for FITC excitation. Sort gates were initially drawn on monovariate flow karyotypes of DAPI fluorescence (not shown) and subsequently on bivariate flow karyotypes of DAPI fluorescence versus GAA-FITC fluorescence as shown in Fig. 1.

Fluorescence in situ hybridization (FISH)

For microscopic observations, 1000 chromosomes were sorted onto microscope slides from each of the sort regions. The slides were left to air-dry in the dark overnight. Then the preparations were used for FISH following the protocol of Kubaláková et al. [19] using a Cy5-labeled probe targeting Afa-family repeats, the chromosomes were already labeled by a GAA microsatellite probe during the FISHIS procedure.

PCR

PCR was done using primers specific for the HOSUT transgene and for markers specific for candidate wheat chromosomes (Table 2). Of the four HOSUT primers, three were designed in the HvSUT1 region (accession no. AJ272309) and one in the HorB1 terminator region (accession no. FN643080). Wheat chromosome-specific markers were designed by Primer3 based on the chromosome sequences from the International Wheat Genome Sequencing Consortium (IWGSC), while preventing the primers from amplifying the sequence from the homoeologous chromosomes. PCR conditions were set as follows: initial denaturation 95 °C for 3 min, 35 cycles of 30 s denaturation at 95 °C, annealing at 58–62 °C (see Table 2 for T_a of the primer pairs) for 30 s and extension at 72 °C for 30 s, followed by final extension at 72 °C for 5 min. The amount of template DNA was 5 ng for each reaction. PCR

Table 2 List of PCR primers for the HOSUT transgene construct and PCR primers for wheat STS markers on chromosomes 1A, 4A, 7A and 5D

Name	Target	Forward primer sequence	Reverse primer sequence	Amplicon size (bp)	Annealing temperature (°C)
HvSUT_6	HOSUT1	AGCGGCGGCGGTCACTGACTG	CCAAAGGACGACACCCCAGCC	265	62
HvSUT_7	HorB1 terminator	ATTAATTCCTCCCCGACCCTGC	CAATGGAGACGGCGCGTGCAA	471	62
HvSUT_11	HOSUT1	GGCGGAACCCGCCGTGCAG	CCTGCGTCTTCCCCATCTGGAAGTA	241	62
HvSUT_RT	HOSUT1	CGGGCGGTCGCAGCTCGCGTCTATT	CATACAGTGACTCTGACCGGCACACA	169	62
Owm121	Chromosome 4A	ATTGCCGTCGCGAACTAGA	CGGGACGAGCTTGACGAT	351	60
Owm126	Chromosome 4A	CCAGTCAGAAATTATTATGAACCTATC	CGCTGTCTCGAGATTGGAGT	342	60
Owm161	Chromosome 4A	TTTTCAAGCAGGTTTTGTGC	TCACTTCTCTTCTTTGCGTTCA	324	60
Owm167	Chromosome 4A	TTTTCTTGGTCAGTATAACCTGTTTTT	TGAGCAGAGAAAAATTTCCAAG	285	60
Owm174	Chromosome 1A	GCATCCTAGTTTCTCTCTCAAGT	AACAAGATCACGAGCGAATTG	157	58
Owm175	Chromosome 1A	AAACCCCTGATACTCATGCG	GTTTCTTGTCATTCATGTCACTTGT	530	58
Owm176	Chromosome 1A	TTCCTGTCTGACTCCGCG	AACCACAACCGTCAACCG	104	58
Owm177	Chromosome 1A	GTAGTCTGCTCCCGAGGAAT	GTCTCTAACCATACATCCATGAAGT	192	58
Owm178	Chromosome 1A	CAACTTCTTCACATCCCGGAA	ATTTGGCCCTATGAGATATAATTACG	306	58
Owm179	Chromosome 1A	ACACTGTGATACCTCTAGATGTATG	CACATTGCCTATAAATTCTAAAAGGTC	425	58
Owm180	Chromosome 5D	CGGACGAGCAGCAGTACC	GCAGATCGGCATAAATTGAATGT	292	58
Owm181	Chromosome 5D	GGAGGTGTTCTAGGTGTACTTACT	AGAGCAATGTCAGAAGTCATCG	240	58
Owm182	Chromosome 5D	TCTCCACCTGCAGAGTCG	CATCAGGCCACAGTGTCAAT	119	58
Owm183	Chromosome 5D	TGTCCACACATTTCCCGTATG	AGTGGTGGATGTGGTTGCT	196	58
Owm184	Chromosome 5D	AGCATGCTCCCAAAGACTATTAC	GTTATGATGGTGGTAGCAATTTGA	400	58
Owm185	Chromosome 5D	GTGAACCTATATGACATCTTACCGG	GGGGCAGTTGTCAAGTATTGC	421	58
Owm186	Chromosome 7A	CTCTCTGTGGCCAATAGTGC	TCTATACCTCAACCCTACATCCA	112	58
Owm187	Chromosome 7A	GGCCACGAATTCCACAAGTA	CTATCGATCAACCAACCATCCA	229	58
Owm188	Chromosome 7A	GTACGAGTGCAGACAGTGTG	ACAATTAATTATACGCCCAGTTAAGC	282	58
Owm189	Chromosome 7A	CGTGCTTTCTTCTTCCTCCG	GCAGGTTAGTTTCTTGTGGTTG	185	58
Owm190	Chromosome 7A	CGCATGGACATTGTTCTAGTCA	GCACTTAGGCACGCTTGAG	517	58
Owm191	Chromosome 7A	CGACGACATTAGGAATATGGGAT	TGCGTGTGGGTGTGCTTA	402	58

products were run on 1.5 % agarose gels. PCR using 100 sorted chromosomes as template was conducted after a few freeze–thaw cycles to disintegrate the chromosomes and the initial denaturation step was prolonged for 7 min.

Whole genome amplification of single chromosomes

DNA amplification of single chromosomes was performed by MDA using a GE Healthcare GenomiPhi V2 kit (GE Healthcare Life Sciences, Little Chalfont, UK) according to Cápal et al. [15]. Five individual chromosomes were flow-sorted into five 0.2 ml PCR tubes from green sort gates from each HOSUT line and their DNA amplified. The amplified DNA was evaluated on 1.5 % agarose gel, purified using magnetic beads (AMPure XP system, Beckman Coulter, Inc., Brea, CA, USA) and the concentration was measured by a spectrophotometer (NanoDrop, Thermo Fisher Scientific Inc., Waltham, MA, USA).

Additional files

Additional file 1: Figure S1. Flow karyotypes (histograms of fluorescence intensity) obtained after the analysis of DAPI-stained chromosomes isolated from three transgenic lines and cv. Chinese Spring of common wheat. Flow karyotypes of the transgenic lines are indistinguishable from each other, and slightly differ in profiles of the major composite peaks from those of Chinese Spring.

Additional file 2: Figure S2. Verification of marker specificity. PCR with a full set of chromosome-specific wheat STS markers was performed using genomic DNA of cv. Chinese Spring and corresponding nullitetrasomic lines for chromosomes 1A, 4A, 5D and 7A. The markers, which resulted in amplification products only in Chinese Spring and not in the nullitetrasomic lines, were used in this study.

Authors' contributions

The study was conceived and designed by TE, PC and JD, experiments were performed by PC, TE, M Ka, M Ku, JV, IMR and EK, manuscript was written by PC, TE, JD and WW. All authors read and approved the final manuscript.

Author details

[1] Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 31, 78371 Olomouc, Czech Republic. [2] Faculty of Agriculture, Ryukoku University, 1-5 Yokotani, Seta Oe-cho, Otsu, Shiga 520-2194, Japan. [3] Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Corrensstrasse 3, 06466 Stadt Seeland, Germany.

Acknowledgements

We thank Zdeňka Dubská for technical assistance with flow cytometric chromosome sorting. This work was supported by the National Program of Sustainability (Award No. LO 2014) and the Czech Science Foundation (Award No. P501-12-G090).

Competing interests

The authors declare that they have no competing interests.

Received: 20 February 2016 Accepted: 19 April 2016
Published online: 25 April 2016

References

1. Ahmad P, Ashraf M, Younis M, Hu X, Kumar A, Akram NA, Al-Qurainy F. Role of transgenic plants in agriculture and biopharming. Biotechnol Adv. 2012;30:524–40.
2. Fischer RA. The importance of grain or kernel number in wheat: a reply to Sinclair and Jamieson. Field Crops Res. 2008;105:15–21.
3. Svitashev S, Anaiev E, Pawlowski WP, Somers DA. Association of transgene integration sites with chromosome rearrangements in hexaploid oat. Theor Appl Genet. 2000;100:872–80.
4. Kumar S, Fladung M. Gene stability in transgenic aspen (Populus). II. Molecular characterization of variable expression of transgene in wild and hybrid aspen. Planta. 2001;213:731–40.
5. Kooter JM, Matzke A, Meyer P. Listening to the silent genes: transgene silencing, gene regulation and pathogen control. Trends Plant Sci. 1999;4:340–7.
6. Svitashev KS, Somers DA. Characterization of transgene loci in plants using FISH: a picture is worth a thousand words. Plant Cell Tiss Org Cult. 2002;69:205–14.
7. Danilova TV, Friebe B, Gill BS. Development of a wheat single gene FISH map for analyzing homoeologous relationship and chromosomal rearrangements within the Triticeae. Theor Appl Genet. 2014;127(3):715–30.
8. Karafiátová M, Bartoš J, Kopecký D, Ma L, Sato K, Houben A, Doležel J. Mapping nonrecombining regions in barley using multicolor FISH. Chromosome Res. 2013;21:739–51.
9. Pedersen C, Zimny J, Becker D, Jähne-Gärtner A, Lörz H. Localization of introduced genes on the chromosomes of transgenic barley, wheat and triticale by fluorescence in situ hybridization. Theor Appl Genet. 1997;94:749–57.
10. Dong J, Pushpa K, Cervera M, Hall TC. The use of FISH in chromosomal localization of transgenes in rice. Methods Cell Sci. 2001;23:105–13.
11. Khrustaleva LI, Kik C. Localization of single-copy T-DNA insertion in transgenic shallots (*Allium cepa*) by using ultra-sensitive FISH with tyramide signal amplification. Plant J. 2001;25:699–707.
12. Weichert N, Saalbach I, Weichert H, Kohl S, Erban A, Kopka J, Hause B, Varshney A, Sreenivasulu N, Strickert M, Kumlehn J, Weschke W, Weber H. Increasing sucrose uptake capacity of wheat grains stimulates storage protein synthesis. Plant Physiol. 2010;152:698–710.
13. Saalbach I, Mora-Ramírez I, Weichert N, Andersch F, Guild G, Wieser H, Koehler P, Stangoulis J, Kumlehn J, Weschke W, Weber H. Increased grain yield and micronutrient concentration in transgenic winter wheat by ectopic expression of a barley sucrose transporter. J Cereal Sci. 2014;60:75–81.
14. Doležel J, Vrána J, Cápal P, Kubaláková M, Burešová V, Šimková H. Advances in plant chromosome genomics. Biotechnol Adv. 2014;32:122–36.
15. Cápal P, Blavet N, Vrána J, Kubaláková M, Doležel J. Multiple displacement amplification of the DNA from single flow-sorted plant chromosome. Plant J. 2015. doi:10.1111/tpj.13035.
16. Vrána J, Kubaláková M, Číhalíková J, Valárik M, Doležel J. Preparation of sub-genomic fractions enriched for particular chromosomes in polyploid wheat. Biol Plant. 2015;59:445–55.
17. Vrána J, Kubaláková M, Šimková H, Číhalíková J, Lysák MA, Doležel J. Flow sorting of mitotic chromosomes in common wheat (*Triticum aestivum* L.). Genetics. 2000;156:2033–41.
18. Giorgi D, Farina A, Grosso V, Gennara A, Ceoloni C, Lucretti S. FISHIS: fluorescence in situ hybridization in suspension and chromosome flow sorting made easy. PLoS One. 2013;8:e57994.
19. Kubaláková M, Kovářová P, Suchánková P, Číhalíková J, Bartoš J, Lucretti S, Watanabe N, Kianian SF, Doležel J. Chromosome sorting in tetraploid wheat and its potential for genome analysis. Genetics. 2005;170:823–9.

APPENDIX V

Utilizing long-term and extensive virulence monitoring data as a basis for *in silico* identification of candidate avirulence effectors of barley powdery mildew fungus

Janáková E, Menardo F, Dreiseitl A, Bourras S, Doležel J, Valárik M, Keller B

Manuscript prepared for submission

Utilizing long-term and extensive virulence monitoring data as a basis for *in silico* identification of candidate avirulence effectors of barley powdery mildew fungus

Eva Janáková[1], Fabrizio Menardo[2], Antonín Dreiseitl[3], Salim Bourras[2], Jaroslav Doležel[1], Miroslav Valárik[1], Beat Keller[2]

[1]Institute of Experimental Botany of the Czech Academy of Sciences, Centre of the Region Haná for Biotechnological and Agricultural Research, Olomouc, Czech Republic

[2]Department of Plant and Microbial Biology, University of Zürich, Zürich, Switzerland

[3]Agrotest Fyto Ltd., Kroměříž, Czech Republic

Abstract

Sustainable control of plant pathogens based on disease resistance is highly desirable to reduce yield losses of cultivated crops. However, employing effective and durable resistance requires prior identification of plant and pathogen interacting components and understanding their interplay. Nowadays, affordable whole genome sequencing provides a basis for dissecting core genes involved in defense responses. Annotation of *Blumeria graminis* f. sp. *hordei* genome revealed a presence of nearly five hundred genes encoding putative candidate effectors. In an effort to predict effectors with potential avirulence function, we conducted a genome wide association study on a set of 90 isolates employing 453,114 high-confidence SNP markers and 35 distinct phenotypes gathered over years of virulence monitoring. Setting stringent criteria yielded ten highly significant associations and closer inspection detected a candidate effector for seven of them including *Mla1*, *Mla3*, *Mla10*, *Mla22*, *Mlg*, *Mlk1* and *MlRu2*. In three cases, the effector was suggested to act as a suppressor of avirulence. A single candidate with presumably dual function designated Avr_{a10}/Svr_{a22} indicated existence of complex interactions influencing avirulence in barley-powdery mildew fungus pathosystem. Complete or partial deletion of effector gene frequently observed among identified candidates exposed a potential pitfall of SNP-based association analysis. This work opens new perspectives for high-throughput identification of avirulence effectors in barley powdery mildew fungus.

Introduction

Protection of agriculture crops from biotic and abiotic stresses is essential to maintain productivity or even secure its growth with respect to constantly increasing human population. Plant-pathogenic fungi causing serious crop diseases belong among the most important groups of biotic stress agents. Nowadays, sustainability has been generally acknowledged as the necessary aspect of any crop protective strategy. A combination of currently available knowledge and resources therefore provides opportunities for accelerated and efficient disease resistance research and breeding (Boyd *et al.* 2013; Dangl *et al.* 2013). Although quantitative resistance has been highly appreciated for its durability, its genetic basis remains poorly understood. The only underlying genes cloned up to now in wheat suggest involvement of diverse mechanisms (Krattinger *et al.* 2009; Fu *et al.* 2009; Moore *et al.* 2015). On the other hand, deployment of major resistance (*R*) genes in breeding has a very long tradition and a growing number of such genes encoding mostly nucleotide-binding leucine-rich repeat (NB-LRR) proteins has been cloned even in complex genomes of cereals (e.g. Sánchez-Martín *et al.* 2016; Steuernagel *et al.* 2016; Yahiaoui *et al.* 2004; Zhou *et al.* 2001). Although the current zigzag model of plant pathogen interactions (Jones and Dangl 2006) provides a satisfying general explanation of limited durability of *R* gene-based resistance frequently observed in agricultural ecosystems (McDonald and Linde 2002), elucidating the molecular basis of these interactions is still a work in progress. An inevitable task consists in identification of pathogen effectors specifically recognized by plant receptors encoded by *R* genes, avirulence (*AVR*) proteins. To date, several hypotheses have been proposed to explain mechanism of such a recognition which finally results in hypersensitive response via activation of downstream signaling pathways. The original elicitor-receptor model (Albersheim and Anderson-Prouty 1975) proposed direct interaction of plant *R* gene product (receptor) located in the host plasma membrane with pathogen *Avr* gene product (elicitor) exposed on the surface of the pathogen to explain an incompatible reaction between the pathogen and the host. Later, van der Biezen and Jones (1998) suggested one particular *Avr* gene product having a specific plant pathogenicity target (a guardee) safeguarded by one matching *R* protein. According to this "guard hypothesis" based on indirect perception, resistance is activated by the *R* protein depending on its interaction with a guardee modified by the pathogen. In spite of accumulating evidence supporting the guard model (Jones and Dangl 2006), new and inconsistent data resulted in

formulation of a decoy model (van der Hoorn and Kamoun 2008). This model presumes evolution of a host protein termed "decoy" with the only function consisting in mimicking effector target to ensure pathogen recognition in the presence of respective *R* protein. The most recent "integrated decoy" hypothesis (Cesari *et al.* 2014) presents an extension of the previous model which considers pairs of plant NB-LRR receptors in situations where both partners are necessary to confer resistance and one of them contains additional non-conserved domain. The decoy is fused to one of the NB-LRR receptors and acts as a bait for pathogen effector which is then directly recognized by the second receptor. Since formulation of this hypothesis, variable integrated domains have been recognized in around 10 % of plant NB-LRR proteins thus paving the way to identify targets of pathogen effectors (Sarris *et al*. 2016).

Powdery mildew fungus of grasses and cereals (*Blumeria graminis*) represents both a causal agent of a serious disease affecting some important crops and a model organism for research on different aspects of fungal pathogens with obligate biotrophic lifestyle. Although available reference genome sequences of *B. graminis formae speciales* (ff. spp.) *hordei* and *tritici* (Spanu *et al.* 2010; Wicker *et al.* 2013) reach only a draft sequence quality, they provided a valuable insight into evolutionary dynamics, gene composition and effector complement of this species. In general, bioinformatic prediction of fungal effectors can be based on several features including presence of an N-terminal secretion signal, relatively small size (usually ≤300 amino acids), frequent cysteine-rich composition and absence of known orthologous proteins outside the genus (Lo Presti *et al.* 2015). In addition, Godfrey *et al.* (2010) identified a common Y/F/WxC motif present in the N-terminal part of *Blumeria graminis* f. sp. *hordei* (*Bgh*) candidate effectors. Characterization of putative effector complements in both sequenced ff. spp. of *B. graminis* revealed that they account for a significant proportion of all identified genes (Kusch *et al.* 2014; Pedersen *et al.* 2012; Spanu *et al.* 2010; Wicker *et al.* 2013). Recently, several of these *in silico* predicted effectors were shown to be involved in avirulence (Bourras *et al.* 2015; Lu *et al.* 2016; Parlange *et al.* 2015; Praz *et al.* 2017) thus adding evidence to the bioinformatic prediction. However, interaction mode of identified avirulence effectors with their corresponding *R* genes is yet to be clarified. Although uncovering a complete set of core components involved in defense response in this pathosystem remains challenging, monitoring effector complement of pathogen populations represents a key for informed deployment of cultivars and engineering durable resistance to achieve sustainable disease control (Dangl *et al*. 2013). This study presents

a potentially high-throughput strategy for *in silico* identification of avirulence effectors benefitting from phenotype monitoring conducted over years on a large set of *Bgh* isolates. Based on results of performed genome-wide association study (GWAS), different aspects of this approach are discussed along with future directions applicable to maximize the usage of available data and the number of identified candidates.

Materials and methods

Pathogen isolates

90 isolates of *Blumeria graminis* f. sp. *hordei* of different geographical origin (Table 1) held at the Agrotest Fyto Ltd. in Kroměříž, Czech Republic were included in the collection selected for association analysis. Isolates were maintained on first leaf segments of seedlings of susceptible barley variety Stirling placed in Petri dishes containing 0.8% agar with 40 $mg \cdot l^{-1}$ benzimidazole. Virulence patterns of individual isolates were evaluated on differential barley varieties (Dreiseitl *et al.* 2006; Dreiseitl 2011a; Dreiseitl and Kosman 2013; Dreiseitl *et al.* 2013; Dreiseitl 2015a). Altogether, 35 distinct phenotypes (Table 2) in a binary data format representing virulence or avirulence were considered for this study. Intermediate infection type was treated as missing data.

DNA extraction and sequencing of Bgh isolates

Each pathogen isolate was propagated as described above to obtain sufficient amount of conidia. Genomic DNA was extracted according to the protocol described by Komínková *et al.* (2016). DNA libraries from individual samples were constructed using the Nextera DNA Library Prep Kit (Illumina Inc., San Diego, CA, USA) and 300 bp or 250 bp paired-end reads were generated by the Illumina MiSeq or HiSeq Sequencing Systems, respectively. Acquired Illumina reads were quality trimmed by Trimmomatic v. 0.32 (Bolger *et al.* 2014) with parameters LEADING:3, TRAILING:3, SLIDINGWINDOW:4:20, MINLEN:50 and the same tool was used to remove Illumina adapter sequences.

Table 1 Origin and sequencing of *Blumeria graminis* f. sp. *hordei* isolates analyzed in this study.

Bgh isolate	Origin*	Year of collection	No. of R1 reads	Genome coverage#	*Bgh* isolate	Origin	Year of collection	No. of R1 reads	Genome coverage
0004	CHN	2003	4,851,638	20.22	7777	CZE	2007	5,315,624	22.15
0020	ZAF	2004	6,190,831	25.80	A17/16	CZE	2016	4,030,481	16.79
0023	AUS	2005	4,031,914	16.80	A1/15	CZE	2015	10,150,564	42.29
0061	URY	2005	5,259,526	21.91	A20/16	CZE	2016	3,870,930	16.13
0235	SWE	1976	2,892,363	12.05	A2/15	CZE	2015	16,568,769	69.04
0323	SWE	1975	4,794,570	19.98	A7/15	CZE	2015	10,871,026	45.30
0422	ZAF	2007	3,528,981	14.70	A7/16	CZE	2016	6,014,985	25.06
1002	DEU	before 1999	3,435,717	14.32	B1-2/14	CZE	2014	6,103,325	25.43
1044	JPN	1953	5,887,634	24.53	B2-2/14	CZE	2014	6,726,354	28.03
1765	CZE	2009	5,462,910	22.76	D2/16	CZE	2016	3,410,011	14.21
1767	CZE	2010	5,248,842	21.87	E2/15	CZE	2015	8,660,780	36.09
1775	CZE	2011	8,247,309	34.36	H-148	ISR	1979	4,407,244	18.36
2567	CZE	2004	4,252,839	17.72	HE4/15	CZE	2015	9,653,038	40.22
3775	CZE	2009	3,969,949	16.54	HE5/15	CZE	2015	9,329,935	38.87
4154	CZE	2010	5,650,110	23.54	I1/15	CZE	2015	9,323,614	38.85
4404	SWE	before 1976	4,670,266	19.46	I3/16	CZE	2016	2,770,139	11.54
4505	CZE	2014	7,577,237	31.57	J-462	ISR	1979	4,594,383	19.14
4517	CZE	2009	3,474,562	14.48	K2/15	CZE	2015	12,574,817	52.40
4531	CZE	2013	7,076,469	29.49	K3/15	CZE	2015	5,627,292	23.45
4535	CZE	2010	4,718,559	19.66	KM10/15	CZE	2015	19,282,674	80.34
4574	CZE	2011	5,859,267	24.41	KM16/15	CZE	2015	5,817,577	24.24
4575	CZE	2010	4,359,052	18.16	KM2/15	CZE	2015	11,052,703	46.05
4711	CZE	2007	5,614,080	23.39	KM3/15	CZE	2015	10,843,074	45.18
4745	CZE	2008	4,683,682	19.52	KM4/15	CZE	2015	10,864,941	45.27
4761	CZE	2001	6,211,363	25.88	KM6/15	CZE	2015	12,744,994	53.10
4773	CZE	2009	5,639,525	23.50	L10/14	CZE	2014	8,242,561	34.34
5375	CZE	2013	5,740,671	23.92	L7/15	CZE	2015	6,372,322	26.55
5435	CZE	2012	4,679,537	19.50	M12/15	CZE	2015	5,798,014	24.16
5515	CZE	2014	9,029,339	37.62	M13/15	CZE	2015	5,675,719	23.65
5535	CZE	2013	4,214,550	17.56	M18/15	CZE	2015	5,626,096	23.44
5541	CZE	2012	5,656,897	23.57	M20/15	CZE	2015	5,171,831	21.55
5551	CZE	2014	6,966,823	29.03	M3/16	CZE	2016	3,594,981	14.98
5565	CZE	2014	6,090,717	25.38	M4/15	CZE	2015	9,599,847	40.00
5715	CZE	1997	4,519,122	18.83	M4/16	CZE	2016	3,288,960	13.70
5765	CZE	1999	3,738,319	15.58	M5/16	CZE	2016	4,830,399	20.13
5774	CZE	2009	5,874,937	24.48	M6/15	CZE	2015	6,087,835	25.37
5775	CZE	2009	5,430,328	22.63	N5/16	CZE	2016	3,430,010	14.29
6040	DNK	1986	5,649,549	23.54	N7/14	CZE	2014	9,023,221	37.60
6045	CZE	1999	4,537,851	18.91	O11/16	CZE	2016	3,048,173	12.70
6545	CZE	2011	6,009,944	25.04	O4/14	CZE	2014	9,514,766	39.64
6577	CZE	2010	4,092,732	17.05	O9/16	CZE	2016	2,998,896	12.50
7455	CZE	2009	6,071,608	25.30	S-016	ISR	1997	5,387,406	22.45
7515	CZE	2012	4,302,046	17.93	Y-035	ISR	1979	5,740,050	23.92

Table 1 Continued.

Bgh isolate	Origin*	Year of collection	No. of R1 reads	Genome coverage#	*Bgh* isolate	Origin	Year of collection	No. of R1 reads	Genome coverage
7541	CZE	2014	6,465,529	26.94	Y-069	ISR	1979	4,314,738	17.98
7557	CZE	2002	5,082,716	21.18	Z1/15	CZE	2015	5,059,420	21.08

*Three-letter country codes indicate origin of individual isolates: CHN = China; ZAF = Republic of South Africa; AUS = Australia; URY = Uruguay; SWE = Sweden; DEU = Germany; JPN = Japan; CZE = Czech Republic; DNK = Denmark; ISR = Israel.

#Mean sequencing genome coverage is 26x.

Generation of SNP markers

Quality trimmed sequencing reads of each isolate were aligned on available genomic scaffolds of *Bgh* strain DH14 (Spanu *et al.* 2010) using Bowtie 2 v. 2.2.9 (Langmead and Salzberg 2012) in end-to-end mode with a maximum of three mismatched base pairs per 100 bp. Alignments were converted into the BAM format, sorted and indexed by SAMtools utilities (Li *et al.* 2009) v. 1.3, the same package was used to remove duplicated reads. Read groups were added to all individual BAM files using Picard tools (http://broadinstitute.github.io/picard/). Subsequently, all BAM files were merged by BamTools toolkit v. 1.0 (Barnett *et al.* 2011) and after indexing by SAMtools, variant calling was performed using FreeBayes v. 9.9.2 (Garrison and Marth 2012). Only position with a minimal mapping quality of 30, minimal base quality of 20, minimal frequency of the alternative allele call of 0.8, minimal quality sum of observations supporting an alternate allele within a single individual of 30, minimal alternate allele count of observation of 1 and 2 within a single individual and within the total population, respectively, were considered. An original set of detected variants in the VCF format was further filtered by VCFtools v. 0.1.15 (Danecek *et al.* 2011) to select biallelic SNPs with a minor allele frequency of 0.02, mean depth values less than 70 and a maximum of 0.1 missing data. In the second step of filtering, vcffilter utility of the vcflib library (Garrison *et al.* submitted) was applied with parameters "QUAL > 1 & SAF > 0 & SAR > 0". A VFC file containing a set of filtered SNP markers was converted into the HapMap format by an in-house Perl script. Sequence scaffolds of the DH14 reference containing only single SNP were eliminated.

Table 2 Phenotypes analyzed within GWAS realized in this study. Altogether, reactions of *Bgh* isolates (Table 1) to 21 resistance genes or alleles and 14 varieties carrying mostly unknown resistances were considered. The parentheses in designation of some resistance genes indicate their tentativeness – e. g. unavailable segregation data or unexplored possibility that the gene is allelic to a named gene.

RESISTANCE GENES			RESISTANCES		
Resistence gene/allele	**Chromosome**	**Reference**	**Differential variety**	**Resistence gene/allele**	**Reference**
Mla1	1H	Briggs and Stanford (1938)	Burštyn	unknown	Dreiseitl (2015b)
Mla3	1H	Moseman and Schaller (1960)	Camilla	*Ml(SI-1)*	designation first used in Germany in 2000
Mla6	1H	Moseman *et al.* (1965)	Gilberta	unknown	Dreiseitl (2013a)
Mla7	1H	Scholz and Nover (1967)	KM-14/2010	unknown	Dreiseitl (2015a)
Mla9	1H	Moseman and Jørgensen (1971)	NORD 07017/69	unknown	Dreiseitl (2015b)
Mla10	1H	Moseman and Jørgensen (1971)	Oowajao	unknown	Dreiseitl (2015b)
Mla12	1H	Giese *et al.* (1981)	Prosa	*Mlg*, unknown	Dreiseitl (2013b)
Mla13	1H	Giese *et al.* (1981)	Sara (Sv 83342)	*Mla3*, *Ml(Tu2)*	Jensen *et al.* (1992)
Mla22	1H	Jørgensen (1991)	SBCC097	unknown	Silvar *et al.* (2013)
MlaN81	1H	Brückner (1987)	SC 44801 N2 (Pop)	unknown	Dreiseitl (unpublished)
Mlat	1H	Moseman and Schaller (1960)	SJ 123063	unknown	Dreiseitl (2015a)
Mlg	4H	Honecker (1931)	SK-4770	unknown	Dreiseitl (2015a)
Ml(IM9)		Jensen *et al.* (1992)	Spilka	unknown	Dreiseitl and Krizanova (2012)
Mlk1	1H	Briggs and Stanford (1938)	Venezia	*Ml(Ve)*	Dreiseitl (2018)
MlLa	2H	Giese et al. (1993)			
Ml(Ln)		Dreiseitl (2011a)			
Ml(Lv)		Dreiseitl (2011b)			
Mlp1	1H?	Stanford and Briggs (1940)			
Ml(Ro)		Dreiseitl (2011c)			
MlRu2	1H	Wiberg (1974)			
Ml(St)		designation first used in Denmark around 1995			

Analysis of population structure

Population structure of selected pathogen isolates was inferred using the Structure software (Pritchard *et al.* 2000) v2.3.4 from a reduced subset of 10,069 SNP markers. Each 45th SNP from the total set was included to avoid linkage between markers. The analysis was carried out assuming an admixture model of population structure and correlated allele frequencies. The number of clusters (*K*) in the range from one to ten was tested using 200,000 burn-in periods followed by 1,000,000 Markov chain Monte Carlo (MCMC) iterations and 10 independent runs were carried out for each *K* to evaluate the variation of the likelihood estimates. The optimal number of clusters was determined using the ΔK method as described in Evanno *et al.* (2004) to detect the uppermost hierarchical level of population structure which was subsequently visualized by Structure Plot v2.0 (http://omicsspeaks.com/strplot2/; Ramasamy *et al.* 2014). Additionally, a principal component analysis (PCA) was carried out using the complete set of SNP markers to further elucidate pathogen population structure. This analysis was implemented in the Genome Association and Prediction Integrated Tool (GAPIT; Lipka *et al.* 2012) R package.

GWAS and identification of candidate effectors

Associations between a filtered set of SNP markers in the HapMap format and avirulence phenotypes were evaluated by the R package GAPIT. In case an informative association was detected, a region of relevant DH14 sequence scaffold surrounding the top-ranking SNP by 100 kbp on either side was manually examined for presence of predicted Candidate Secreted Effector Proteins (CSEPs; Kusch *et al.* 2014; Pedersen *et al.* 2012; Spanu *et al.* 2010) using the Integrative Genomics Viewer (IGV) visualization tool (Robinson *et al.* 2011; Thorvaldsdottir *et al.* 2013). Each CSEP identified within the range was subjected to closer inspection regarding presence of any DNA polymorphism showing correlation with given phenotype.

Results

Pathogen resequencing and generation of SNP markers

Genomes of a total of 90 pathogen isolates were sequenced with an estimated genome coverage of row reads ranging from 11.54x to 80.34x (26x in average, Table 1). Quality-trimmed reads were mapped onto the DH14 reference genome (Spanu *et al.* 2010) and variant calling on the merged alignment data from all isolates yielded 1,107,880 polymorphic sites. Two subsequent filtering steps resulted in a set of 453,114 high-confidence biallelic SNP markers.

Analysis of population structure

In order to provide insight into the population genetic structure of analyzed pathogen isolates, a Bayesian algorithm implemented in the software STRUCTURE was employed. Since the determination of correct number of clusters based solely on the highest value of $L(K)$ (an estimate of posterior probability of the data for a given K) can be misleading as demonstrated by Evanno *et al.* (2005), the ΔK method was selected as an appropriate statistic to estimate the number of clusters. The highest value of ΔK (4,433.5) was observed at K = 2 (Fig. 1) implying that the uppermost level of population structure is best characterized by two distinct groups (Fig. 2). The first one includes a vast majority of isolates while a group of five Israeli isolates along with an ancient isolate originating from Japan (1044, also known as Race I) has been clearly separated. Due to its unique character, this subgroup is hereafter referred to as "group 2".

Additionally, PCA on the complete set of SNP markers revealed a clear separation of isolates based on their geographical origin with a potential influence of temporal origin (Table 1). The first two principal components accounted for 11.5 and 3.6 % of genetic variation, respectively. PC1 discriminated the same group of isolates that had been detached also by STRUCTURE analysis whereas PC2 separated all other isolates collected in different countries than the Czech Republic with two exceptions: Swedish isolates 0235 and 0323 clustered with predominating Czech isolates (PC1 -31.2 and -36.5, PC2 -50.4 and -59.9, respectively). Apart from the major cluster of Czech isolates, three outliers (KM3/15, KM4/15 and KM6/15) were markedly separated by PC2 as well (Fig. 3). Altogether, composite results obtained by PCA and

the STRUCTURE software suggest existence of several hierarchical levels of population structure within the analyzed set of pathogen isolates.

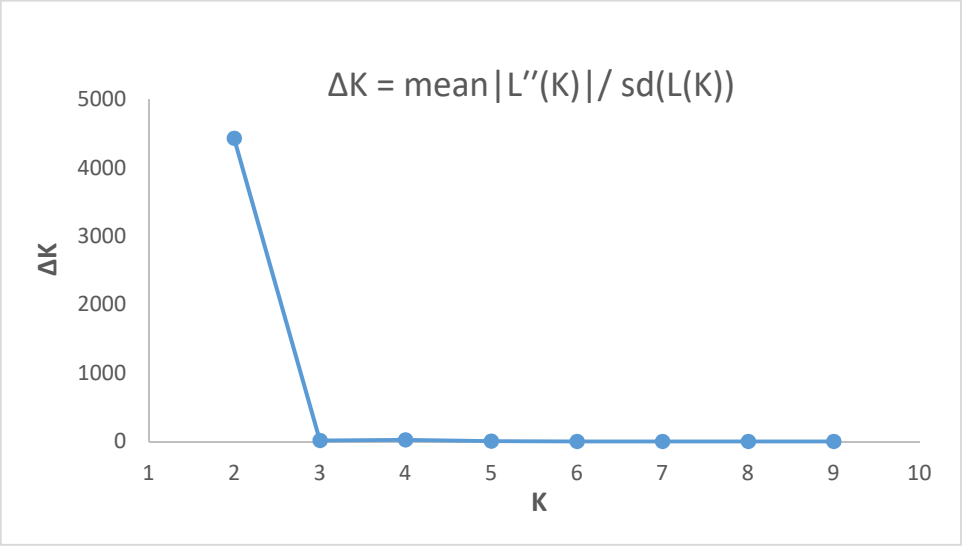


Fig. 1 Distribution of ΔK values calculated according to Evanno *et al.* (2005) for K ranging from two to nine clusters. The *ad hoc* ΔK statistic is based on the second order rate of change in the log probability of data between successive K values. The maximal value of ΔK distribution reached at $K = 2$ indicates that the uppermost level of population structure consists of two groups. sd = standard deviation

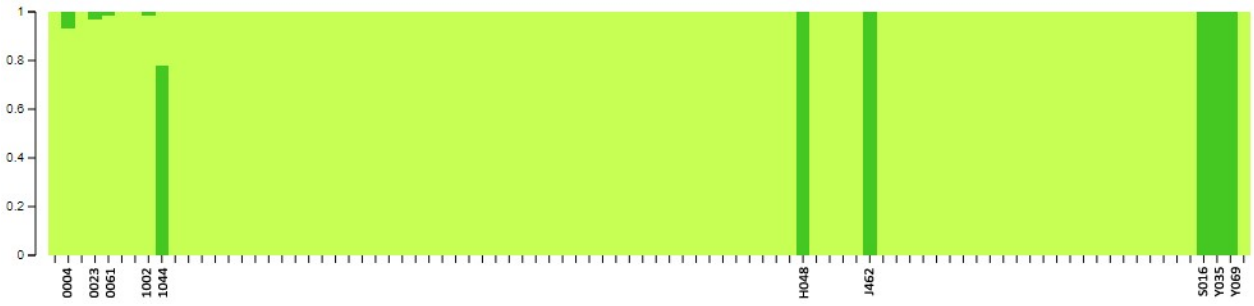


Fig. 2 The uppermost hierarchical level of population structure of 90 pathogen isolates based on 10,069 SNP markers. Segment of different colors represent membership fractions to each of two detected clusters.

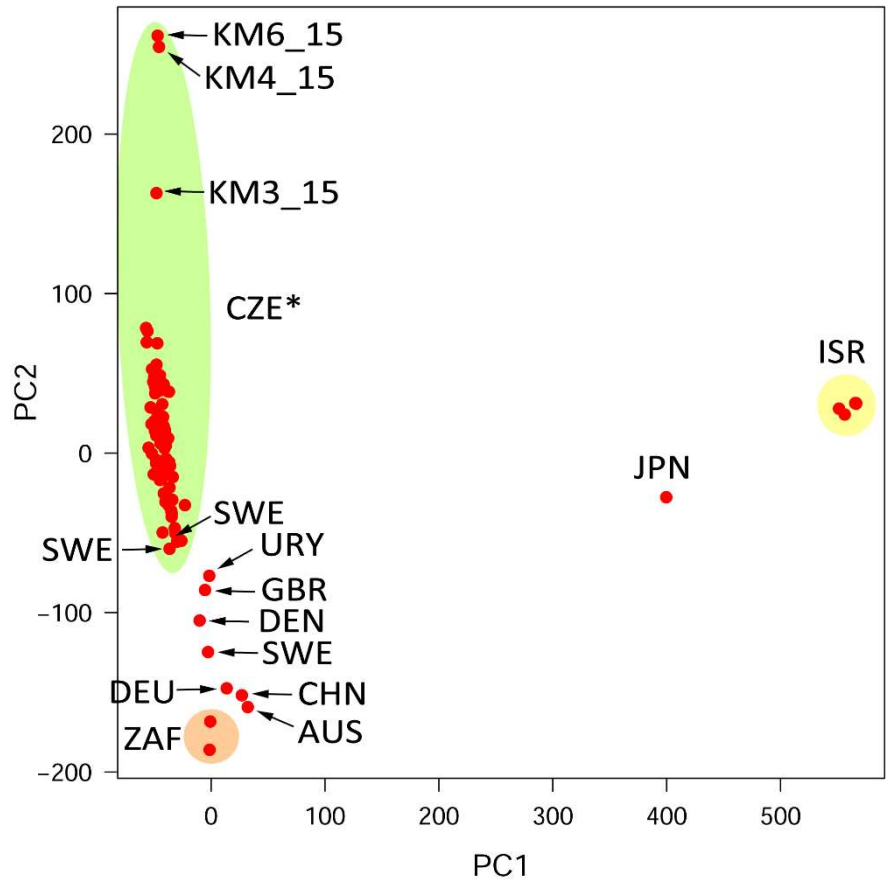


Fig. 3 Principal component analysis plot based on 453,114 SNP markers illustrating 90 *Blumeria graminis* f. sp. *hordei* isolates included in the whole genome association study and a reference strain DH14 (GBR). Three-letter country codes indicate origin of individual isolates. A cluster of Czech isolates marked by an asterisk includes also two isolates originating from Sweden, 0235 and 0323 characterized by PC1 -31.2 and -36.5, PC2 -50.4 and -59.9, respectively. Remaining isolates collected in other countries than CZE were clearly separated. Besides their geographical origin, analyzed isolates differ by their year of collection (Table 1).

Genome-wide association study

A GWAS on 453,114 SNP markers and 35 traits (Table 2) was carried out to identify pathogen effectors involved in avirulence. Detected associations with avirulence phenotypes were considered significant only for False Discovery Rate (FDR)-corrected *P*-values (Benjamini and Hochberg 1995) below 0.001. Following this criterion, a total of 15 suggestive marker-trait associations were identified. Out of them, associations with *Mla6*, *Mlap1* and variety Spilka were

excluded due to extreme deviations in corresponding quantile-quantile (QQ) plots (a plot illustrating relationship between the observed distribution of the test statistic and the expected, null distribution) suggesting systemic bias. Such a bias can be caused by population substructure confounding the GWAS model which consequently may not sufficiently account for spurious associations. Additionally, association with *Mla12* was excluded as well due to low number of SNPs with significant FDR-adjusted *P*-values and high background signals resulting in absence of clear peak. Lastly, apart from the same problems as in case of *Mla12*, association with variety SC 44801 N2 was not considered owing to a rare incidence of virulent phenotype in the dataset (only four isolates). Manhattan plots and corresponding QQ plots of remaining 10 high-confidence associations are displayed in Fig. 4 and Fig. 5, respectively. In the next step, Manhattan plots were visually inspected and solitary SNPs with small *p*-values localized either in regions with large *p*-values or within very short sequence scaffolds (below 3,000 bp) were excluded from further analysis due to their potentially spurious character or limited use for identification of candidate effector, respectively. Only large and distinct peaks demonstrating strong association of the genomic region with the trait were therefore considered. The best-correlated SNP marker for each high-confidence peak identified within all 10 associations is described in Table 3.

Identification of candidates for avirulence effectors

Following manual inspection of genomic regions determined by the best correlated SNPs, all previously annotated CSEPs (Kusch *et al.* 2014; Pedersen *et al.* 2012; Spanu *et al.* 2010) present within 100 kbp on either side were subjected to closer examination in terms of possessing polymorphism correlating with phenotype. A single candidate for avirulence effector was identified for seven associations (*Mla1*, *Mla3*, *Mla10*, *Mla22*, *Mlag*, *Mlk1* and *MlRu2*), all of them were characterized by the presence of a signal peptide. Both SNPs and presence/absence variation (PAV) were detected as a putative source of phenotype variation. In three cases (*Mla22*, *Mlg* and *Mlk1*), function of suppressor of avirulence was proposed based on PAV pattern in individual virulent/avirulent isolates. Suggested candidates for avirulence effectors are summarized in Table 4.

A previously predicted effector gene designated *bgh00029* was identified as the only candidate for *Avr*$_{a1}$. The same conclusion was drawn by Lu *et al.* (2016) who performed also functional validation of this newly identified avirulence gene. The gene is located 38,842 bp from the top-ranking SNP and contains one intron. A putative effector protein encoded by this gene consists of 118 amino acids (AA). A potential causal SNP was detected in position 320,085 of scaffold no. 24. Cytosine in avirulent isolates substituted for guanine in virulent isolates caused an amino acid change from proline to arginine. However, isolates of group 2 (all of them virulent to *Mla1*) were found to carry adenine at the same position resulting in glutamine in translated polypeptide sequence. This SNP therefore was not included in the marker set employed for GWAS due to its multiallelic nature. Besides this rare variant, two more nonsynonymous SNPs were detected within *bgh00029* allele of group 2 and one of isolates carried also approximately 20 bp deletion in 1st exon. The predictive potential of GWAS was well demonstrated by the second and the third best SNP localized only 2,662 bp and 124 bp downstream *Avr*$_{a1}$, respectively.

Detailed analysis of the *Mla3* association revealed *bgh00316* gene to be the only candidate for *Avr*$_{a3}$. Unlike *Avr*$_{a1}$, this gene consists of three exons and is longer (228 AA). In this case, a GWAS top-ranking SNP in position 124,723 of scaffold 51 was found to be at the same time a putative causal polymorphism located in the second exon of the gene. The nonsynonymous SNP leads to AA substitution G163C (position including a signal peptide is given) in virulent isolates. An exception to this pattern was observed in three isolates (6040, KM3_15 and Y069), however, all of them contained different SNP or deletion accountable for loss of avirulence. Additionally, multiple other SNPs were detected in isolates of group 2.

A potential candidate for avirulence effector *Avr*$_{a10}$, *bgh03730*, resembles *Avr*$_{a1}$ in both sequence length and exon-intron structure. This gene localized at the beginning of scaffold 2 (the longest one in *Bgh* reference genome assembly) was detected within a distance of mere 1,307 bp from the GWAS top-ranking SNP. Interestingly, it was found to be completely deleted in 42 isolates. The size of observed deletion was estimated to be ~6 kbp based on IGV visualization of mapped reads and absence of effector gene coincides with virulence. Moreover, a nonsynonymous SNP was detected in second exon of *bgh03730* gene in remaining isolates and isolates with consequent F77I substitution showed virulent phenotype as well. However, five isolates (4535, 5765, 5774, 5551, A7_16) did not support this observation.

Unexpectedly, the same effector gene, *bgh03730*, was identified in case of *Mla22* association. However, since its presence correlated with virulence (regardless the previously described SNP variant) and absence with avirulence of corresponding isolates, it is tempting to conclude that this time, the GWAS yielded a candidate suppressor of avirulence, Svr_{a22}. Detailed inspection revealed only two isolates (5775 and D2_16) which contradicted this hypothesis.

Second potential suppressor of avirulence was detected in case of *Mlk1* association. Putative Svr_{k1}, a predicted effector gene *bghG000389000001001* composed of two exons, was localized only 2,212 bp from the top-ranking SNP detected in scaffold no. 2. As in previous case, PAV pattern was in good concordance with virulence and avirulence. This time, the size of deletion was found to be ~4 kbp in most isolates. Only three exceptions (isolates 0023, 0422 and L10_14) were observed and two of them (0023 and 0422) showed atypically large deletions of ~45 and ~52 kbp, respectively.

The last candidate for suppressor of avirulence, *bgh03686* localized only 419 bp from the top-ranking SNP, was found to best explain the *Mlg* association. Again, a well-correlating PAV pattern was observed for ~800 bp deletion including most of the first exon and complete second exon of the candidate gene. The only detected exception was isolate 0023 with virulent phenotype in spite of deleted Svr_{8}.

Finally, a candidate for *Avr* gene corresponding to *MlRu2 R* gene was identified. Proposed Avr_{Ru}, *bgh03736*, begins 1,081 bp downstream of the GWAS best SNP. A short indel in the first exon was suggested as putative causal polymorphism, detailed examination revealed expansion of trinucleotide repeat 'AAT' in virulent isolates. Avirulent isolates shared the sequence length with the DH14 reference and only one copy of the motif was present. On the other hand, two or more copies were observed in virulent isolates. Exceptions to this rule included isolates 1775, B2-2_14, L10_14, KM10_15 and O9_16 which showed reference-like allele but virulent phenotypes. Moreover, large deletions of the complete locus (~12 or ~14 kbp) including candidate Avr_{Ru} were proposed to be responsible for virulent phenotype of isolates 1044 and 7455, A7_15, E2_15, respectively.

Table 3 Summary of top-ranking SNP markers detected by GWAS based on the strongest marker-trait association with avirulence to resistance genes/alleles. Only large and distinct peaks of high-confidence associations were considered.

Resistance gene/allele	Marker	DH14 reference scaffold	Position	P-value	FDR-adjusted P-value
Mla1	snp325694	24	281,149	3.95E-13	1.79E-07
Mla3	snp402845	51	124,723	1.73E-12	7.82E-07
Mla10	snp20574	2	139,273	8.00E-11	1.68E-05
Mla13	snp158596	7	1,524,941	2.28E-09	9.98E-04
Mla22	snp20572	2	139,151	4.32E-14	1.96E-08
Mlat	snp399157	47	358,861	2.73E-09	7.74E-05
	snp425124	183	35,448	2.73E-09	7.74E-05
	snp400280	49	157,685	1.09E-07	7.68E-04
Mlg	snp123074	5	1,321,238	5.15E-12	2.34E-06
Mlk1	snp20781	2	180,557	1.23E-13	1.85E-08
	snp393100	44	418,740	1.44E-07	4.86E-04
Ml(Lv)	snp316523	22	1,682,725	1.86E-09	5.38E-05
MlRu2	snp347016	28	427,f636	6.43E-10	1.07E-04

Table 4 Summary of candidate effectors potentially involved in avirulence towards corresponding *Ml* genes/alleles tested in GWAS. AVR = avirulence effector; SVR = suppressor of avirulence.

Phenotype	Candidate effector*	Putative function	Scaffold	Position	Distance from top-ranking SNP	No. of exons	No. of AA
Mla1	*bgh00029/CSEP0008*	AVR	24	319,991–320,421	38,842 bp	2	118
Mla3	*bgh03316/CSEP0096*	AVR	51	124,470–125,273	0	3	228
Mla10	*bgh03730/CSEP0141*	AVR	2	137,552–137,966	1,307 bp	2	118
Mla13	-	-	-	-	-	-	-
Mla22	*bgh03730 CSEP0141*	SVR	2	137,552–137,966	1,185 bp	2	118
Mlat	-	-	-	-	-	-	-
Mlg	*bgh03686/CSEP0128*	SVR	5	1,321,657–1,322,137	419 bp	2	134
Mlk1	*bghG000389000001001/CSEP0285*	SVR	2	177,993–178,345	2,212 bp	2	99
Ml(Lv)	-	-	-	-	-	-	-
MlRu2	*bgh03736/CSEP0145*	AVR	28	428,717–429,130	1,081 bp	2	119

* The first designation gives ID in Blugen database (http://www.blugen.org), the second one CSEP ID

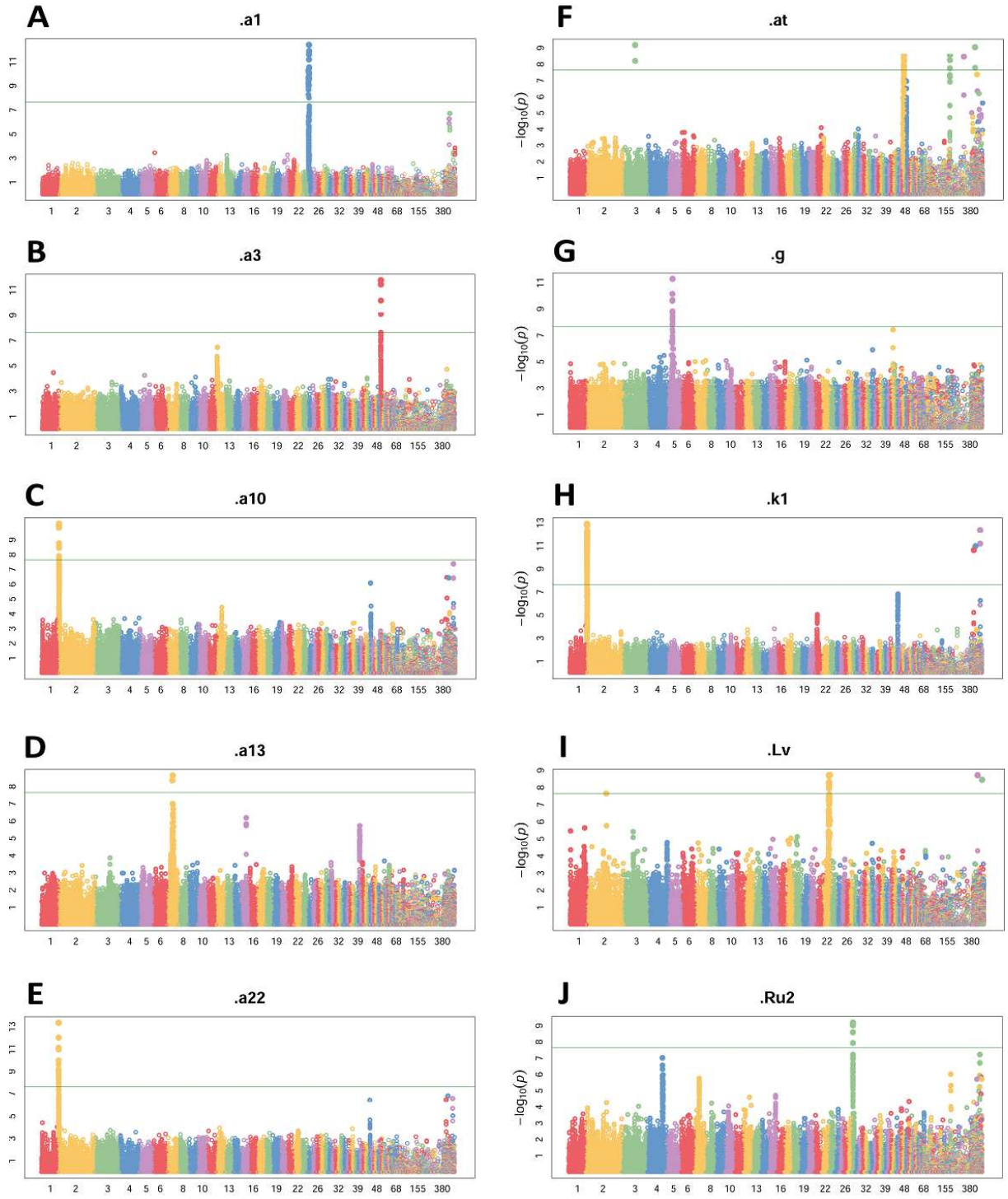


Fig. 4 Manhattan plots providing a graphical summary of GWAS results for (**A**) *Mla1*, (**B**) *Mla3*, (**C**) *Mla10*, (**D**) *Mla13*, (**E**) *Mla22*, (**F**) *Mlat*, (**G**) *Mlg*, (**H**) *Mlk1*, (**I**) *Ml(Lv)* and (**J**) *MlRu2*. The X-axis represents positions of 453,114 tested SNP markers within sequence scaffolds marked by different colors, the Y-axis is the negative base 10 logarithm of the *P*-values (from the F-test testing null hypothesis of no association between the SNP marker and trait).

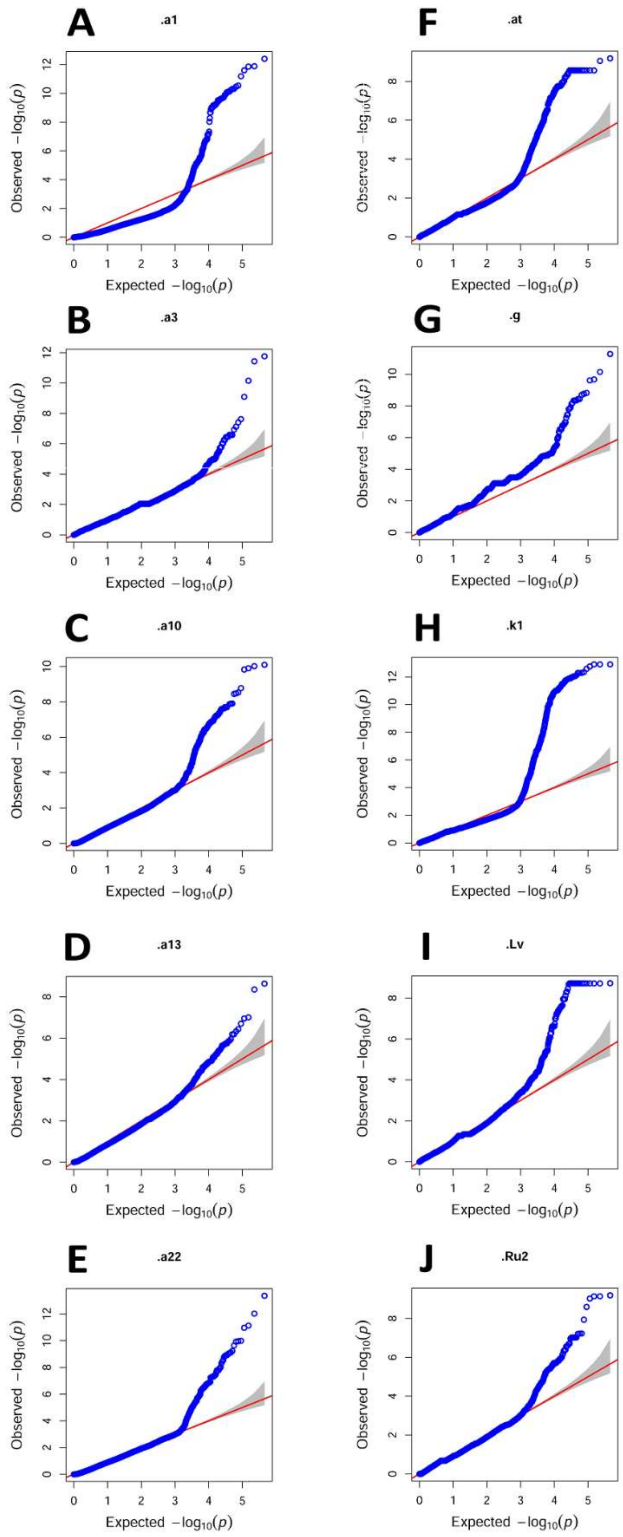


Fig. 5 The QQ plots demonstrating distribution of *P* values corresponding to GWAS results for **A**) *Mla1*, (**B**) *Mla3*, (**C**) *Mla10*, (**D**) *Mla13*, (**E**) *Mla22*, (**F**) *Mlat*, (**G**) *Mlg*, (**H**) *Mlk1*, (**I**) *Ml(Lv)* and (**J**) *MlRu2*. The Y-axis represents negative base 10 logarithm of the observed *P*-values, the X-axis shows negative base 10 logarithm of the expected *P*-values under the assumption of a uniform [0,1] distribution.

Discussion

During recent years, the onset of bioinformatic era has been turning the spotlight on whole-genome sequence data. However, while limitations in sequencing large and complex genomes have been gradually eliminated, elucidating the genetic basis of different traits keeps strict requirements with respect to obtaining phenotype data of appropriate quality and amount – a task that may become a bottleneck of such analysis. Here, we demonstrate that years of extensive virulence monitoring proved to be an invaluable resource for seeking candidate effectors of barley powdery mildew fungus which are involved in race-specific resistance interactions.

An important consideration when performing association analysis is population structure of the analyzed dataset. Including geographically distant individuals has the potential to maximize the genetic variation but also to introduce genetic heterogeneity. As a consequence, a non-causative marker can show stronger association with phenotype than the real underlying variants (Korte and Farlow 2013). Analysis of population structure within our set of *Bgh* isolates revealed presence of two major groups of unequal size (Fig. 2). A small subgroup of six isolates collected between 1953 and 1997 in Israel and Japan showed extreme divergence and generally, different polymorphisms were observed during manual inspection of predicted candidate effector genes than in remaining isolates. High number of SNPs and unique haplotypes in this distinct subgroup could be a reason to completely exclude such isolates from the GWAS to avoid confusion of the results (as practiced e.g. by Lu *et al.* 2016). However, by including even such a divergent group, we demonstrated the power and robustness of our analysis. A finer level of population subdivision uncovered by PCA showed clear separation of most isolates originating from different countries from the main cluster comprising Czech isolates (Fig. 3). An influence of population stratification on GWAS results leading to spurious associations may be indicated by inspection of QQ plots. Indeed, we observed mild (Fig. 5) to more serious (results not shown) deviations from theoretical distribution. To deal with this problematics and optimize the association analysis, different approaches and models might be tested and compared (e.g. Segura *et al.* 2012).

In this study, large-scale whole genome resequencing of *Bgh* isolates was performed to mine intraspecific variation in search of candidate effectors associated with avirulence

phenotypes. We conducted the GWAS and examined polymorphism in determined regions. Employing a total of 90 pathogen isolates, 453,114 SNP markers and 35 distinct resistance phenotypes resulted in the most extensive GWAS reported so far in *B. graminis*. Setting stringent criteria for filtration of results yielded 10 high-confidence associations representing a success rate of 28.6 %. For majority of these selected associations, a putative candidate potentially involved in avirulence interaction could be identified. Interestingly, two candidates (Avr_{a1} and Svr_{g}) show structural homology to ribonucleases (Pedersen *et al.* 2012) which was reported also for Avr_{a13} (Lu *et al.* 2016) and *AvrPm2* (Praz *et al.* 2017). Comparison of detected associations with available genetic mapping data based on analysis of segregating populations may be indicative of correctness of our results. According to Brown and Jessop (1995), linkage was detected between Avr_{a10} and Avr_{k1}. Moreover, Caffier *et al.* (1996) found that these two genes were very closely linked to Avr_{a22} as well. These findings are consistent with our results suggesting a single candidate gene for Avr_{a10}/Svr_{a22} localized 40 kpb from candidate Svr_{k1}.

Although the GWAS was bases only on SNP markers, the most frequent type of polymorphism observed within candidate genes was complete or partial indel. Point mutation was suggested as the sole causal polymorphism only in two cases and once, expansion of trinucleotide repeat motif was observed. Loss of avirulence function by various mechanisms including transposable element insertion, segmental or complete deletion and point mutation has been reported in other phytopathogenic fungi as well (e.g. Zhang *et al.* 2015). It is obvious that in such a situation, relying solely on SNP markers when performing GWAS can be misleading. A clear illustration of confusion this may produce is *Mla13* association. Based on the described procedure of manual inspection, we failed to identify any corresponding candidate for avirulence effector. However, later on, Lu *et al.* (2016) reported successful identification and validation of Avr_{a13} represented by previously predicted effector gene *bghG002861000001001/CSEP0372* and a 326-bp insertion close to 3' end of the gene was accounted for the loss of avirulence phenotype. Subsequent re-examination of our data revealed that Avr_{a13} is localized 144,365 bp from the GWAS top-ranking SNP. Such a distance was out of set range but other SNPs were found much closer, i.e. 2,791 bp, 66,785 bp, 10,322 bp or even 236 bp from the gene. Altogether, four of the ten best SNPs occurred within the range of 100 kbp. This example points out the necessity of more complex evaluation of available results since other types of polymorphism may be elusive to SNP-based GWAS.

The consequence of frequent indels affecting *B. graminis* avirulence genes (as found by this study, Lu *et al.* 2016 and Praz *et al.* 2017) consists in another serious limitation of this approach. When evaluating detected association, we relied completely on *in silico* predicted genes encoding candidate secreted effector proteins (Kusch *et al.* 2014; Pedersen *et al.* 2012; Spanu *et al.* 2010) which had been annotated in reference genome sequence of *Bgh* strain DH14 (Spanu *et al.* 2010). However, such annotation is dependent on presence of avirulence (or suppressor of avirulence) gene within this strain and in case of their partial or complete deletion, even precise GWAS result is not sufficient to identify a relevant candidate. The solution of this problem consists in employing annotated assemblies of different isolates with contrasting virulence patterns as a reference for read mapping. However, obtaining a high-quality assembly of such a repeat rich genome (~64 % of transposable elements was estimated by Spanu *et al.* 2010) remains a challenging task. Consequently, the current reference sequence assembly is highly fragmented and contains many gaps which can possibly harbor some genes of interest surrounded by repetitive elements (as shown by Plissonneau *et al.* 2016). Altogether, these limitations may explain lacking candidate genes involved in avirulence for three out of ten high-quality associations analyzed in this study.

If we focus on analyzed phenotypes with no significant association detected, several explanations are applicable. The first one may concern especially differential varieties which carry more than one resistance gene with major impact on observed phenotype. In such a case, GWAS may not be sufficient to clearly indicate genomic position of multiple components. Most likely, this explains the fact that all high-confidence associations identified within this study are connected with response to specific resistance gene or allele rather than to varieties with unknown or multiple sources of resistance. Second, some phenotypes may not have yielded association of appropriate quality due to their extremely rare occurrence within the pathogen population which presents problems for GWAS (Korte and Farlow 2013). And finally, some avirulence phenotypes of plant pathogenic fungi are mediated by complex interactions with more than one fungal gene involved (Bourras *et al.* 2016; Petit-Houdenot and Fudal 2017). With increasing number of cloned fungal avirulence determinants, exceptions to the gene-for-gene model (Flor 1971) have been reported including such that may not be captured by GWAS. To date, different scenarios have been described. Ma *et al.* (2015) identified a pair of interacting fungal effector genes expressed under shared promotor and both of them were required for

immunity conferred by single *R* gene in *Fusarium oxysporum* f. sp. *lycopersici* and tomato pathosystem. An intriguing example of complex interaction is represented by dual function of single pathogen effector. This scenario has been proposed by Houterman *et al.* (2008) and Plissonneau *et al.* (2016) who demonstrated that an effector secreted by plant pathogenic fungus can both trigger and suppress *R* gene-based immunity. In compliance with these observation, a common Avr_{a10}/Svr_{a22} candidate was identified in this study. Interestingly, Plissonneau *et al*. (2017) found that certain point mutations in such *Avr* gene which prevent recognition by corresponding *R* gene do not interfere with a suppression function of the same effector. Again, this is in agreement with the case of Avr_{a10}/Svr_{a22} effector where a nonsynonymous SNP co-occurred with a loss of avirulence towards *Mla10* but only presence-absence polymorphism of the effector appeared to be linked to its function suppressing avirulence towards *Mla22*. The evolutionary basis of this complex interaction can be explained by conservation of certain effector required for full virulence. Since evasion of host recognition through gene loss would inflict a fitness penalty to pathogen, it has evolved another effector capable of suppressing the recognition (Houterman *et al.* 2008).

Even within some of the high-confidence associations with putative avirulence or suppressor gene identified, several exceptions disturbing otherwise clear correlation pattern were registered. They could be generated by various mechanisms including those acting on different than DNA level, such as transcriptional, post-transcriptional or post-translational regulation. Another potential cause consists in presence of more components influencing the avirulence interactions in the genetic background of such isolates. Accordingly, five analyzed isolates with lost or mutated candidate Avr_{a10} showed avirulent phenotype on *Mla10* barley. Such a relatively high fraction can be explained by a presence of additional *Avr* factor or *Avr* modifier polymorphic in global population as suggested by McNally *et al.* (2018) who observed similar ratio of unexpectedly avirulent isolates with presumably inactive *Avr* allele in powdery mildew of wheat. Moreover, in case of two putative suppressors of avirulence (Svr_{k1}, Svr_{g}), several isolates with partially or completely deleted suppressor gene showed virulent phenotype. This could be caused by lost or mutated third component in the system, the unknown *Avr* gene. It is very tempting to hypothesize that at least in case of Svr_{k1}, avirulence gene might be localized in close proximity of the suppressor since both isolates presenting this exception (0023 and 0422) carry much larger deletion than remaining virulent isolates (45 or

52 kbp versus 4 kbp, respectively) which affects also two other CSEPs. On the other hand, four virulent isolates with intact candidate *Avr*$_{Ru2}$ gene suggest presence of unknown active suppressor in respective genomes.

To demonstrate the reliability of results presented in this work, it is essential to perform functional validation assays with identified candidate genes. Employing *Agrobacterium*-mediated transient expression (agroinfiltration) in *Nicotiana benthamiana* leaves was described by Ma *et al.* (2012) as an efficient approach for co-expression of an *R* gene with a corresponding *Avr* gene which triggers host-defence response. Later on, this protocol was successfully adopted for *B. graminis* ff. spp. (Bourras *et al.* 2015, Praz *et al.* 2017). However, it requires prior cloning of respective *R* gene or allele. In barley, only a number of *Mla* alleles have been cloned so far (Halterman *et al.* 2001, Halterman *et al.* 2003, Halterman and Wise 2004, Seeholzer *et al.* 2010, Shen *et al.* 2003, Zhou *et al.* 2001). For us, this makes possible validation of *Avr*$_{a1}$, *Avr*$_{a3}$ and *Avr*$_{a10}$. Since *Avr*$_{a1}$ has been already validated (Lu *et al.* 2016), it can provide a positive control. Validation of *Svr*$_{a22}$ remains problematic in spite of available *Mla22* allele since no respective avirulence gene affected by this suppressor has been identified so far. Moreover, some of *R* proteins encoded by the *Mla* locus were found to require a zinc-binding protein *RAR1* for their function (Jørgensen 1996). This complex interaction may interfere with validation by agroinfiltration in the *N. benthamiana* system and different approach should be employed such as agroinfiltration directly in barley, as demonstrated by Lu *et al.* (2016).

References

Albersheim P, Anderson-Prouty AJ (1975) Carbohydrates, proteins, cell surfaces, and the biochemistry of pathogenesis. Annu Rev Plant Phys 26:31-52

Barnett DW, Garrison EK, Quinlan AR, Strömberg MP, Marth GT (2011) BamTools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics 27:1691-1692

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate – a practical and powerful approach to multiple testing. J Roy Stat Soc B Met 57:289-300

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-2120

Bourras S, McNally KE, Ben-David R, Parlange F, Roffler S, Praz CR, Oberhaensli S, Menardo F, Stirnweis D, Frenkel Z, Schaefer LK, Flückiger S, Treier G, Herren G, Korol AB, Wicker T, Keller B (2015) Multiple avirulence loci and allele-specific effector recognition control the *Pm3* race-specific resistance of wheat to powdery mildew. Plant Cell 37:2991-3012

Bourras S, McNally KE, Müller MC, Wicker T, Keller B (2016) Avirulence genes in cereal powdery mildews: the gene-for-gene hypothesis 2.0. Front Plant Sci 7:241

Boyd LA, Ridout C, O'Sullivan DM, Leach JE, Leung H (2013) Plant-pathogen interactions: disease resistance in modern agriculture. Trends Genet 29:233-240Briggs FN, Stanford EH (1938) Linkage of factors for resistance to mildew in barley. J Genet 37:107-117

Brown JKM, Jessop AC (1995) Genetics of avirulences in *Erysiphe graminis* f. sp. *hordei*. Plant Pathol 44:1039-1049

Brückner F (1986) Possibilities of the use of the 'Nepal 81' cultivar to spring barley breeding for resistance to powdery mildew. Genet Šlecht 22:97-102 (in Czech)

Caffier V, de Vallavieille-Pope C, Brown JKM (1996) Segregation of avirulences and genetic basis of infection types in *Erysiphe graminis* f. sp. *hordei*. Phytopathology 86:1112-1121

Cesari S, Bernoux M, Moncuquet P, Kroj T, Dodds PN (2014) A novel conserved mechanism for plant NLR protein pairs: the "integrated decoy" hypothesis. Front Plant Sci 5:606

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, 1000 Genomes Project Analysis Group (2011) The Variant Call Format and VCFtools. Bioinformatics 27:2156-2158

Dangl JL, Horvath DM, Staskawicz BJ (2013) Pivoting the plant immune system from dissection to deployment. Science 341:746-751

Dreiseitl A (2006) Powdery mildew resistance of winter barley varieties in Czech official trials. Czech J Genet Plant 42:38-43

Dreiseitl A (2011a) Changes in the population of *Blumeria graminis* f. sp. *hordei* in the Czech Republic from 2009 to 2010. Plant Protect Sci 47:43-51

Dreiseitl A (2011b) Presence of the newly designated powdery mildew resistance Landi in some winter barley cultivars. Czech J Genet Plant 47:64-68

Dreiseitl A (2011c) Resistance of 'Laverda' to powdery mildew and its presence in some winter barley cultivars. Cereal Res Commun 39:569-576

Dreiseitl A (2011d) Resistance of 'Roxana' to powdery mildew and its presence in some European spring barley cultivars. Plant Breeding 130:419-422

Dreiseitl A (2012) Identity of barley powdery mildew resistances *Bw* and *Ru2*. Czech J Genet Plant 48:185-188

Dreiseitl A (2013a) Genes for resistance to powdery mildew in European winter barley cultivars registered in the Czech Republic and Slovakia to 2010. Plant Breeding 132:558-562

Dreiseitl A (2013b) Postulation of genes for resistance to powdery mildew in spring barley cultivars registered in the Czech Republic from 1996 to 2010. Euphytica 191:183-189

Dreiseitl A (2015a) Changes in virulence frequencies and higher fitness of simple pathotypes in the Czech population of *Blumeria graminis* f. sp. *hordei*. Plant Protect Sci 51:67-73

Dreiseitl A (2015b) Rare virulences of barley powdery mildew found in aerial populations in the Czech Republic from 2009 to 2014. Czech J Genet Plant 51:1-8

Dreiseitl A (2018) Resistance of barley variety 'Venezia' and its reflection in *Blumeria graminis* f. sp. *hordei* population. Euphytica DOI: 10.1007/s10681-018-2123-5

Dreiseitl A, Dinoor A, Kosman E (2006) Virulence and diversity of *Blumeria graminis* f. sp. *hordei* in Israel and in the Czech Republic. Plant Dis 90:1031-1038

Dreiseitl A, Fowler RA, Platz GJ (2013) Pathogenicity of *Blumeria graminis* f. sp. *hordei* in Australia in 2010 and 2011. Australas Plant Path. 42:713-721reiseitl A, Kosman E (2013) Virulence phenotypes of *Blumeria graminis* f. sp. *hordei* in South Africa. Eur J Plant Pathol 136:113-121

Dreiseitl A, Kosman E (2013) Virulence phenotypes of *Blumeria graminis* f. sp. *hordei* in South Africa. Eur J Plant Pathol 136:113-121

Dreiseitl A, Križanová K (2012) Powdery mildew resistance genes in spring barley cultivars registered in Slovakia from 2000 to 2010. Cereal Res Commun 40:494-501

Evanno G, Regnaut S, Goudet J (2004) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611-2620

Flor HH (1971) Current status of gene-for-gene concept. Annu Rev Phytopathol 9:275-296

Fu D, Uauy C, Distelfeld A, Blechl A, Epstein L, Chen XM, Sela H, Fahima T, Dubcovsky J (2009) A novel kinase-START gene confers temperature-dependent resistance to wheat stripe rust. Science 323:1357-1360

Garrison E, Marth G (2012) Haplotype-based variant detection from short-read sequencing. *arXiv preprint arXiv:1207.3907 [q-bio.GN]*

Garrison E, Yandell M, Shapiro M, Marth G, Durbin R, Eichler EE, The VCFLIB Consortium, Kronenberg ZN (submitted) VCFLIB: and ensemble of methods for variant manipulation and population genetics.

Giese H, Holm-Jensen AG, Jensen HP, Jensen J (1993) Localization of the Laevigatum powdery mildew resistance gene to barley chromosome 2 by the use of RFLP markers. Theor Appl Genet 85:897-900

Giese H, Jørgensen JH, Jensen HP, Jensen J (1981) Linkage relationship of ten powdery mildew resistance genes on barley chromosome 5. Hereditas 95:43-50

Godfrey D, Böhlenius H, Pedersen C, Zhang Z, Emmersen J, Thordal-Christensen H (2010) Powdery mildew fungal effector candidates share N-terminal Y/F/WxC-motif. BMC Genomics 11:317

Halterman DA, Zhou FS, Wi FS, Wise RP, Schulze-Lefert P (2001) The MLA6 coiled-coil, NBS-LRR protein confers *AvrMla6*-dependent resistance specificity to *Blumeria graminis* f. sp. *hordei* in barley and wheat. Plant J 25:335-348

Halterman DA, Wei FS, Wise RP (2003) Powdery mildew-induced *Mla* mRNAs are alternatively spliced and contain multiple upstream open reading frames. Plant Physiol 131:558-567

Halterman DA, Wise RP (2004) A single-amino acid substitution in the sixth leucine-rich repeat of barley MLA6 and MLA13 alleviates dependence on RAR1 for disease resistance signaling. Plant J 38:215-226

Honecker L (1931) Beiträge zum Mehltauproblem bei der Gerste mit besonderer Berücksichtigung der züchterischen Seite. Pflanzenbau 8:78-84

Houterman PM, Cornelissen JC, Rep M (2008) Suppression of plant resistance gene-based immunity by a fungal effector. Plos Pathog 4:e1000061

Jensen HP, Christensen E, Jørgensen JH (1992) Powdery mildew resistance genes in 127 northwest European spring barley varieties. Plant Breeding 108:210-228

Jones JDG, Dangl JL (2006) The plant immune system. Nature 444:323-329

Jørgensen JH (1991) Powdery mildew resistance genes Mla22 and Mla23 in barley. In: Barley Genetics VI. Proc. 6th Int. Barley Genet Symp., Munksgaard, Copenhagen, pp 596-598

Jørgensen JH (1996) Effect of three suppressors on the expression of powdery mildew resistance genes in barley. Genome 39:492-498

Komínková E, Dreiseitl A, Malečková E, Doležel J, Valárik M (2016) Genetic diversity of *Blumeria graminis* f. sp. *hordei* in Central Europe and its comparison with Australian population. Plos One 11:e0167099

Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. Plant Methods 9:29

Krattinger SG, Lagudah ES, Spielmeyer W, Singh RP, Huerta-Espino J, McFadden H, Bossolini E, Selter LL, Keller B (2009) A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. Science 323:1360-1363

Kusch S, Ahmadinejab N, Panstruga R, Kuhn H (2014) *In silico* analysis of the core signaling proteome from the barley powdery mildew pathogen (*Blumeria graminis* f. sp. *hordei*). BMC Genomics 15:843

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357-359

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup (2009) The Sequence Alignment/Map (SAM) format and SAMtools. Bioinformatics 25:2078-2079.

Lipka AE, Tian F, Wang QS, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang ZW (2012) GAPIT: genome association and prediction integrated tool. Bioinformatics 28:2397-2399

Lo Presti L, Lanver D, Schweizer G, Tanaka S, Liang L, Tollot M, Zuccaro A, Reissmann S, Kahmann R (2015) Fungal effectors and plant susceptibility. Annu Rev Plant Biol 66:513-545

Lu XL, Kracher B, Isabel ML Saur, Bauer S, Ellwood SR, Wise R, Yaeno T, Maekawa T, Schulze-Lefert P (2016) Allelic barley MLA immune receptors recognize sequence-unrelated avirulence effectors of the powdery mildew pathogen. P Natl Acad Sci USA 113:E6486-E6495

Ma LS, Houterman PM, Gawehns F, Cao LX, Sillo F, Richter H, Clavijo-Ortiz MJ, Schmidt SM, Boeren S, Vervoot J, Cornelissen BJC, Rep M, Takken FLW (2015) The *AVR2-SIX5* gene pair is required to activate *I-2*-mediated immunity in tomato. New Phytol 208:507-518

Ma LS, Lukasik E, Gawehns F, Takken FLW (2012) The use of agroinfiltration for transient expression of plant resistance and fungal effector proteins in *Nicotiana benthamiana* leaves. In: Bolton MD, Thomma BPHJ (eds) Plant Fungal Pathogens: Methods and Protocols, Humana Press, NJ, USA, pp 61-74

McDonald BA, Linde C (2002) Pathogen population genetics, evolutionary potential and durable resistance. Annu Rev Phytopathol 40:349-379

McNally KE, Menardo F, Lühti L, Praz CR, Müller MC, Kunz L, Ben-David R, Chandrasekhar K, Dinoor A, Cowger C, Meyers E, Xue M, Zeng F, Gong S, Yu D, Bourras S, Keller B (2018) Distinct domains of the AVRPM3$^{A2/F2}$ avirulence protein from wheat powdery mildew are involved in immune receptor recognition and putative effector function. New Phytol DOI: 10.1111/nph.15026

Moore JW, Herrera-Foessel S, Lan CX, Schnippenkoetter W, Ayliffe M, Huerta-Espino J, Lillemo M, Viccars L, Milne R, Periyannan S, Kong XY, Spielmeyer W, Talbot M, Bariana H, Patrick JW, Dodds P, Singh R, Lagudah E (2015) A recently evolved hexose transporter variant confers resistance to multiple pathogens in wheat. Nat Genet 47:1494-1498

Moseman JG, Jørgensen JH (1971) Identification of genes at the *Mla* locus in barley for resistance to *Erysiphe graminis* f. sp. *hordei*. Crop Sci 11:547-550

Moseman JG, Macer RCF, Greeley LW (1965) Genetic studies with cultures of *Erysiphe graminis* f. sp. *hordei* virulent on *Hordeum spontaneum*. Trans Brit Mycol Soc 48:479-489

Moseman JG, Schaller CW (1960) Genetics of the allelic series at the *Mla* locus in barley and cultures of *Erysiphe graminis* f. sp. *hordei* that differentiate these alleles. Phytopathology 50:736-741

Parlange F, Roffler S, Menardo F, Ben-David R, Bourras S, McNally KE, Oberhaensli S, Stirnweis D, Buchmann G, Wicker T, Keller B (2015) Genetic and molecular characterization of a locus involved in avirulence of *Blumeria graminis* f. sp. *tritici* on wheat *Pm3* resistance alleles. Fungal Genet Biol 82:181-192

Pedersen C, van Themaat EVL, McGuffin LJ, Abbott JC, Burgis TA, Barton G, Bindschedler LV, Lu XL, Maekawa T, Wessling R, Cramer R, Thordal-Christensen H, Panstruga R, Spanu PD (2012) Structure and evolution of barley powdery mildew effector candidates. BMC Genomics 13:694

Petit-Houdenot Y, Fudal I (2017) Complex interactions between fungal avirulence genes and their corresponding plant resistance genes and consequences for disease resistance management. Front Plant Sci 8:1072

Plissonneau C, Blaise F, Ollivier B, Leflon M, Carpezat J, Rouxel T, Balesdent MH (2017) Unusual evolutionary mechanism to escape effector-triggered immunity in the fungal phytopathogen *Leptosphaeria maculans*. Mol Ecol 26:2183-2198

Plissonneau C, Daverdin G, Ollivier B, Blaise F, Degrave A, Fudal I, Rouxel T, Balesdent MH (2016) A game of hide and seek between avirulence genes *AvrLm4-7* and *AvrLm3* in *Leptosphaeria maculans*. New Phytol 209:1613-1624

Praz CR, Bourras S, Zeng FS, Sánchez-Martín J, Menardo F, Xue MF, Yang LJ, Roffler S, Böni R, Herren G, McNally KE, Ben-David R, Parlange F, Oberhaensli S, Flückiger S, Schäfer LK, Wicker T, Yu DZ, Keller B (2017) *AvrPm2* encodes an RNase-like avirulence effector which is conserved in two different specialized forms of wheat and rye powdery mildew fungus. New Phytol 213:1301-1314

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-959

Sánchez-Martín J, Steuernagel B, Ghosh S, Herren G, Hurni S, Adamski N, Vrána J, Kubaláková M, Krattinger S, Wicker T, Doležel J, Keller B (2016) Genome Biol 17:221

Sarris PF, Cevik V, Dagdas G, Jones JDG, Krasileva KV (2016) Comparative analysis of plant immune receptor architectures uncovers host proteins likely targeted by pathogens. BMC Biology 14:8

Scholz F, Nover I (1967) Genetische Untersuchungen mit einer vollständig mehltauresistenten Gerstenlinie. Kulturpflanze 15:243-254

Seeholzer S, Tsuchimatsu T, Jordan T, Bieri S, Pajonk S, Yang W, Jahoor A, Shimizu KK, Keller B, Schulze-Lefert P (2010) Diversity at the *Mla* powdery mildew resistance locus from cultivated barley reveals sites of positive selection. Mol Plant Microbe In 23:497-509

Segura V, Vilhjálmsson BJ, Platt A, Korte A, Seren Ü, Long Q, Nordborg M (2012) An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nat Genet 44: 825-830

Shen QH, Zhou FS, Brieri S, Haizel T, Shirasu K, Schulze-Lefert P (2003) Recognition specificity and RAR1/SGT1 dependence in barley *Mla* disease resistance genes to the powdery mildew fungus. Plant Cell 15:732-744

Silvar C, Kopahnke D, Flath K, Serfling A, Perovic D, Casas AM, Igartua E, Ordon F (2013) Resistance to powdery mildew in one Spanish barley landrace hardly resembles other previously identified wild barley resistances. Eur J Plant Pathol 136:459-468

Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, Stuber K, van Themaat EVL, Brown JKM, Butcher SA, Gurr SJ *et al.* (2010) Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. Science 330:1543-1546

Stanford EH, Briggs FN (1940) Two additional factors for resistance to mildew in barley. J Agric Res 61:231-236

Steuernagel B, Periyannan SK, Hernández-Pinzón I, Witek K, Rouse MN, Yu G, Hatta A, Ayliffe M, Bariana H, Jones JDG, Lagudah ES, Wulff BBH (2016) Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. Nat Biotechnol 34:652-655

Ramasamy RK, Ramasamy S, Bindroo BB, Naik VG (2014) STRUCTURE PLOT: a program for drawing elegant STRUCTURE bar plots in user friendly interface. SpringerPlus 3:431

Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP (2011) Integrative genomics viewer. Nat Biotechnol 29:24-26

Thorvaldsdottir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14:178-192

van der Biezen EA, Jones JDG (1998) Plant disease-resistance proteins and the gene-for-gene concept. Trends Biochem Sci 23:454-456

van der Hoorn RAL, Kamoun S (2008) From guard to decoy: a new model for perception of plant pathogen effectors. Plant Cell 20:2009-2017

Wiberg A (1974) Genetical studies of spontaneous sources of resistance to powdery mildew in barley. Hereditas 77:89-148

Wicker T, Oberhaensli S, Parlange F, Buchmann JP, Shatalina M, Roffler S, Ben-David R, Doležel J, Šimková H, Schulze-Lefert P, Spanu PD, Bruggmann R, Amselem J, Quesneville H, van Themaat EVL, Paape T, Shimizu KK, Keller B (2013) The wheat powdery mildew genome shows the unique evolution of an obligate biotroph. Nat Genet 45:1092-1096

Yahiaoui N, Srichumpa P, Dudler R, Keller B (2004) Genome analysis at different ploidy levels allows cloning of the powdery mildew resistance gene *Pm3b* from hexaploid wheat. Plant J 37:528-538

Zhang SL, Wang L, Wu WH, He LY, Yang XF, Pan QH (2015) Function and evolution of *Magnaporthe oryzae* avirulence gene *AvrPib* responding to the rice blast resistance gene *Pib*. Sci Rep-UK 5:11642

Zhou FS, Kurth J, Wei FS, Elliot C, Valè G, Yahiaoui N, Keller B, Somerville S, Wise R, Schulze-Lefert P (2001) Cell-autonomous expression of barley *Mla1* confers race-specific resistance to the powdery mildew fungus via a *Rar1*-independent signaling pathway. Plant Cell 13:337-350

APPENDIX VI

13th International Wheat Genetics Symposium. Tulln, Austria; April 23-28, 2017

Positional cloning of powdery mildew resistance gene introgressed to bread wheat from *Triticum militinae*

Eva Komínková[1], Irena Jakobson[2], Hilma Peusha[2], Michael Abrouk[1], Kadri Järve[2], Jaroslav Doležel[1], Miroslav Valárik[1]

[1]Institute of Experimental Botany of the Czech Academy of Sciences, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 31, Olomouc, CZ-78371, Czech Republic

[2]Department of Chemistry and Biotechnology, Tallinn University of Technology, Akadeemia tee 15, 19086 Tallinn, Estonia

Abstract

Introgressions from related species have frequently been employed in crop improvement programs. Powdery mildew resistance locus *QPm.tut-4A* originating from tetraploid *Triticum militinae* significantly improves seedling and adult plant resistance of hexaploid bread wheat (*T. aestivum*). The resistance locus was mapped to the distal end of 4AL chromosome arm using a mapping population comprising 7500 lines which was created from a cross between resistant introgression line 8.1 and susceptible cv. Chinese Spring and 2053 lines of a mapping population derived from a cross between line with introgression and cv. Chinese Spring carrying a recessive *Ph1* locus. The gene region was identified by chromosome walking employing 4AL chromosome arm-specific BAC libraries of cv. Chinese Spring and the introgression line 8.1. The *QPm.tut-4A* region (650 kbp) shows low sequence conservation between the parental lines as a consequence of insertions of different transposable elements, which resulted in suppression of recombination. The locus comprises about 80 putative protein coding genes in cv. Chinese Spring. To reduce the number of potential candidates, a TILLING population of 975 lines was established and six mutants susceptible to powdery mildew were selected. Chromosome 4A will be flow-sorted from these lines, sequenced and the data obtained

will be used to identify a candidate resistance gene. This work has been supported by grant award LO1204 from the National Program of Sustainability I, by the Czech Science Foundation (award 14-07164S) and by institutional research funding IUT 193 of the Estonian Ministry of Education and Research.

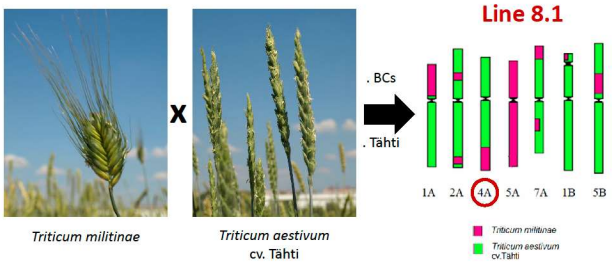
Positional cloning of powdery mildew resistance gene introgressed to bread wheat from *Triticum militinae*

Eva Komínková[1], Irena Jakobson[2], Hilma Peusha[2], Michael Abrouk[1], Kadri Järve[2], Jaroslav Doležel[1], Miroslav Valárik[1]

[1]Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Olomouc, Czech Republic
[2]Department of Chemistry and Biotechnology, Tallinn University of Technology, Tallinn, Estonia

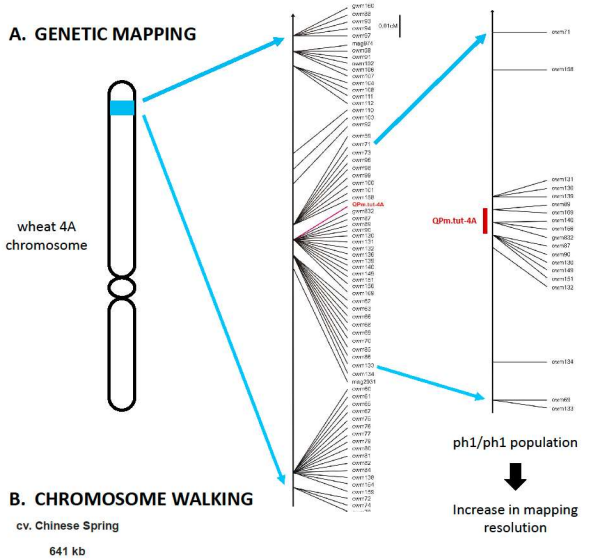
INTRODUCTION

Introgressions from related species have frequently been employed in crop improvement programs. Powdery mildew resistance locus *QPm.tut-4A* originating from tetraploid *Triticum militinae* significantly improves race non-specific seedling and adult plant resistance of hexaploid bread wheat (*T. aestivum*). The major resistance QTL was mapped to the distal end of 4AL chromosome arm and was subjected to map-based cloning approach to identify the responsible gene.



Triticum militinae — *Triticum aestivum* cv. Tähti

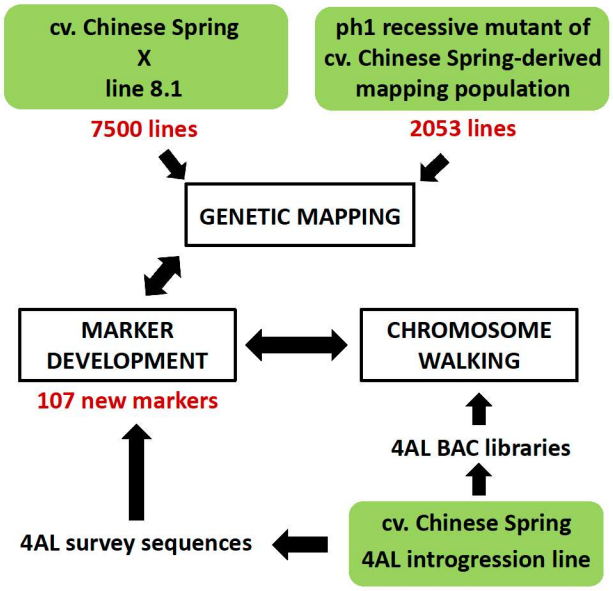
RESULTS



Physical map of the *QPm.tut-4A* locus

- 480 kb
- Significant sequence divergence of corresponding region in *T. aestivum* cv. Chinese Spring
- Different composition of transposable elements fraction
- Suppression of recombination
- 12 predicted protein-coding genes
- A cluster of 6 putative genes showing homology with resistance gene domains

MATERIAL & METHODS



CONCLUSIONS

This study reports a significant progress in an effort to clone a *QPm.tut-A4* gene improving race non-specific resistance of hexaploid bread wheat to powdery mildew. With a physical map and candidate genes available, the next steps will be focused on identification of the causal gene and its validation. To reduce the number of potential candidates, a TILLING population of 975 lines was established and seven mutants susceptible to powdery mildew were selected. Chromosome 4A will be flow-sorted from these lines, sequenced and the data obtained will be used to identify a candidate resistance gene.

In the meantime, a possibility to employ *QPm.tut-4A* in breeding is being assessed by transferring the minimal introgression into the genomes of different bread wheat cultivars.

REFERENCES

Jakobson I, Reis D, Tiidema A, Peusha H, Timofejeva L, Valárik M, Kladivová M, Šimková H, Doležel J, Järve K (2012): Fine mapping, phenotypic characterization and validation of non-race-specific resistance to powdery mildew in a wheat-*Triticum militinae* introgression line. Theor Appl Genet 125: 609-623

ACKNOWLEDGEMENTS

This work has been supported by grant award LO1204 from the National Program of Sustainability I, by the Czech Science Foundation (award 14-07164S) and by institutional research funding IUT 193 of the Estonian Ministry of Education and Research.

ÚEB

TALLINN UNIVERSITY OF TECHNOLOGY

APPENDIX VII

14th International Cereal Rusts and Powdery Mildews Conference. Helsingør, Denmark; July 5-8, 2015

Molecular description of *Blumeria graminis* f. sp. *hordei* isolates

Eva Komínková[1], Eva Malečková[1], Hana Vanžurová[1], Antonín Dreiseitl[2], Jaroslav Doležel[1], Miroslav Valárik[1]

[1]Institute of Experimental Botany of the Czech Academy of Sciences, Centre of the Region Haná for Biotechnological and Agricultural Research, Olomouc, Czech Republic

[2]Agrotest Fyto Ltd., Kroměříž, Czech Republic

Abstract

The air-born fungus *Blumeria graminis* f. sp. *hordei* (*Bgh*) is a causal agent of barley powdery mildew. The pathogen attracts substantial attention due to its destructiveness. However, molecular diversity studies based on "house-keeping" genes do not provide sufficient resolution when applied to isolates from geographically limited regions. This study focused on developing a more efficient genotyping system capable to discriminate between closely related isolates. Whole genome sequence data were employed to design a panel of molecular markers based on microsatellites and insertion sites of transposable elements which represent an abundant part of the genome. A genotyping marker panel comprising 16 SSR, 14 SNP and 2 ISBP/RJM markers was applied on a set of 97 isolates originating from the Czech Republic, 50 Australian isolates and a collection of 11 isolates representing global *Bgh* diversity. The marker panel provided significant resolution of studied isolates, most of them showing unique genotype profiles. The analysis of phylogenetic relationship performed by neighbor-joining algorithm for 97 Czech isolates resulted in 87 separate clades and revealed high diversity of the pathogen population within a small geographical area. After supplementing with data on virulence of individual isolates, this study might open new opportunities of studying the host-pathogen relationship and patterns of the pathogen spatial distribution. This work has been supported by the Czech Ministry of Education, Youth and Sports (grant awards LD14105, LO1204).

Molecular description of *Blumeria graminis* f. sp. *hordei* isolates

Eva Komínková[1], Eva Malečková[1], Hana Vanžurová[1], Antonín Dreiseitl[2], Jaroslav Doležel[1] & Miroslav Valárik[1]

[1]Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 31, CZ-783 71 Olomouc, Czech Republic
[2]Agrotest Fyto Ltd., Havlíčkova 2787, CZ-767 01 Kroměříž, Czech Republic

INTRODUCTION

The air-born fungus *Blumeria graminis* f. sp. *hordei* (*Bgh*) is a causal agent of barley powdery mildew. The pathogen attracts substantial attention due to its destructiveness (Fig. 1). However, molecular diversity studies based on "house-keeping" genes do not provide sufficient resolution when applied to isolates from geographically limited regions. This study focuses on developing a more efficient genotyping system capable to discriminate between closely related isolates. Furthermore, it demonstrates application of new markers on a set of isolates including Czech Bgh population together with selected isolates originating from different parts of the world.

Figure 1: *Blumeria graminis* DC., currently considered to be the 6th most important fungal plant pathogen (Dean *et al.*, 2012).
A) Colonies of fungus visible on leaf surface. B) Conidiophores. C) Conidia.

MATERIAL & METHODS

Whole genome sequence data of *Bgh* strain DH14 (Spanu *et al.*, 2010)

Searching microsatellites using WebSat tool (Martins *et al.*, 2009)

Searching transposable elements junctions: BLAST with RepBase (Jurka *et* Kohany, 2011)

Designing primers, testing polymorphism on 14 *Bgh* isolates, sequencing

Set of new markers – SSR, ISBP/RJM, SNP

DNA extraction from conidia, genotyping, phylogenetic analysis by neighbor-joining algorithm using software package PHYLIP 3.69 (Felsenstein, 2005)

CONCLUSIONS

- Whole genome sequence data of *Blumeria graminis* f. sp. *hordei* were employed to design a panel of molecular markers based on microsatellites and insertion sites of transposable elements. Altogether, 32 polymorphic markers of three different types (SSR, SNP, ISBP/RJM) were developed and used for further analysis.
- Genotyping of 158 Bgh isolates revealed high genetic variability allowing unambiguous identification in most of the cases. Resulting genotype profiles were used for phylogenesis inference analysis to demonstrate one of possible applications of these data.
- After supplementing with data on virulence of individual isolates, this study might open new opportunities of studying the host-pathogen relationship and patterns of the pathogen spatial and time distribution.

RESULTS

Based on the approach described above, a genotyping marker panel comprising 16 SSR, 14 SNP and 2 ISBP/RJM markers was developed (Fig. 2).

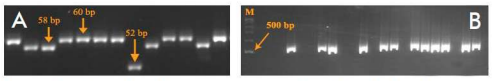


Figure 2: Examples of new markers amplified from *Bgh* isolates.
A) SSR marker *obm28*. **B)** ISBP/RJM marker *obm14*. PCR products were visualized by ethidium bromide staining after electrophoretic separation on 6% and 4% polyacrylamid gel, respectively.

The final marker panel comprising 32 polymorphic markers provided significant resolution of 158 studied isolates, most of them showed unique genotype profiles. The analysis of phylogenetic relationship performed by neighbor-joining algorithm for 97 Czech isolates resulted in 87 separate clades and revealed high diversity of the pathogen population within a small geographical area (Fig. 3, 4).

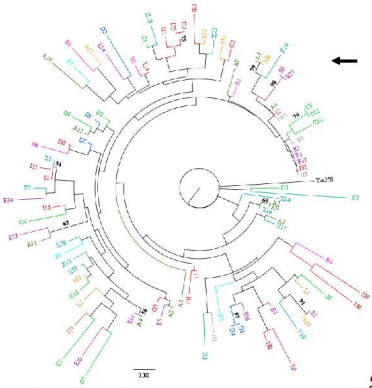


Figure 3: Tree phylogram obtained from neighbor-joining inference analysis performed for 97 Czech *Bgh* isolates.
Bootstrap values exceeding 50 % are indicated above branches. Groups of isolates are separated by different colours corresponding to the region of origin within the Czech republic (Fig. 4).

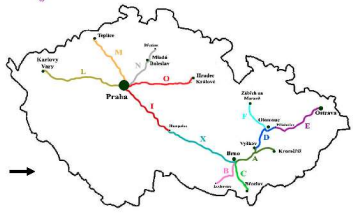


Figure 4: Distribution of studied isolates within the Czech republic.
Colours of individual sections correspond to colours of isolates on Fig. 3. Letters (A, B, C, D, E, F, I, L, M, N, O, X) represent groups of isolates collected in respective region.

REFERENCES

Dean, R., van Kan, J. A. L., Pretorius, Z. A., Hammond-Kosack, K. E., di Pietro, A., Spanu, P. D., Rudd, J. J., Dickman, M., Kahmann, R., Ellis, J., Foster, G. D. (2012): The Top 10 fungal pathogens in molecular plant pathology. Molecular Plant Pathology 13 (4): 414 – 430.

Felsenstein, J. (2005): PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Jurka, J., Kohany, O. (2011): LTR retrotransposons from barley powdery mildew. Repbase Reports 11 (9), 2 333 – 2 333. Accesible online at http://www.girinst.org/.

Martins, W. S., Lucas, D. C. S., Neves, K. F. S., Bertioli, D. J. (2009): WebSat – a web software for microsatellite marker developement. Bioinformation 3 (6): 282 – 283. Accesible online at http://wsmartins.net/websat/.

Spanu, P. D., Abbott, J. C, Amselem J., Burgis, T. A., Soanes, D. M. et al. (2010): Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. Science 330 (6010): 1 543 – 1 546.

ACKNOWLEDGEMENTS

This work has been supported by the Czech Ministry of Education, Youth and Sports (grant awards LD14105, LO1204).

ÚEB C. R. HANÁ

APPENDIX VIII

Plant Biotechnology: Green for Good III. Olomouc, Czech Republic; June, 15-18, 2015

Positional cloning of powdery mildew resistance gene introgressed to bread wheat from *Triticum militinae*

Eva Komínková[1], Barbora Klocová[1], Michael Abrouk[1], Irena Jakobson[2], Hilma Peusha[2], Hana Šimková[1], Jan Šafář[1], Kadri Järve[2], Jaroslav Doležel[1], Miroslav Valárik[1]

[1]Institute of Experimental Botany of the Czech Academy of Sciences, Centre of the Region Haná for Biotechnological and Agricultural Research, Olomouc, Czech Republic

[2]Department of Gene Technology, Tallinn University of Technology, Tallinn, Estonia

Abstract

Introgressions from wild relatives are a valuable source of new genes for crop improvement. Following this approach, a gene pool of bread wheat cv. Tähti was enriched by segments of *T. militinae* genome. The resulting introgressive line 8.1 showed improved race non-specific resistance to powdery mildew. The major resistance QTL called *QPm.tut-4A* was identified on the distal end of the long arm of chromosome 4A. The gene locus was fine mapped to 0.013 cM region using 51 new markers and 7500 individuals from Chinese Spring x line 8.1 mapping population. However, a physical map of the region still comprises three unconnected BAC contigs. To increase the mapping resolution, a mapping population homozygous for ph1 gene was established. The recombination frequency in this population increased by 27-fold and four new recombinants in the region were obtained using only 1256 individuals. In addition, over 15 Mbp of DNA sequences were acquired by sequencing MTP BAC clones spanning the locus and flanking regions. These advances offer a solid base for cloning the resistance gene. This work has been supported by grant LO1204 from the National Program of Sustainability I, by the Czech Science Foundation (14-07164S) and by Estonian Ministry of Agriculture.

Positional cloning of powdery mildew resistance gene introgressed to bread wheat from *Triticum militinae*

Eva Komínková[1], Barbora Klocová[1], Michael Abrouk[1], Irena Jakobson[2], Hilma Peuscha[2], Hana Šimková[1], Jan Šafář[1], Kadri Järve[2], Jaroslav Doležel[1], Miroslav Valárik[1]

[1]Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Olomouc, Czech Republic
[2]Department of Gene Technology, Tallinn University of Technology, Tallinn, Estonia

INTRODUCTION

Since the Neolitic revolution leading to development of agriculture 12,000 years ago, cereals has played a key role in human nutrition. Wheat, the most widely grown cereal grain occupying 17 % of the total cultivated lands in the world, provides staple food for 35 % of the world's population. With the human population estimated to reach 9.6 billion by 2050, wheat production will have to increase by 60 % in the next four decades. Meeting the growing demand represents a major challenge for breeders and scientific community worldwide. Introgressions from wild relatives represent a valuable source of new genes for crop improvement. Following this approach, a gene pool of hexaploid bread wheat *Triticum aestivum* cv. Tähti was enriched by segments of tetraploid *Triticum militinae* genome. The resulting introgressive line 8.1 showed improved race non-specific resistance to the fungal pathogen *Blumeria graminis* f. sp. *tritici*, a causal agent of powdery mildew disease of wheat. The major resistance QTL called *QPm.tut-4A* was identified on the distal end of the long arm of chromosome 4A. The goal of this study was to delimit the locus by employing two different mapping populations and developing new molecular markers in the region of interest. This progress will facilitate physical mapping of the locus and cloning the gene.

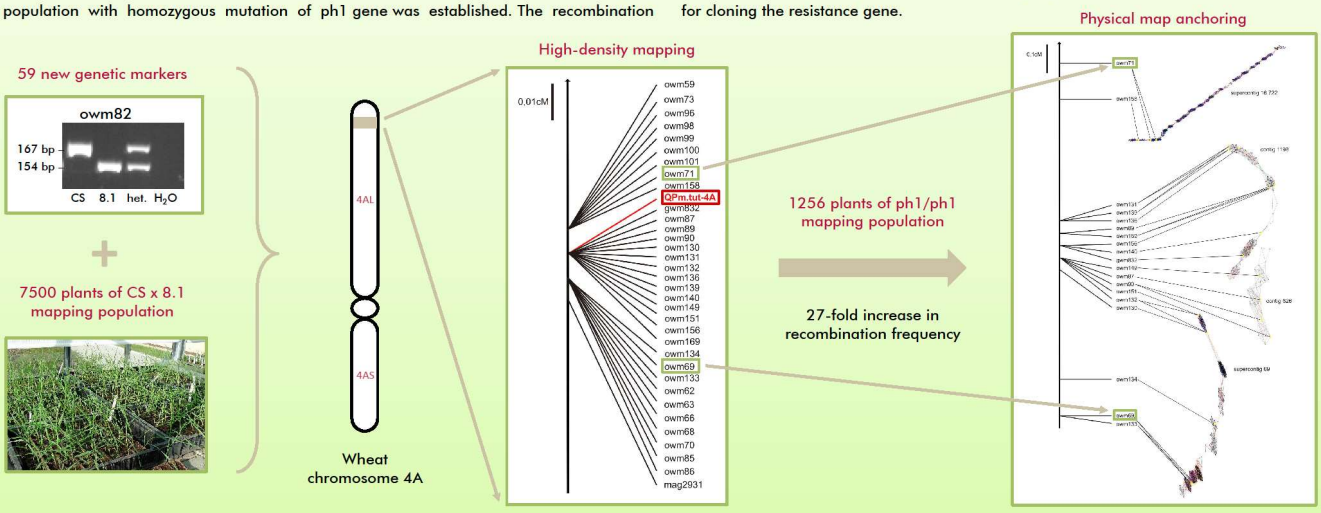
Triticum aestivum cv. Tähti (AABBDD)
SUSCEPTIBLE TO POWDERY MILDEW

Triticum militinae (AAGG)
RESISTANT TO POWDERY MILDEW

Triticum aestivum line 8.1 (AABBDD*)
IMPROVED SEEDLING AND ADULT PLANT RESISTANCE

RESULTS

The gene locus was fine mapped to 0.013 cM region using 59 new genetic markers (CAPS, markers based on length polymorphism) and 7500 individuals of *T. aestivum* cv. Chinese Spring (CS) x line 8.1 mapping population. However, a physical map of the region still comprises several unconnected BAC contigs. To increase the mapping resolution, a mapping population with homozygous mutation of ph1 gene was established. The recombination frequency in this population increased by 27-fold and four new recombinants in the region were obtained using only 1256 individuals. Phenotyping of these critical recombinants is under progress. In addition, over 15 Mb of DNA sequences were acquired by sequencing MTP BAC clones spanning the locus and flanking regions. These advances offer a solid base for cloning the resistance gene.



CONCLUSIONS

- Mapping population carrying mutated ph1 alleles provides a highly efficient solution for fine mapping of regions with insufficient incidence of natural recombinations. This approach resulted in significant increase of mapping resolution while number of plants used for mapping was several times lower compared to the original Ph1/Ph1 mapping population.
- Due to recent progress, cloning the QPm.tut-4A gene gets within reach.

ACKNOWLEDGEMENTS

This work has been supported by grant LO1204 from the National Program of Sustainability I, by the Czech Science Foundation (14-07164S) and by Estonian Ministry of Agriculture.