Advances in wheat and pathogen genomics - implications for disease control

Beat Keller^{1*}, Thomas Wicker¹, Simon G. Krattinger^{2*}

¹Department of Plant and Microbial Biology, University of Zurich, Zollikerstrasse 107, 8008 Zurich, Switzerland

²King Abdullah University of Science and Technology (KAUST), Biological and Environmental Sciences & Engineering Division (BESE), Thuwal, 23955-6900, Saudi Arabia

Corresponding authors: bkeller@botinst.uzh.ch; simon.krattinger@kaust.edu.sa

Keywords: genomics-assisted breeding, genome sequence, gene cloning, pathogenomics, disease resistance gene, pathogenic fungi

ABSTRACT

The gene pool of wheat and its wild and domesticated relatives contains a plethora of resistance genes that can be exploited to make wheat more resilient to pathogens. Only few of these genes have been isolated and studied at the molecular level. In recent years, we have seen a shift from “classical” breeding to “genomics-assisted breeding”, which makes use of the enormous advancements in DNA sequencing and high-throughput molecular marker technologies for wheat improvement. These genomic advancements have the potential to transform wheat breeding in the near future and to significantly increase the speed and precision at which new cultivars can be bred. This review highlights the genomic improvements that were made in wheat and its pathogens over the past years and discusses their implications for disease resistance breeding.

INTRODUCTION

The cereal crop plants including maize, rice, wheat, barley, sorghum, millets, oats and rye play a pivotal role in providing food for humankind. Tetraploid pasta wheat (*Triticum durum*) and hexaploid bread wheat (*Triticum aestivum*) are among the most widely cultivated cereal plants. In 2016, the global wheat production reached a new all-time record of 760 million tons (FaoStat). Despite this impressive number, the production of wheat and other crops is constantly at risk and threatened by biotic and abiotic stresses that impair crop yield and quality. Ever since the dawn of agriculture, crop diseases, pests, droughts and other stresses have affected crop production, with sometimes dramatic impacts on human history. Plant diseases caused by pathogenic viruses, bacteria, oomycetes and fungi have been a major cause of crop losses. In the last years, the emergence of new and highly virulent races of the wheat stem rust fungus, collectively referred to as Ug99 after their first occurrence in Uganda in the year 1999, caused serious damage to wheat production in Africa and the Middle East (81). In 2016, wheat blast, a disease previously restricted to Brazil and neighboring South American countries, ravaged wheat fields in Bangladesh (44). These examples highlight the impact of plant diseases on food security and human health. Not surprisingly, one of the main objectives in breeding and agriculture has consisted in the selection and cultivation of crops that are able to withstand disease. The basis of disease resistance however was not comprehensively understood until the description of the laws of inheritance by Gregor Mendel in the mid-19th century, which laid the foundation for modern genetics and plant breeding as we know it nowadays.

Latest since the Green Revolution in the 1960s (73), the genetics of disease resistance in crops has been systematically studied and dissected into its single genetic components, which are the disease resistance genes encoded in a plant's genome. In wheat, this systematic effort has led to the cataloging of several hundred different disease resistance genes (61). However, only a small fraction of these genes has been cloned and their molecular function studied in detail.

The wheat genome is among the largest and most complex of all cultivated plants. With 17 billion nucleotides (Gb), the bread wheat genome is five times larger than the human genome. Bread wheat is a hexaploid species that arose through natural hybridization of three closely related wild grass species (76). In addition, the proportion of repetitive sequences within the wheat genome is higher than 80%. The huge and complex genome of wheat has greatly complicated the isolation of

individual genes in the past. The first disease resistance genes cloned in wheat were *Lr10* and *Lr21* that both confer race-specific resistance against the fungal leaf rust disease (*Puccinia triticina*). Both genes encode intracellular immune receptors of the nucleotide binding-site – leucine rich repeat (NLR) family (26; 34). The cloning of *Lr10* and *Lr21* occurred in 2003, almost eight years after the first disease resistance gene was isolated from rice (82). In comparison to the wheat genome, the rice genome with its 380 million nucleotides is considerably smaller (40) which makes gene cloning in rice more feasible. The difficult nature of the wheat genome however has inspired wheat researchers to think of novel genomic approaches that facilitate genomics-assisted breeding, including gene cloning. These efforts have resulted in the development of extensive genomic resources, including high-quality genome sequences, high-throughput genotyping platforms and several rapid gene cloning strategies. The results of these achievements have already become apparent when looking at the increasing rate at which disease resistance genes are isolated in wheat (Figure 1). Here, we review these recent advances in genomics of wheat and its pathogens and we will discuss implications for wheat disease resistance breeding.

ADVANCES IN WHEAT GENOMICS – IMPLICATIONS FOR THE GENETIC AND MOLECULAR UNDERSTANDING OF DISEASE RESISTANCE

Gene cloning strategies in the ‘pre-reference genome era’

The ‘cloning’ or ‘isolation’ of a gene describes the process of pinpointing it to its exact location within the genome, which will ultimately reveal its nucleotide sequence and the protein it encodes for. The cloning of disease resistance genes forms the basis for many important breeding objectives: (i) it allows to design perfect gene-based molecular markers for marker-assisted selection, (ii) it enables the rapid transfer of cloned genes through transgenesis or their modification through genome editing, (iii) it enables diagnostic applications, e.g. to determine which gene is present in which wheat cultivar, and (iv) it forms the basis to unravel the molecular function of plant-pathogen interactions, knowledge that will ultimately help to protect wheat from pathogens. However, gene cloning from the vast wheat genome has presented a great challenge in the past. The majority of all wheat disease resistance genes cloned so far were isolated through ‘map-based cloning’ (Figure 1), an approach that aims to stepwise narrow down the location of a gene by making use of genetic recombination in bi-parental populations segregating for the gene

of interest. A critical unit in this process is the size of DNA fragments that are easily accessible to sequencing. Since it was not possible until recently to sequence whole wheat genomes, wheat DNA was fragmented and tiny segments of the wheat genome were randomly cloned into bacterial artificial chromosomes (BAC). Each BAC clone can accommodate a DNA piece of 100-200 kb in size, which corresponds to as little as 0.001% of a hexaploid wheat genome. Hence, around 500,000 BAC clones are necessary to obtain a representative library of a complete wheat genome (56). In order to delimit a gene of interest to an attainable piece of DNA, the gene thus needed to be mapped to a very small genetic interval (<0.5 cM) that corresponded to one or a few overlapping BAC clones. Large mapping populations consisting of several thousand plants were required to achieve the necessary genetic resolutions. For example, the *Lr10* leaf rust resistance gene was fine-mapped to a 0.13 cM genetic interval by using a mapping population consisting of 3,120 F₂ plants (83), which allowed to narrow down *Lr10*'s location to a physical interval of 450 kb that was spanned by four overlapping BAC clones. While map-based cloning works well in telomeric regions with high-recombination rates, cloning genes from centromeric regions that often span megabases devoid of recombination, is nearly impossible using map-based cloning. For similar reasons, the cloning of genes from regions that were introgressed into wheat from wild grasses is very challenging. Below we will highlight and discuss recent key advances in genomics (6) that had and will have significant impact on the speed, costs and precision at which wheat genes can be isolated in the future.

Sequencing the wheat genome – the rocky road towards a high quality reference sequence

The wheat genome has long been considered to be too large and too complex to be sequenced at high quality. First whole genome shotgun assemblies of the hexaploid bread wheat landrace Chinese Spring and the wild diploid wheat relatives einkorn (*Triticum urartu*) and goatgrass (*Aegilops tauschii*) were generated using short-read Roche 454 and Illumina sequencing (10; 45; 53). While these assemblies provided a first snapshot of the wheat genome and its gene space in particular, they were still highly fragmented and consisted of ten thousands of unordered scaffolds. The scaffold N50 lengths of these 'early' assemblies was in the range of a few kb. In 2014, the International Wheat Genome Sequencing Consortium (IWGSC) adopted a novel strategy by generating a whole draft genome sequence of Chinese Spring from single isolated chromosome arms (41). This approach was based on the ability to isolate single wheat chromosome arms from

aneuploid Chinese Spring lines by flow cytometry (21), which allowed to unambiguously assign sequence scaffolds to individual chromosome arms. The order of the genic sequences was inferred through a method named ‘genome zipper’, which makes use of the fact that gene order is highly conserved among different grass species (60). Thus, the gene order from grass species with completed high-quality genome sequences such as rice (40), sorghum (66) and *Brachypodium distachyon* (89) was used to infer the gene order in wheat. However, this approach did not allow to capture large structural rearrangements between wheat and other grasses and is also prone to errors in regions where gene order is poorly conserved or which contain large chromosomal translocations. In the same year as the Chinese Spring draft genome sequence, a high-quality genome assembly of the 1-gigabase Chinese Spring chromosome 3B was reported (18). For this assembly, a physical map was initially produced by painstakingly fingerprinting and ordering 67,968 BAC clones (67; 74). The resulting sequence assembly had a scaffold N50 of 892 kb, which provided a much improved contiguousness compared to the whole genome shotgun assemblies that were published around the same time (41).

The end of the rocky road - chromosome-scale assemblies of wheat genomes

A quantum leap towards the improvement of cereal genomes was recently achieved through several technological and conceptual advancements like integrating whole-genome shotgun libraries of various insert-sizes (32), the use of long-read sequencing technologies (46), the use of chromosome contact maps (Hi-C) (59) and the development of better assembly algorithms (19). The first high-quality genome sequence of a wheat relative was generated from the wild emmer (*Triticum turgidum* ssp. *dicoccoides*) accession ‘Zavitan’ (4). The sequencing of several whole genome shotgun libraries resulted in a 10.1-gigabase assembly with a scaffold N50 of nearly 7 Mb (i.e. 50% of the assembly is represented in sequence scaffolds of 7 Mb), which is 50-times longer than a large BAC clone and represented an improvement by a factor of hundred to thousand over the previous whole genome assemblies. Integration of additional Hi-C chromosome conformation capture sequencing allowed to assemble these scaffolds into 14 pseudomolecules matching the 14 chromosomes of tetraploid wild emmer.

Similar strategies recently allowed to generate a high-quality genome assembly of the hexaploid wheat landrace Chinese Spring (42). The 21 resulting pseudomolecules have a cumulative length of 14.1 Gb and thus represent near-complete chromosomes. The total size of the assembly is close

to previous estimates which were based on cytometric measurements and which ranged from 15.9 Gb (3) to 17.3 Gb (5). The difference to previous estimates is likely due to the presence of highly repetitive sequences (e.g. ribosomal DNA clusters) which were collapsed during the assembly and possibly also due to an overestimate in the original numbers (42). The quality of the assembly surpasses that of previous ones with an N90 of 4.1 Mb (i.e. 90% of the assembly is represented in scaffold of at least 4.1 Mb). This near-complete assembly allowed for the first time a genome-wide assessment of genic and repetitive regions. Gene content of the three sub-genomes was found to be very similar with each containing approximately 34,000 to 35,000 genes. Furthermore, gene order and spacing is well conserved between sub-genomes. Nevertheless, about 15% of “homeologous triads” (i.e. groups of orthologs across the A, B and D sub-genomes) showed duplications of genes in at least one sub-genome, indicating that many gene loci are prone to copy number variation.

In the close future, a high-quality assembly of the tetraploid durum wheat cultivar Svevo will become available (International Durum Wheat Sequencing Consortium, unpublished). Together with genome sequences of diploid wheat progenitors such as the one of *Ae. tauschii* (55), future comparative analyses can be performed across genomes of three ploidy states. Furthermore, an effort is underway to sequence a total of ten wheat genotypes at the same level of quality as the IWGSC reference sequence for Chinese Spring (<http://www.wheatinitiative.org/activities/associated-programmes/10-wheat-genomes-project>). Comparative analysis across these ten genotypes will provide insight into differences in gene content, large-scale chromosomal rearrangements and haplotype diversity.

Whole genome sequences assist identification and cloning of disease resistance genes

The availability of near-complete chromosome sequences marked a milestone for gene cloning because it became possible to ‘navigate’ within wheat genomes and to get insight into large contiguous genomic regions. This will make map-based cloning much simpler and faster in the future. Furthermore, the best markers identified in genome-wide association studies (GWAS) can be rapidly anchored to large contigs, allowing to define candidate genes for the trait of interest. Finally, these assemblies allow to rapidly identify potential disease resistance genes across entire genomes based on homology with already known disease resistance genes. For example, many disease resistance genes encode NLRs (20). Based on this knowledge, Steuernagel et al. (2018)

developed an *in silico* NLR-Annotator algorithm that allows to detect NLR genes based on the presence of conserved motifs that are typically found in NLRs. A systematic evaluation of the Chinese Spring reference sequence (42) revealed the presence of 3,400 NLR loci in the Chinese Spring genome, serving as potential candidates for disease resistance genes (Figure 2). NLR density was higher towards the telomeres and rapidly decreased towards the centromeric regions. Interestingly, a literature survey with known and mapped race-specific leaf and stem rust resistance loci revealed that many of them coincided with regions that contained only a few NLR genes in Chinese Spring (Steuernagel et al. 2018; Figure 2). These results show the enormous value and potential of high-quality wheat genome sequences for gene cloning and a rapid increase of the number of cloned disease resistance genes can be expected in the near future. Furthermore, it has been found that orthologous genes can have a conserved function in disease resistance across different species. An example is the mildew A (*Mla*) locus that was originally identified in barley where it confers resistance against the fungal barley powdery mildew disease (*Blumeria graminis* f.sp. *hordei*). The *Mla* locus consists of multiple copies of paralogous NLR genes (94), but the allelic series of the active *Mla* resistance gene in different genotypes is derived from only one of them (79). Subsequent to the identification of *Mla* in barley, orthologous wheat and rye genes of *Mla* were identified that confer resistance against wheat powdery mildew (*B. graminis* f.sp. *tritici*) and stem rust (*P. graminis* f.sp. *tritici*) (48; 58; 68). The availability of high-quality wheat reference genomes will greatly simplify the identification of orthologs of known resistance genes from other cereals and to functionally test them. The identification of resistance gene orthologs is of great relevance for practical breeding as was shown by studies on a specific wheat suppressor of the rye *Pm8* powdery mildew resistance gene (36; 37). In this work, the wheat ortholog *Pm3CS* (a susceptible allele of the *Pm3* allelic series of resistance genes) of *Pm8* was shown to suppress *Pm8*. Resistance gene suppression has been observed frequently in wheat resistance breeding and homoeologous or orthologous suppressor genes might be quite common. Such putative suppressor genes can now be identified and functionally tested in the transient *Nicotiana benthamiana* system (37) and they could be eliminated by gene editing.

Rapid gene cloning approaches from non-reference wheat cultivars

Despite the progress in generating high-quality reference genomes, one particular problem remains; the gene of interest might only occur in a very specific wheat cultivar or landrace from a

genebank collection and might be absent from the reference genotypes. It is known that sequence variation between different wheat cultivars can be extensive, including large structural rearrangements (47; 57). In particular, regions harboring disease resistance genes have been found to be rapidly evolving and diverse (43). Thus, while high-quality reference sequences of particular genotypes are invaluable resources to define and characterize a region of interest, it might not be possible to identify the actual coding sequence of a specific gene from the reference cultivars. One exception is the durable multi-pathogen resistance gene *Lr34*, which is present in Chinese Spring and codes for an ABC transporter protein (52). The *Lr34* gene was cloned in 2009 by mapping the gene to a 0.15 cM genetic interval. A considerable amount of time was spent in assembling the corresponding 370 kb physical interval from four overlapping BAC clones. With the availability of the Chinese Spring reference sequence, the same task could now be completed within a few minutes instead of several months. Chinese Spring also contains the *Zymoseptoria tritici* resistance gene *Stb6* that was recently found to encode a wall-associated kinase (WAK)-like receptor-like kinase (75). To establish the physical interval of the *Stb6* target region the chromosome survey sequences of Chinese Spring were used, which allowed to directly identify candidate genes (75). However, these two examples are exceptions and most disease resistance genes are not present in Chinese Spring or the wild ancestors that have been sequenced at high quality.

As described above, map-based cloning approaches in wheat have been slow and tedious in the past. Several recent strategies have been developed to overcome the limitations associated with classical map-based cloning (91). In addition, these approaches aimed to facilitate gene cloning directly from the cultivar of interest in order to eliminate the problem of genotype-to-genotype variation. One of these approaches, MutRenSeq, makes use of the fact that disease resistance genes often encode NLR proteins. Based on this knowledge, an exome capture approach was developed that allowed to enrich for NLR-specific wheat sequences (84). The enrichment library was based on 60,000 NLR baits that were designed on publicly available wheat and barley genome sequences. As a proof-of-concept, MutRenSeq was applied to rapidly clone the two wheat stem rust resistance genes *Sr22* and *Sr45* (84). To isolate *Sr22*, captured NLR short-read sequence contigs of six EMS-derived *Sr22* loss-of-function mutants in the genetic background of the *Sr22* carrying Australian wheat cultivar Schomburgk were compared to the respective wild-type contigs, which revealed a single contig with independent mutations in five of the six loss-of-function mutants. A similar strategy was used to isolate *Sr45* from the wheat line CS1D5406. The great advantages of

MutRenSeq are that it does not require fine-mapping and that it can be applied to any wheat cultivar. However, as the molecular identification of *Stb6* gene showed (75), not all (race-specific) disease resistance genes encode NLR proteins and this has to be considered when applying MutRenSeq.

Other approaches utilized genome complexity reduction through the isolation of individual wheat chromosomes. The race-specific wheat powdery mildew resistance gene *Pm2* for example was cloned by isolating and sequencing chromosome 5D of a *Pm2* carrying wheat line and of six independent EMS-derived *Pm2* loss-of-function mutants using flow cytometry and short-read sequencing. Comparison of sequence contigs of the six mutants to the *Pm2* donor line allowed to delimit the causal mutations to a single gene encoding a NLR protein. This approach that combines EMS mutagenesis, chromosome flow sorting and short-read sequencing was named MutChromSeq (Figure 3a) (78). A third rapid gene cloning approach used chromosome flow sorting in combination with long-range scaffolding to rapidly generate high-quality assemblies of target intervals identified through genetic mapping (Figure 3b). By using this approach, the leaf rust resistance gene *Lr22a*, previously mapped to chromosome 2D (31), was isolated. To sidestep the tedious process of assembling the *Lr22a* target region through multiple rounds of BAC library screening, chromosome 2D from the *Lr22a* containing wheat line ‘CH Campala *Lr22a*’ was isolated and the corresponding physical interval was rapidly established by using a combination of short-read sequencing with chromosome contact maps of *in vitro* reconstituted chromosomes (Dovetail Genomics Chicago method) (71; 90). The resulting assembly had a scaffold N50 of 9.76 Mb. In other words, this approach generated scaffolds of the target region that were 50-100-fold longer than a BAC clone. *Lr22a* was found on a single scaffold of 6.39 Mb that contained both *Lr22a* flanking markers. Using five independent EMS-derived loss-of-function mutants, *Lr22a* was shown to encode a NLR protein (90). Hence, this approach named ‘targeted chromosome-based cloning via long-range assembly (TACCA)’ eliminates the need to construct BAC libraries for gene cloning. MutRenSeq, MutChromSeq and TACCA provide a significant reduction in the time and costs needed to isolate genes. While it was not unusual to spend 5-10 years on gene cloning project in the past, these novel approaches allow to isolate genes in less than two years.

Other recent developments allow to reduce the time and efforts needed to validate the phenotypic effects of specific candidate genes. For example, the exome captured coding sequences of 2,735

EMS mutagenized plants, 1,535 in the background of the tetraploid durum variety Kronos and 1,200 in the genetic background of the UK hexaploid wheat cultivar Cadenza, were recently sequenced and deposited in a public database. This work revealed more than 10 million mutations in 48,172 and 73,895 gene models in Kronos and Cadenza, respectively (51). This resource was already successfully used to validate candidate genes in the *Stb6* target region, which helped to determine the WAK-like receptor-like kinase as the *Stb6* gene (75).

Genomics of genebanks – unlocking the treasure of diversity to increase disease resistance

A greatly underutilized resource for the improvement of disease resistance consists in the hundreds of thousands of wheat accessions that are stored and maintained in genebanks worldwide (7; 54). These collections hold an enormous richness in genetic diversity, novel disease resistance genes and alleles. Wheat breeders have used old landraces and wild wheat progenitor to increase the diversity of elite wheat gene pools, for example through the generation of synthetic wheat or through alien introgressions of chromosome fragments from wild grass species. These attempts aim to compensate for the loss of diversity that went along domestication and modern breeding (87). One of the most extensive synthetic wheat breeding programs has been performed at the International Wheat and Maize Improvement Center (CIMMYT), where more than 1,000 synthetic wheat lines from more than 900 different *Aegilops tauschii* accessions have been produced since the 1990s (22). One of the major drawbacks of using old landraces and wild wheat progenitors in modern breeding is that undesired traits resulting in reduced yield or quality are often co-introduced when landraces or wild progenitors are crossed with elite wheat cultivars. The use of genomic information for marker-assisted selection would allow to break this linkage drag and to specifically introduce desired genes. However, even with the recent advances in genome sequencing, the generation of whole genome assemblies from thousands of wheat accessions is not yet feasible. Hence, to assess the genetic variation in large collections different genotyping methods are required. For example, several array-based SNP platforms have recently been developed that allow to rapidly and cost-effectively genotype many different wheat accessions with a pre-defined set of markers. These arrays contain between 9,000 and 819,571 SNP markers and are available from different SNP genotyping platforms such as the Affimetrix's Axiom or Illumina's Infinium technology (2; 16; 72; 92; 96). For example, the wheat 9K iSelect Beadchip Assay was used to genotype 2,994 diverse hexaploid wheat accessions including landraces and

modern cultivars (16). The so called wheat breeders' array containing 35,143 SNP markers was tested on a set of 1,779 hexaploid wheat accessions (2). Other approaches to rapidly genotype many wheat accessions include exome capture (49) and genotyping-by-sequencing (69). Current projects aim to obtain a comprehensive genomic insight into entire cereal genebank collections. For example, the goal of the 'Biodiversity informatics to bridge the gap from genome information to educated utilization of genetic diversity hosted in Genebanks' (BRIDGE) project consists in the genotyping of 20,000 barley accessions held in the IPK genebank by using genotyping-by-sequencing (<http://apex.ipk-gatersleben.de/apex/f?p=164:1>). Similarly, the 'Seeds of Discovery' project of CIMMYT aims to genotype wheat accessions of the CIMMYT genebank and to complement these data with phenotypic information. The combination of these genotypic data with phenotypic information will allow to discover marker-trait associations through genome-wide association studies (GWAS) and to uncover novel disease resistance genes. While in the past, producing the genotypic information might have been the bottleneck, the limitation in the number of accessions that can be evaluated nowadays comes from the space and effort needed for accurate phenotyping.

Beside SNP genotyping platforms, approaches like the NLR capture arrays can be used to assess the diversity of NLR genes in genebank collections. Since this approach allows to greatly reduce the size and complexity of the wheat genome, the sequencing of the NLRome from a single wheat cultivar becomes relatively cheap and can be applied to many wheat accessions.

GENOMICS OF WHEAT PATHOGENS - IMPLICATIONS FOR RESISTANCE BREEDING AND DISEASE CONTROL

The increasing pace of genome-based molecular identification of disease resistance genes in wheat is complemented by an increased understanding of the molecular components from the pathogen involved in the resistance interactions. The combination of these different fields of research and a complete molecular comprehension of pathosystems will allow breeders to integrate pathogen-based information into the development of sustainable strategies for the use of resistance genes. The major wheat pathogens are fungal organisms and some of their genomes have been completely sequenced. Depending on the technology used and the genome size of the respective pathogen, the available genome sequences are of different quality in terms of number and contiguity of

fragments: the complete genomes of *Stagonospora nodorum* (30), wheat stem rust (*Puccinia graminis f.sp. tritici*) (23), *Zymoseptoria tritici* (88), wheat stripe rust (*Puccinia striiformis f.sp. tritici*) (15), wheat powdery mildew (*Blumeria graminis f.sp. tritici*) (95), wheat leaf rust (*Puccinia triticina*) (50) and the transcriptome of wheat blast (44) have been determined. These genome sequences have been complemented for some pathogens by other genomic tools such as BAC libraries and physical maps (65; 95) and all these resources have been used for the development of pathogen-based strategies to improve resistance. Here, only a few selected studies will be presented and discussed in more detail to demonstrate the applications of such knowledge for resistance breeding.

The molecular analysis of wheat-pathogen interaction has focused mostly on the host, but a more complete understanding of the pathosystems depends on the identification of the molecular factors from the pathogen recognized by the immune receptors encoded by wheat resistance genes. The molecular identification of such recognized pathogen factors, also called avirulence (*Avr*) proteins, has lagged behind in wheat pathogens. Only recently, the new genomic tools have allowed the isolation of a number of avirulence genes. GWAS have been successful for the identification of virulence/avirulence genes in surprisingly small populations (28; 70) and population genomics of fungal wheat pathogens is a highly productive area of research (29).

In wheat powdery mildew, two *Avr* genes have been isolated using map-based cloning approaches (8) or combined methods integrating GWAS (70). The activity of the cloned genes as specific *Avr* genes was confirmed in transient expression assays in *Nicotiana* which result in cell death when corresponding NLR and *Avr* genes are co-infiltrated. This transient assay will also be of great importance for a future identification of many additional *Avr* genes. Basically, it should be possible to identify in the near future all *Avr* genes for which the corresponding immune receptor gene has been cloned from the wheat genome. Interestingly, for wheat powdery mildew it was found that a suppressor gene is also involved in recognition specificity (8; 9). Thus, the pathogen can “hide” the presence of an *Avr* gene by expressing a suppressor gene and thus avoids recognition. If the corresponding immune receptor is not present in a host, the pathogen can easily regain the *Avr* function (possibly an effector function increasing virulence) by removing the suppressor gene by recombination.

The *AvrPm2* gene is recognized by the wheat *Pm2* resistance gene. It was found that *AvrPm2* is not only present in wheat mildew, but that a highly similar gene is also present in rye powdery

mildew and the product of this rye mildew gene is recognized by the Pm2 immune receptor. Possibly this provides a general explanation for the successful introgression breeding of resistance genes from wheat relatives into wheat such as the wheat-rye translocations: a subset of *Avr* genes might be present in mildews specialized on different hosts and, therefore, NLR resistance genes are active in different hosts against different host-specialized mildews (62; 70).

The recent identification of the *AvrSr35* and *AvrSr50* genes in wheat stem rust (17; 77) as well as of *AvrStb6* in *Zymoseptoria* (98) represent important steps forward for our understanding of these host-pathogen interactions. These discoveries were all based on whole-genome analysis and describe the first *Avr* genes from these species. In *Zymoseptoria*, the identification of *AvrStb6* complements the molecular identification of the *Stb6* resistance gene (75). Very interestingly, *AvrStb6* encodes a secreted, small effector protein, very similar to the fungal *Avr* proteins recognized by intracellular immune receptors of the NLR class (8; 70). However, as described above, the corresponding *Stb6* gene encodes a receptor-like kinase with an extracellular WAK-like domain. From work in the model plant *Arabidopsis thaliana*, it has previously been established that WAK domains generally bind oligosaccharides (12). If we assumed a direct interaction between *Stb6* and *AvrStb6*, this would likely occur in the apoplastic space and *Stb6* would thus perceive a protein ligand rather than an oligosaccharide. The molecular details of this putative interaction remain to be determined and represent a challenging novel frontier in disease resistance biology. Clearly, it is highly important to further study WAK resistance genes, given that there are several hundred WAK encoding genes in the wheat genome, with *Stb6* being the only member with an identified function (42). Also, WAKs were recently found to be key players in disease resistance against fungal pathogens in other cereals, including rice and maize (33; 38; 99).

The molecular identification of genes encoding avirulence as well as suppressor genes will result in the development of specific markers for them, allowing to easily determine their presence and evolution in pathogen populations. This will simplify the analysis of pathogen avirulence patterns over time and in specific agricultural regions and contribute to the rational resistance gene deployment in agriculture. Furthermore, and possibly in a more distant future, molecular knowledge on recognized *Avr* proteins might allow the molecular breeder to select combinations of genes with complementary specificities and lower risks for pathogen adaptation. Pathogens can rapidly evolve and adapt to disease resistance, particularly in agricultural ecosystems (64). Studies at the genomic level have recently revealed the evolutionary history of

two newly emerged pathogens. One of this is the wheat blast pathogen that originated in South America and is now threatening wheat production in parts of Asia. It was found that wheat blast has evolved by the functional loss of a host specificity determinant (39).

Around the year 2000 a novel powdery mildew form was discovered that is able to grow on the rye-wheat amphiploid crop species Triticale. It was found that this novel form of mildew was the product of a hybridization event between the mildew forms specialized on wheat and rye. Thus, pathogen evolution mirrored host evolution of Triticale (63). This finding also indicates that such hybridization events among different powdery mildew forms have to be considered for the use of resistance gene introgressions from wild wheat relatives. Obviously, wheat mildew can exchange genetic material with mildews from other grass species, resulting in novel virulence properties and a powerful evolutionary mechanism to overcome specific resistance genes introgressed from wheat relatives.

Pathogen genomics has also proven to be a powerful tool in the identification of the causal pathogen races of new stripe rust epidemics in the UK (35). In an approach called “field pathogenomics”, transcriptome sequencing was performed on infected wheat leaves from the field. This showed that the pathogen population dramatically changed within a few years and increased in diversity. Field pathogenomics is a very promising way for genotyping a large number of isolates and can rapidly deliver a view on an emerging epidemic. Thus, it has the potential for a wide application to many pathogens and at a global scale.

As a final example of pathogen genomics, the global, transcriptional response in pathogens to the presence of the durable, quantitative *Lr34* resistance gene is discussed here (52). As mentioned above, *Lr34* confers resistance against several pathogen species. Transcriptomes of leaf rust and powdery mildew pathogens growing on partially resistant, *Lr34*-containing wheat or barley host plants, respectively, were compared to pathogens growing on (near-)isogenic host genotypes without *Lr34* (86). Surprisingly, there was no transcriptional response in the pathogens to the presence of *Lr34*, despite slower and reduced growth. Thus, pathogens do not react to slower growth caused by *Lr34* and it is tempting to speculate that this might be a general property of quantitative and partial disease resistance genes which could be used as a predictive feature of durable resistance. This hypothesis can now be tested for additional such quantitative resistance genes for which isogenic lines are available.

The rapid developments in fungal pathogenomics exemplified above complement the molecular analysis on the host side, and breeding strategies can now integrate the knowledge on pathogen evolution and molecular biology. There is a rapidly developing field of research based on pathogen-informed strategies which takes into consideration these findings at the molecular, functional, evolutionary and population level for improvement in resistance breeding strategies(13).

GENOMIC-ASSISTED DISEASE RESISTANCE BREEDING – A LOOK INTO THE NEAR FUTURE

Despite considerable international effort, only around 6% of the known and mapped wheat disease resistance genes have been cloned so far, not to mention the yet undiscovered genes hidden in genebank collections. With the availability of high-quality wheat reference sequences, high-throughput marker technologies and several rapid gene isolation strategies, the cloning of disease resistance genes will become easier, quicker and less costly. Hence, it is now the right time to discuss the cloning of most of the 400 known disease resistance genes (61) in an international effort and to study the critical sequence polymorphisms that distinguish susceptible from resistant alleles. This work could reveal perfect functional molecular markers of an extensive gene set that could be used for breeding. Wheat breeders could choose their ‘preferred’ resistance genes based on the diseases prevalent in a particular region and ‘stack’ these genes in single cultivars through marker-assisted selection (Figure 4). A particular challenge however remains that a considerable number of segregating genes from different sources need to be combined in a single cultivar to achieve durable resistance. These genes then need to be maintained in subsequent crosses to avoid the ‘thinning out’ of disease resistance. A strategy that could overcome this problem is through the use of ‘gene cassettes’, where several disease resistance genes are cloned into the same vector and jointly transformed into susceptible cultivars (24; 97). The bottleneck of this strategy has so far been limitations in transformation technology since the resulting constructs would likely be several dozen kb in length. As a different approach, disease resistance could also be engineered by the modification of specific genes through genome editing. A relatively simple approach would be to use genome editing in order to eliminate susceptibility genes. It has been shown that some necrotrophic fungi require host specific susceptibility genes in order to establish infection (25; 80).

450 These susceptibility genes show a remarkable similarity to genes that confer resistance against
451 biotrophic fungi and it has been suggested that necrotrophic pathogens can hijack the plant's
452 disease resistance mechanisms and provoke cell death through hypersensitive response (HR). HR
453 is a common outcome of resistance reactions against biotrophic pathogens that depend on living
454 host tissue, but HR does provide the basis for necrotrophic fungi to proliferate on dead host tissue.
455 For example, the wheat *Tsn1* gene encodes a protein with a leucine-rich repeat domain similar to
456 the one found in NLR proteins. *Tsn1* is required for the successful inoculation by the fungal
457 pathogens that cause *Stagonospora nodorum* blotch and tan spot (27). These fungi produce the
458 host-selective toxin ToxA and it has been hypothesized that ToxA perception by Tsn1 triggers HR.
459 Similarly, the wheat *Snn1* gene, coding for a wall-associated kinase (WAK), has been shown to
460 perceive the toxin SnTox1 of the necrotrophic pathogen *Parastagonospora nodorum* (80). The
461 removal of susceptibility genes like *Tsn1* or *Snn1* through marker-assisted selection or genome
462 editing would permanently increase disease resistance against necrotrophic fungi. However, it is
463 likely that genes like *Tsn1* or *Snn1* may have important functions in wheat plants other than serving
464 as entry points for pathogens. It therefore needs to be carefully evaluated whether the systematic
465 removal of susceptibility genes compromises wheat yield or quality. Genome editing might also
466 be particularly useful in polyploid wheat species because it allows simultaneous targeting of the
467 homoeologous gene copies on all the three sub-genomes. For example, the *mlo* resistance of
468 diploid barley provides one of the most durable sources of resistance against the fungal powdery
469 mildew disease (*B. graminis* f.sp *hordei*). The *mlo* resistance in barley arose through natural and
470 induced loss-of-function mutations in the *Mildew resistance locus O* (*Mlo*) gene coding for a
471 plasma membrane-localized protein with unknown biochemical function (14). Natural *mlo*
472 resistance is not known in polyploid wheat because it would require simultaneous knock-outs in
473 all three homoeologous *Mlo* copies. However, it was recently shown that targeting the three
474 homoeologs of *Mlo* through genome editing or EMS-mutagenesis resulted in increased resistance
475 of wheat against wheat powdery mildew (*B. graminis* f.sp. *tritici*) (1; 93). Work on the *Pm3*
476 powdery mildew resistance gene of wheat that codes for a NLR protein, showed that single amino
477 acid polymorphisms can result in an increased resistance spectrum or stronger disease signaling
478 (11; 85). These studies showed that it would be possible to engineer broad-spectrum NLR genes,
479 although the targeted modification of specific amino acids is more difficult to achieve than loss-
480 of-function modifications. Last, the emerging field of pathogenomics described above will allow

to precisely monitor the presence of avirulence genes in pathogen populations. This knowledge can then be used to choose and grow wheat cultivars with disease resistance genes that are effective against the pathogen populations prevalent in a certain region. Hence, instead of “hoping” that wheat cultivars will be resistant against the prevalent pathogens, the choices of which disease resistance genes to deploy can be based on genomic information from pathogens, which will greatly increase the effectiveness of disease resistance in the field.

SUMMARY POINTS

1. Based on recently developed genomic resources gene isolation in wheat has become easier and there is a rapidly increasing number of cloned disease resistance genes.
2. The first high quality reference sequences of wheat genomes are available and reveal a large number of candidate resistance genes of the NLR and WAK class.
3. New approaches based on sequencing complete, flow-sorted chromosomes or array-based technologies allow to rapidly isolate resistance genes based on mutants. These approaches are independent of the chromosomal position, do not depend on recombination and can be used in any genotype of interest.
4. Based on pathogenomics, there has been rapid progress in the isolation of avirulence genes in several fungal wheat pathogens. For these genes, the corresponding resistance genes have been isolated and the molecular interactions can be studied.
5. Knowledge on molecularly identified susceptibility or resistance suppressor genes can be used for selection against their presence or for inactivation by gene editing.
6. Genomics both on the host and the pathogen side is providing resistance breeding with novel tools such as perfect markers, isolated genes for gene cassettes and pathogen-informed breeding strategies.

FUTURE ISSUES

1. Genome sequences from dozens of wheat genomes with a similar quality as the reference genome are needed to define the pan-genome for resistance gene loci. This will include genotypes of important wild grass donors of resistance genes.
2. A large international collaboration is needed to isolate all resistance genes described in the wheat gene catalog and which are important for breeding. Novel genes identified in

genebank material need to be included. Such a goal has become feasible based on novel and future chromosome and genome-wide strategies based on mutants.

3. Quantitative trait loci for durable resistance should be approached in a similar way. This needs the development of near-isogenic lines and high quality phenotyping to identify mutants in such lines.
4. For each major wheat disease, appropriate sets of lines should be phenotyped and genotyped for GWAS. This will allow identification and isolation of novel genes which have not yet been genetically described.
5. For all disease resistance genes important in breeding, the corresponding avirulence genes should be isolated to get a comprehensive view on the molecular interactions and to develop pathogen-informed breeding strategies.
6. Field pathogenomics should be performed for all major wheat diseases in all major growth regions of wheat. This will result in a global view on pathogen diversity and evolution.
7. Research is needed to identify the best strategies for the integration of the rapidly developing knowledge in different fields into efficient regional breeding programs.

DISCLOSURE STATEMENT

The authors are not aware of any affiliation, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGEMENTS

The work on this review was supported by the King Abdullah University of Science and Technology and by a Grant from the Swiss National Science Foundation (310030_163260).

FIGURE LEGENDS

Figure 1. Genomic advancements and their implications for the cloning of disease resistance genes in wheat. Key advances in genomics are indicated in red, such as the introduction of next-generation sequencing (NGS) technologies in 2007, the release of the wheat draft genome sequence in 2014 and the Chinese Spring reference sequence (RefSeq v1.0) in 2017. In addition, novel gene cloning approaches such as MutRenSeq, MutChromSeq and TACCA that are discussed in the text are indicated. Genes isolated by map-based cloning are shown in blue, genes isolated through homology-based cloning in green and genes isolated with novel gene cloning approaches in orange. The trend line corresponds to an exponential growth curve.

Figure 2. Distribution of NLR genes across wheat chromosomes. Shown are group 1 chromosomes as representatives for the whole genome. NLR genes are enriched in distal regions of chromosomes. Locations of individual genes are indicated with blue lines. Density of NLR genes is indicated as a bar plot in windows of 3 Mb. The inset shows the structure of the *Pm3* powdery mildew resistance locus on chromosome 1A. The gray boxes which indicate the genes are not drawn to scale.

Figure 3. Cloning of target genes using next-generation sequencing technologies. a.

Chromosomes from individual mutagenized wheat plants are isolated via cytometric flow-sorting. Illumina shotgun sequences are then assembled into (usually several 1000) sequence contigs. The target gene is identified by mapping of Illumina reads of the individual mutants on to the sequence contigs (i.e. the target gene carries mutations in all independent mutants). **b.** The chromosome containing the target gene is isolated via flow-sorting and sequenced with Illumina technology. Contact maps (e.g. through re-constituted chromatin) are used to construct megabase-size scaffolds. The target interval on a single scaffold is identified by mapping genetic markers on to the scaffolds identifies the target interval on a single scaffold. The target gene is identified by mapping of Illumina reads of the individual mutants on to the sequence scaffolds.

Figure 4. Genomic-informed breeding strategies that could be used to increase disease resistance in wheat.

LITERATURE CITED

1. Acevedo-Garcia J, Spencer D, Thieron H, Reinstadler A, Hammond-Kosack K, et al. 2017. *mlo*-based powdery mildew resistance in hexaploid bread wheat generated by a non-transgenic TILLING approach. *Plant Biotechnology Journal* 15:367-78
2. Allen AM, Winfield MO, BurrIDGE AJ, Downie RC, Benbow HR, et al. 2017. Characterization of a Wheat Breeders' Array suitable for high-throughput SNP genotyping of global accessions of hexaploid bread wheat (*Triticum aestivum*). *Plant Biotechnology Journal* 15:390-401
3. Arumuganathan K, Earle ED. 1991. Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter* 9:208-18
4. Avni R, Nave M, Barad O, Baruch K, Twardziok SO, et al. 2017. Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science* 357:93-7
5. Bennett MD, Smith JB. 1976. Nuclear dna amounts in angiosperms. *Philosophical Transactions of the Royal Society B Biological Sciences* 274:227-74
6. Bevan MW, Uauy C, Wulff BB, Zhou J, Krasileva K, Clark MD. 2017. Genomic innovation for crop improvement. *Nature* 543:346-54
7. Bhullar NK, Zhang Z, 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 Biology* 10:88

- 598 8. Bourras S, McNally KE, Ben-David R, Parlange F, Roffler S, et al. 2015. Multiple avirulence loci and
599 allele-specific effector recognition control the *Pm3* race-specific resistance of wheat to powdery
600 mildew. *Plant Cell* 27:2991-3012
- 601 9. Bourras S, McNally KE, Muller MC, Wicker T, Keller B. 2016. Avirulence genes in cereal powdery
602 mildews: The gene-for-gene hypothesis 2.0. *Frontiers in Plant Science* 7:241
- 603 10. Brechley R, Spannagl M, Pfeifer M, Barker GLA, D'Amore R, et al. 2012. Analysis of the
604 breadwheat genome using whole-genome shotgun sequencing. *Nature* 491:705-10
- 605 11. Brunner S, Hurni S, Streckeisen P, Mayr G, Albrecht M, et al. 2010. Intragenic allele pyramiding
606 combines different specificities of wheat *Pm3* resistance alleles. *Plant J* 64:433-45
- 607 12. Brutus A, Sicilia F, Macone A, Cervone F, De Lorenzo G. 2010. A domain swap approach reveals a
608 role of the plant wall-associated kinase 1 (WAK1) as a receptor of oligogalacturonides. *Proc Natl*
609 *Acad Sci U S A* 107:9452-7
- 610 13. Burdon JJ, Zhan J, Barrett LG, Papaix J, Thrall PH. 2016. Addressing the challenges of pathogen
611 evolution on the world's arable crops. *Phytopathology* 106:1117-27
- 612 14. Buschges R, Hollricher K, Panstruga R, Simons G, Wolter M, et al. 1997. The barley *Mlo* gene: A
613 novel control element of plant pathogen resistance. *Cell* 88:695-705
- 614 15. Cantu D, Govindarajulu M, Kozik A, Wang M, Chen X, et al. 2011. Next generation sequencing
615 provides rapid access to the genome of *Puccinia striiformis* f. sp. *tritici*, the causal agent of wheat
616 stripe rust. *PLoS One* 6:e24230
- 617 16. Cavanagh CR, Chao S, Wang S, Huang BE, Stephen S, et al. 2013. Genome-wide comparative
618 diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces
619 and cultivars. *Proc Natl Acad Sci U S A* 110:8057-62
- 620 17. Chen J, Upadhyaya NM, Ortiz D, Sperschneider J, Li F, et al. 2017. Somatic recombination in wheat
621 stem rust leads to virulence for Ug99-effective *Sr50* resistance. *Science* submitted
- 622 18. Choulet F, Alberti A, Theil S, Glover N, Barbe V, et al. 2014. Structural and functional partitioning
623 of bread wheat chromosome 3B. *Science* 345:1249721
- 624 19. Clavijo BJ, Venturini L, Schudoma C, Accinelli GG, Kaithakottil G, et al. 2017. An improved assembly
625 and annotation of the allohexaploid wheat genome identifies complete families of agronomic
626 genes and provides genomic evidence for chromosomal translocations. *Genome Res* 27:885-96
- 627 20. Dodds PN, Rathjen JP. 2010. Plant immunity: towards an integrated view of plant-pathogen
628 interactions. *Nature Reviews Genetics* 11:539-48
- 629 21. Dolezel J, Vrana J, Safar J, Bartos J, Kubalakova M, Simkova H. 2012. Chromosomes in the flow to
630 simplify genome analysis. *Funct Integr Genomic* 12:397-416
- 631 22. Dreisigacker S, Kishii M, Lage J, Warburton M. 2008. Use of synthetic hexaploid wheat to increase
632 diversity for CIMMYT bread wheat improvement. *Aust J Agr Res* 59:413-20
- 633 23. Duplessis S, Cuomo CA, Lin YC, Aerts A, Tisserant E, et al. 2011. Obligate biotrophy features
634 unraveled by the genomic analysis of rust fungi. *Proc Natl Acad Sci U S A* 108:9166-71
- 635 24. Ellis JG, Lagudah ES, Spielmeier W, Dodds PN. 2014. The past, present and future of breeding rust
636 resistant wheat. *Frontiers in Plant Science* 5:641
- 637 25. Faris JD, Zhang ZC, Lu HJ, Lu SW, Reddy L, et al. 2010. A unique wheat disease resistance-like gene
638 governs effector-triggered susceptibility to necrotrophic pathogens. *Proc Natl Acad Sci U S A*
639 107:13544-9
- 640 26. Feuillet C, Travella S, Stein N, Albar L, Nublat A, Keller B. 2003. Map-based isolation of the leaf rust
641 disease resistance gene *Lr10* from the hexaploid wheat (*Triticum aestivum* L.) genome. *Proc Natl*
642 *Acad Sci U S A* 100:15253-8
- 643 27. Friesen TL, Stukenbrock EH, Liu Z, Meinhardt S, Ling H, et al. 2006. Emergence of a new disease as
644 a result of interspecific virulence gene transfer. *Nat Genet* 38:953-6

28. Gao Y, Liu Z, Faris JD, Richards J, Brueggeman RS, et al. 2016. Validation of genome-wide association studies as a tool to identify virulence factors in *Parastagonospora nodorum*. *Phytopathology* 106:1177-85
29. Grunwald NJ, McDonald BA, Milgroom MG. 2016. Population genomics of fungal and oomycete pathogens. *Annu Rev Phytopathol* 54:323-46
30. Hane JK, Lowe RG, Solomon PS, Tan KC, Schoch CL, et al. 2007. Dothideomycete plant interactions illuminated by genome sequencing and EST analysis of the wheat pathogen *Stagonospora nodorum*. *Plant Cell* 19:3347-68
31. Hiebert CW, Thomas JB, Somers DJ, McCallum BD, Fox SL. 2007. Microsatellite mapping of adult-plant leaf rust resistance gene *Lr22a* in wheat. *Theor Appl Genet* 115:877-84
32. Hirsch CN, Hirsch CD, Brohammer AB, Bowman MJ, Soifer I, et al. 2016. Draft assembly of elite inbred line PH207 provides insights into genomic and transcriptome diversity in maize. *Plant Cell* 28:2700-14
33. Hu K, Cao J, Zhang J, Xia F, Ke Y, et al. 2017. Improvement of multiple agronomic traits by a disease resistance gene via cell wall reinforcement. *Nat Plants* 3:17009
34. 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* 164:655-64
35. Hubbard A, Lewis CM, Yoshida K, Ramirez-Gonzalez RH, de Vallavieille-Pope C, et al. 2015. Field pathogenomics reveals the emergence of a diverse wheat yellow rust population. *Genome Biol* 16:23
36. Hurni S, Brunner S, Buchmann G, Herren G, Jordan T, et al. 2013. Rye *Pm8* and wheat *Pm3* are orthologous genes and show evolutionary conservation of resistance function against powdery mildew. *Plant J* 76:957-69
37. Hurni S, Brunner S, Stirnweis D, Herren G, Peditto D, et al. 2014. The powdery mildew resistance gene *Pm8* derived from rye is suppressed by its wheat ortholog *Pm3*. *Plant J* 79:904-13
38. Hurni S, Scheuermann D, Krattinger SG, Kessel B, Wicker T, et al. 2015. The maize disease resistance gene *Htn1* against northern corn leaf blight encodes a wall-associated receptor-like kinase. *Proc Natl Acad Sci U S A* 112:8780-5
39. Inoue Y, Vy TTP, Yoshida K, Asano H, Mitsuoka C, et al. 2017. Evolution of the wheat blast fungus through functional losses in a host specificity determinant. *Science* 357:80-3
40. International Rice Genome Sequencing Project. 2005. The map-based sequence of the rice genome. *Nature* 436:793-800
41. International Wheat Genome Sequencing Consortium. 2014. A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science* 345:1251788
42. International Wheat Genome Sequencing Consortium. 2018. Shifting the limits in wheat research and breeding through a fully annotated and anchored reference genome sequence. submitted
43. Isidore E, Scherrer B, Chalhoub B, Feuillet C, Keller B. 2005. Ancient haplotypes resulting from extensive molecular rearrangements in the wheat A genome have been maintained in species of three different ploidy levels. *Genome Res* 15:526-36
44. Islam MT, Croll D, Gladieux P, Soanes DM, Persoons A, et al. 2016. Emergence of wheat blast in Bangladesh was caused by a South American lineage of *Magnaporthe oryzae*. *BMC Biology* 14:84
45. Jia J, Zhao S, Kong X, Li Y, Zhao G, et al. 2013. *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature* 496:91-5
46. Jiao Y, Peluso P, Shi J, Liang T, Stitzer MC, et al. 2017. Improved maize reference genome with single-molecule technologies. *Nature* 546:524-7
47. Jordan KW, Wang S, Lun Y, Gardiner LJ, MacLachlan R, et al. 2015. A haplotype map of allohexaploid wheat reveals distinct patterns of selection on homoeologous genomes. *Genome Biol* 16:48

693 48. Jordan T, Seeholzer S, Schwizer S, Toller A, Somssich IE, Keller B. 2011. The wheat *Mla* homologue
 694 *TmMla1* exhibits an evolutionarily conserved function against powdery mildew in both wheat and
 695 barley. *Plant J* 65:610-21
 696 49. King R, Bird N, Ramirez-Gonzalez R, Coghill JA, Patil A, et al. 2015. Mutation scanning in wheat by
 697 exon capture and next-generation sequencing. *PLoS One* 10:e0137549
 698 50. Kiran K, Rawal HC, Dubey H, Jaswal R, Devanna BN, et al. 2016. Draft genome of the wheat rust
 699 pathogen (*Puccinia triticina*) unravels genome-wide structural variations during evolution.
 700 *Genome Biology and Evolution* 8:2702-21
 701 51. Krasileva KV, Vasquez-Gross HA, Howell T, Bailey P, Paraiso F, et al. 2017. Uncovering hidden
 702 variation in polyploid wheat. *Proc Natl Acad Sci U S A* 114:E913-E21
 703 52. Krattinger SG, Lagudah ES, Spielmeier W, Singh RP, Huerta-Espino J, et al. 2009. A putative ABC
 704 transporter confers durable resistance to multiple fungal pathogens in wheat. *Science* 323:1360-
 705 3
 706 53. Ling HQ, Zhao S, Liu D, Wang J, Sun H, et al. 2013. Draft genome of the wheat A-genome progenitor
 707 *Triticum urartu*. *Nature* 496:87-90
 708 54. Longin CF, Reif JC. 2014. Redesigning the exploitation of wheat genetic resources. *Trends in Plant*
 709 *Science* 19:631-6
 710 55. Luo MC, Gu YQ, Puiu D, Wang H, Twardziok SO, et al. 2017. Genome sequence of the progenitor
 711 of the wheat D genome *Aegilops tauschii*. *Nature* 551:498-502
 712 56. Ma Z, Weining S, Sharp PJ, Liu C. 2000. Non-gridded library: a new approach for BAC (bacterial
 713 artificial chromosome) exploitation in hexaploid wheat (*Triticum aestivum*). *Nucleic Acids Res*
 714 28:e106
 715 57. Mago R, Tabe L, Vautrin S, Simkova H, Kubalakova M, et al. 2014. Major haplotype divergence
 716 including multiple germin-like protein genes, at the wheat *Sr2* adult plant stem rust resistance
 717 locus. *BMC Plant Biology* 14:379
 718 58. Mago R, Zhang P, Vautrin S, Simkova H, Bansal U, et al. 2015. The wheat *Sr50* gene reveals rich
 719 diversity at a cereal disease resistance locus. *Nat Plants* 1:15186
 720 59. Mascher M, Gundlach H, Himmelbach A, Beier S, Twardziok SO, et al. 2017. A chromosome
 721 conformation capture ordered sequence of the barley genome. *Nature* 544:427-33
 722 60. Mayer KF, Taudien S, Martis M, Simkova H, Suchankova P, et al. 2009. Gene content and virtual
 723 gene order of barley chromosome 1H. *Plant Physiology* 151:496-505
 724 61. McIntosh RA, Yamazaki Y, Dubcovsky J, Rogers J, Morris C, et al. 2008. Catalogue of gene symbols
 725 for wheat. In *Proceedings of the 11th International Wheat Genetics Symposium*, ed. R Appels, R
 726 Eastwood, E Lagudah, P Langridge, M Mackay, et al. Sydney: Sydney University Press. Number of.
 727 62. Menardo F, Praz CR, Wicker T, Keller B. 2017. Rapid turnover of effectors in grass powdery mildew
 728 (*Blumeria graminis*). *BMC Evolutionary Biology* 17:223
 729 63. Menardo F, Praz CR, Wyder S, Ben-David R, Bourras S, et al. 2016. Hybridization of powdery
 730 mildew strains gives rise to pathogens on novel agricultural crop species. *Nat Genet* 48:201-5
 731 64. Moller M, Stukenbrock EH. 2017. Evolution and genome architecture in fungal plant pathogens.
 732 *Nat Rev Microbiol* 15:756-71
 733 65. Parlange F, Oberhaensli S, Breen J, Platzer M, Taudien S, et al. 2011. A major invasion of
 734 transposable elements accounts for the large size of the *Blumeria graminis* f.sp. *tritici* genome.
 735 *Funct Integr Genomic* 11:671-7
 736 66. Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, et al. 2009. The *Sorghum bicolor*
 737 genome and the diversification of grasses. *Nature* 457:551-6
 738 67. Paux E, Sourdille P, Salse J, Sautenac C, Choulet F, et al. 2008. A physical map of the 1-gigabase
 739 bread wheat chromosome 3B. *Science* 322:101-4

740 68. Periyannan S, Moore J, Ayliffe M, Bansal U, Wang XJ, et al. 2013. The gene *Sr33*, an ortholog of
741 barley *Mla* genes, encodes resistance to wheat stem rust race Ug99. *Science* 341:786-8
742 69. Poland JA, Brown PJ, Sorrells ME, Jannink JL. 2012. Development of high-density genetic maps for
743 barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS One*
744 7:e32253
745 70. Praz CR, Bourras S, Zeng F, Sanchez-Martin J, Menardo F, et al. 2017. *AvrPm2* encodes an RNase-
746 like avirulence effector which is conserved in the two different specialized forms of wheat and rye
747 powdery mildew fungus. *New Phytologist* 213:1301-14
748 71. Putnam NH, O'Connell BL, Stites JC, Rice BJ, Blanchette M, et al. 2016. Chromosome-scale shotgun
749 assembly using an in vitro method for long-range linkage. *Genome Res* 26:342-50
750 72. Rasheed A, Hao Y, Xia X, Khan A, Xu Y, et al. 2017. Crop breeding chips and genotyping platforms:
751 progress, challenges, and perspectives. *Molecular Plant* 10:1047-64
752 73. Reynolds MP, Borlaug NE. 2006. Impacts of breeding on international collaborative wheat
753 improvement. *Journal of Agricultural Science* 144:3-17
754 74. Safar J, Simkova H, Kubalakova M, Cihalikova J, Suchankova P, et al. 2010. Development of
755 chromosome-specific BAC resources for genomics of bread wheat. *Cytogenet Genome Res*
756 129:211-23
757 75. Saintenac C, Lee WS, Cambon F, Rudd JJ, King R, et al. 2018. A pattern recognition receptor-like
758 protein conserved in wheat controls gene-for-gene resistance to an extracellular fungal pathogen.
759 *Nat Genet* online
760 76. Salamini F, Ozkan H, Brandolini A, Schafer-Pregl R, Martin W. 2002. Genetics and geography of
761 wild cereal domestication in the Near East. *Nature Reviews Genetics* 3:429-41
762 77. Salcedo A, Rutter W, Wang S, Akhunova A, Bolus S, et al. 2017. Variation in the *AvrSr35* gene
763 determines *Sr35* resistance against wheat stem rust race Ug99. *Science* submitted
764 78. Sanchez-Martin J, Steuernagel B, Ghosh S, Herren G, Hurni S, et al. 2016. Rapid gene isolation in
765 barley and wheat by mutant chromosome sequencing. *Genome Biol* 17:221
766 79. Seeholzer S, Tsuchimatsu T, Jordan T, Bieri S, Pajonk S, et al. 2010. Diversity at the *Mla* powdery
767 mildew resistance locus from cultivated barley reveals sites of positive selection. *Molecular Plant-
768 Microbe Interactions* 23:497-509
769 80. Shi G, Zhang Z, Friesen TL, Raats D, Fahima T, et al. 2016. The hijacking of a receptor kinase-driven
770 pathway by a wheat fungal pathogen leads to disease. *Science Advances* 2:e1600822
771 81. Singh RP, Hodson DP, Huerta-Espino J, Jin Y, Bhavani S, et al. 2011. The emergence of Ug99 races
772 of the stem rust fungus is a threat to world wheat production. *Annu Rev Phytopathol* 49:465-81
773 82. Song WY, Wang GL, Chen LL, Kim HS, Pi LY, et al. 1995. A receptor kinase-like protein encoded by
774 the rice disease resistance gene, *Xa21*. *Science* 270:1804-6
775 83. Stein N, Feuillet C, Wicker T, Schlagenhauf E, Keller B. 2000. Subgenome chromosome walking in
776 wheat: A 450-kb physical contig in *Triticum monococcum* L. spans the *Lr10* resistance locus in
777 hexaploid wheat (*Triticum aestivum* L.). *Proc Natl Acad Sci U S A* 97:13436-41
778 84. Steuernagel B, Periyannan SK, Hernandez-Pinzon I, Witek K, Rouse MN, et al. 2016. Rapid cloning
779 of disease-resistance genes in plants using mutagenesis and sequence capture. *Nature
780 Biotechnology* 34:652-5
781 85. Stirnweis D, Milani SD, Jordan T, Keller B, Brunner S. 2014. Substitutions of two amino acids in the
782 nucleotide-binding site domain of a resistance protein enhance the hypersensitive response and
783 enlarge the PM3F resistance spectrum in wheat. *Mol Plant Microbe In* 27:265-76
784 86. Sucher J, Menardo F, Praz CR, Boni R, Krattinger SG, Keller B. 2017. Transcriptional profiling reveals
785 no response of fungal pathogens to the durable, quantitative *Lr34* disease resistance gene of
786 wheat. *Plant Pathology* online

87. Tanksley SD, McCouch SR. 1997. Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 277:1063-6
88. Testa A, Oliver R, Hane J. 2015. Overview of genomic and bioinformatic resources for *Zymoseptoria tritici*. *Fungal Genetics and Biology* 79:13-6
89. The International Brachypodium Initiative. 2010. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* 463:763-8
90. Thind AK, Wicker T, Simkova H, Fossati D, Moullet O, et al. 2017. Rapid cloning of genes in hexaploid wheat using cultivar-specific long-range chromosome assembly. *Nature Biotechnology* 35:793-6
91. Uauy C, Wulff BBH, Dubcovsky J. 2017. Combining traditional mutagenesis with new high-throughput sequencing and genome editing to reveal hidden variation in polyploid wheat. *Annual Review of Genetics* 51:435-54
92. Wang SC, Wong DB, Forrest K, Allen A, Chao SM, et al. 2014. Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnology Journal* 12:787-96
93. Wang Y, Cheng X, Shan Q, Zhang Y, Liu J, et al. 2014. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nature Biotechnology* 32:947-51
94. Wei F, Gobelmann-Werner K, Morroll SM, Kurth J, Mao L, et al. 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* 153:1929-48
95. Wicker T, Oberhaensli S, Parlange F, Buchmann JP, Shatalina M, et al. 2013. The wheat powdery mildew genome shows the unique evolution of an obligate biotroph. *Nat Genet* 45:1092-6
96. Winfield MO, Allen AM, Burrige AJ, Barker GL, Benbow HR, et al. 2016. High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. *Plant Biotechnology Journal* 14:1195-206
97. Wulff BB, Moscou MJ. 2014. Strategies for transferring resistance into wheat: from wide crosses to GM cassettes. *Frontiers in Plant Science* 5:692
98. Zhong Z, Marcel TC, Hartmann FE, Ma X, Plissonneau C, et al. 2017. A small secreted protein in *Zymoseptoria tritici* is responsible for avirulence on wheat cultivars carrying the *Stb6* resistance gene. *New Phytologist* 214:619-31
99. Zuo W, Chao Q, Zhang N, Ye J, Tan G, et al. 2015. A maize wall-associated kinase confers quantitative resistance to head smut. *Nat Genet* 47:151-7