Cloning of the wheat Yr15 resistance gene sheds light on the plant tandem kinase-pseudokinase family

Valentina Klymiuk1,2, Elitsur Yaniv1,2,3,4, Lin Huang1,2,5, Dina Raats1,2,6, Andrii Fatiukha1,2, Shisheng Chen7, Lihua Feng1,2, Zeev Frenkel1, Tamar Krugman1, Gabriel Lidzbarsky1,1, Wei Chang3,4, Marko J. Jääskeläinen3,4, Christian Schudoma6, Lars Paulin3, Pia Laine3, Harbans Bariana8, Hanan Sela1,9, Kamran Saleem10, Chris Khadgi Sørensen10, Mogens S. Hovmøller10, Assaf Distelfeld1,11, Boulos Chalhoub12, Jorge Dubcovsky7,13, Abraham B. Korol1,2, Alan H. Schulman3,4,14 & Tzion Fahima1,2

Yellow rust, caused by Puccinia striiformis f. sp. tritici (Pst), is a devastating fungal disease threatening much of global wheat production. Race-specific resistance (R)-genes are used to control rust diseases, but the rapid emergence of virulent Pst races has prompted the search for a more durable resistance. Here, we report the cloning of Yr15, a broad-spectrum R-gene derived from wild emmer wheat, which encodes a putative kinase-pseudokinase protein, designated as wheat tandem kinase 1, comprising a unique R-gene structure in wheat. The existence of a similar gene architecture in 92 putative proteins across the plant kingdom, including the barley RPG1 and a candidate for Ug8, suggests that they are members of a distinct family of plant proteins, termed here tandem kinase-pseudokinases (TKPs). The presence of kinase-pseudokinase structure in both plant TKPs and the animal Janus kinases sheds light on the molecular evolution of immune responses across these two kingdoms.
heat was domesticated in the Fertile Crescent over 10,000 years ago. Today, wheat is cultivated more extensively than any other food crop, covering more than 244 million ha and yielding more than 750 million tons annually (Food and Agriculture Organization Corporate Statistical Database (FAOSTAT)). It is a major component in human diet, providing starch, proteins, vitamins, dietary fiber, and phytochemicals.

Wheat crops are subjected to numerous types of biotic and abiotic stresses, with fungal diseases comprising one of the most serious threats to wheat production. Yellow (stripe) rust disease of wheat is caused by the fungus *Puccinia striiformis f. sp. triticii* (*Pst*). Due to the rapid evolution of the pathogen and development of new virulent races, severe yield losses occurred during recent decades in Europe and worldwide. Given that 88% of wheat production is susceptible to yellow rust, more than five million tons of wheat harvest, valued at around one billion USD, are estimated to be lost annually due to this pathogen. These epidemics demonstrate that disease management relying on single race-specific resistance genes has become ineffective. As an alternative, broad-spectrum resistance genes are much sought after by breeders, because such genes generally can provide robust protection against diseases, in particular against yellow rust.

The plant surveillance system can detect pathogen-secreted molecules and activate powerful defense responses, which often trigger local programmed cell death (PCD) at the site of attempted colonization, known as a hypersensitive response (HR). In cereals, three different resistance mechanisms have been identified: (i) surface-localized receptors that recognize microbial-associated or pathogen-associated molecular patterns (PAMPs) and activate PAMP-triggered immunity; (ii) race-specific resistance conferred by intracellular immune receptors that recognize pathogen effector proteins and activate effector-triggered immunity; (iii) broad-spectrum quantitative (partial) disease resistance via various molecular pathways.

Although hundreds of disease resistance (*R*)-gene loci have been genetically mapped to wheat chromosomes, only a small number of yellow rust resistance genes (*Yr*) have been isolated so far. *Yr10, Yr18/Lr34/Sr57, Yr36, Yr46/Lr67/Sr55, Yr5/YrSP, and Yr7* have been cloned. Most of the cloned *R*-genes in wheat encode intracellular nucleotide-binding leucine-rich-repeat receptors (NLRs), which recognize pathogen-secreted effectors delivered into the host cytoplasm. Nevertheless, among the cloned *Yr* genes only three are NLRs (*Yr10, Yr5/YrSP, and Yr7*), while the others represent different protein families, which act through unique mechanisms and result in broad-spectrum resistance to yellow rust (*Yr36*), and even in multi-pathogen (partial) resistance against the three wheat rusts and powdery mildew (*Yr18/Lr34/Sr57/Pm38, Yr46/Lr67/Sr55/Pm39*). *Yr36* encodes a wheat kinase start 1 (WKS1) protein having a non-RD kinase domain fused to a START lipid-binding domain. WKS1 contributes to cell death, restricting pathogen growth and sporulation. However, this response takes several days longer than typical HR responses, thereby resulting in partial resistance. *Yr46/Lr67* is a hexose transporter involved in sugar uptake. Changes in two critical amino acids in the resistant allele of *Yr46/Lr67* probably alter the hexose transport in infected leaves and may explain the reduced growth of multiple biotrophic pathogen species. *Yr18/Lr34* encodes an ATP-binding cassette (ABC) transporter of a yet unknown substrate.

The genetic bottlenecks associated with wheat polyploidization and domestication, as well as selections in agroecosystems, led to a decrease in wheat genetic diversity and an increase in its vulnerability to biotic and abiotic stresses. One of the ways to improve wheat tolerance to these stresses is to recruit the adaptive potential of the wild wheat germplasm. Wild emmer wheat (*Triticum turgidum* ssp. *dicoccoides*), the tetraploid progenitor of hexaploid common wheat (*Triticum aestivum*), has valuable adaptive diversity to various diseases, including yellow rust.

*Yr15* is a yellow rust resistance gene discovered in the 1980s in WEW accession G25 (G25). *Yr15* confers broad-spectrum resistance against a worldwide collection of more than 3000 genetically diverse *Pst* isolates, including modern races, such as 'Warrior' (race DK09/11), which is currently threatening wheat production. Although some *Pst* races were reported to be virulent on *Yr15* in 2000s, since 2003 no virulence on *Yr15* was reported among new incoming *Pst* isolates (Global Rust Reference Center (GRRC)).

Previously, we had mapped *Yr15* on chromosome arm 1BS distal to the nucleolar organizer locus *Nor-B1* using restriction fragment length polymorphism and simple sequence repeat markers. The development of near isogenic lines (NILs) for *Yr15* in the common wheat Avocet background and the identification of DNA markers linked to this gene have accelerated the introgression of *Yr15* into a variety of pre-breeding and commercial materials, including cultivars, such as Clearwhite 515, Expresso, Patwin 515, and Seahawk, that are widely grown in western United States. In light of the above, *Yr15* will find use in wheat breeding in many other wheat-growing regions.

Here, we describe the identification of the *Yr15* gene sequence by map-based cloning and demonstrate its function in conferring resistance to yellow rust in durum and common wheats, by both ethyl methanesulfonate (EMS) mutagenesis of resistant *Yr15* introgression lines (ILs) and transgenic complementation of susceptible varieties. We show that *Yr15* encodes a protein composed of putative kinase and pseudokinase domains in tandem, which we name here as WTK1 (wheat tandem kinase 1). Orthologs, paralogs, and homologs of WTK1 were found in other cereal species. Furthermore, we discovered similar tandem kinase-pseudokinase (TKP) structures in a wide range of plant taxa and suggest that they belong to the same protein family, designated here as TKP family.

**Results**

**Variation in resistance response of *Yr15* ILs.** A set of *Yr15* ILs inoculated with *Pst* isolate #5006 showed a range of resistance responses, probably due to differences in their genetic backgrounds (Supplementary Fig. 1a). Infection types (*IT*) for these ILs ranged from 0 (no visible response, e.g., B70) to 2 (large chlorotic blotches, e.g., Sel46 (*Yr15*)), with clear HR visible to the naked eye in most of the *Yr15* ILs (Supplementary Fig. 1a).

Previous studies showed that although *Yr15* conferred broad-spectrum resistance against a worldwide collection of *Pst* isolates, a few isolates that appeared in Europe more than a decade ago (e.g., isolate DK92/02) showed virulence on *Yr15* IL. *Avocet* + *Yr15*2. Therefore, we have challenged a set of *Yr15*, *Yr5*, and *Yr5/Yr15* ILs with *Pst* races known to be virulent on *Yr15* (*DK92/02* or *Yr5*), and *Yr5* (*AU85569*). Indeed, *DK92/02* was virulent on *Avocet* + *Yr15* (*IT = 5*–*7*), and *AU85569* was virulent on *Avocet* + *Yr5* (*IT = 7*), while both of these races were virulent (IT = 8–9) on *Avocet* S NIL (Supplementary Fig. 1b). However, *Yr15* with *Yr5* in four different backgrounds (*YecoraRojo*, *Yr5*, *Patwin* Yr5Yr15, *Summit* Yr5Yr15, and *Dirkwin* Yr5Yr15) showed full protection against both virulent isolates (Supplementary Fig. 1b).

**Mechanism of resistance associated with HR.** Microscopic observations of the interactions between plant host cells and the invading pathogen showed the formation of fungal structures within leaf tissues, including substomatal vesicles, primary infection hyphae, haustoria mother cells (HMCs), and haustoria.
already at 1 day post inoculation (dpi) in both the susceptible common wheat genotype Avocet S and its resistant NIL Avocet + Yr15 (Supplementary Fig. 2a). Moreover, high-resolution close-up images presented in Fig. 1a–c clearly show HMCs, very thin penetrating hyphae (neckband) and haustoria feeding structures inside invaded mesophyll cells of Avocet + Yr15 at 3 dpi. Substantially larger fungal colonies and much more abundant pathogen feeding structures were observed in Avocet S compared with Avocet + Yr15 NIL, 3–14 dpi (Fig. 1d, e and Supplementary Fig. 2a). Relative quantification of fungal development and colonization of Avocet S (Fig. 1d) fitted the standard logistic population growth model well and displayed highly significant increase in biomass ($p < 10^{-15}$, likelihood ratio test), which is typical for the growth of microorganisms under favorable conditions\textsuperscript{23}. By contrast, colonization of Avocet + Yr15 was negligible and did not change significantly 1 to 14 dpi ($p = 0.21$, likelihood ratio test) (Fig. 1d). At 7 dpi fungal biomass residing within infected leaf tissues began to show significant differences between NILs ($p < 0.001$, $t$ test, $n = 16$) (Fig. 1d), when macroscopic evidence of HR became to be visible on Avocet + Yr15 leaves, but not on Avocet S (Fig. 1f and Supplementary Fig. 2b). The differences in fungal growth between the NILs can be attributed to the strong HR response in Avocet + Yr15, which restricts the fungal growth at very early stages of infection. The host cells that are interacting with fungal feeding structures show clear indications of HR, as evident from the collapse and distortion of these cells, strong autofluorescence as well as the aggregation of the chloroplasts within the infected cells (Fig. 1g and Supplementary Fig. 3a–d). These results suggest that HR plays a central role in restricting the development of fungal feeding structures.

**Map-based cloning of Yr15.** A recombinant inbred mapping population segregating for Yr15 was generated by crossing the susceptible recurrent durum line D447 with the resistant durum Yr15 ILs (B9, B10), into which Yr15 had been introgressed from G25\textsuperscript{20}. We screened 8573 F2 plants with DNA markers to develop a high-resolution map of the Yr15 gene region. First, we exploited...
collinearity between the Yr15 region and the syntenic regions in model organisms Brachypodium distachyon, Oryza sativa (rice), and Sorghum bicolor (Supplementary Fig. 1a–d and Supplementary Tables 1–4). The nearest collinearity-derived markers, uhw264 and uhw258, defined Yr15 within an interval of 0.3 cM. The target interval containing Yr15 was then saturated with 24 markers developed from either bacterial artificial chromosome (BAC) sequences positioned on the physical map of chromosome arm 1BS of T. aestivum cv. Chinese Spring (CS)24 (Supplementary Table 5), the genome assembly of WEW accession Zavitan (Zavitan)25 (Supplementary Fig. 4c–e and Supplementary Table 6), or G25 BAC sequences. Using these markers, we mapped Yr15 between markers uhw300 and uhw273 (0.013 cM) (Fig. 2a–c).

A pooled BAC library, constructed from Yr15 donor G25, was screened with the closely linked or co-segregating markers uhw297, uhw292, uhw296, and uhw273 (Fig. 2c), which were derived from the CS contig corresponding to the Yr15 region.

Fig. 2 Map-based cloning of Yr15. a, b Genomic sequence of CS (a) and Zavitan (b) 1BS pseudomolecules harboring the WTK1 region. c Genetic map of the 1BS region carrying Yr15. Marker uhw301 was developed from the WTK1 sequence. d G25 BAC-based physical map covering the Yr15 region. Assembly of BAC sequences yielded contigs G25ctg1 (287 kb) and G25ctg2 (131 kb). Marker BAC141RP2_1 was used to connect G25-141 and G25-40 BAC clones. e, f Structure of WTK1 alleles from G25 (e) and from CS and Zavitan (f). Features that are unique to CS are marked with superscript CS, while those that are present only in Zavitan are marked with superscript Za

Fig. 3 Validation of WTK1 as Yr15. a Loss of resistance to Pst in WTK1 kinase-like domain mutants. EMS4 and EMS6 carry mutations in KinI and KinII, respectively. b–c Wtk1 IF1 transcription levels (b) and Pst resistance phenotypes (c) in independent transgenic events Tyr15-K2 and Tyr15-F1. Error bars denoting standard error of mean (s.e.m.) are based on three technical replicates. d Wtk1 expression survey in root, stem, and leaf of B9 by RT-PCR. e Transcript levels of the Wtk1 IF1 in mock-inoculated or Pst-inoculated B9 plants. Error bars denoting s.e.m. are based on six biological replicates. Asterisks indicate the level of significance by t test, n = 12: p < 0.05 (*), p < 0.01 (**), p < 0.001 (***).
Sequencing of the six BACs isolated with these markers yielded two non-overlapping contigs (G25ctg1, 286,738 bp and G25ctg2, 131,485 bp; Fig. 2d). Sequence analysis of these contigs revealed the presence of three putative candidate genes. One of them contained predicted protein domains that previously have been associated with plant responses to pathogens and therefore was selected for validation. This gene contains two distinct kinase-like domains arranged in tandem and is designated here as WTK1 (Fig. 2e, f). Alignment of the WTK1 full-length complementary DNA (cDNA) against the WTK1 genomic sequence (4657 bp) indicated that the gene contains six exons coding for a 665 amino acid protein.

Validation of the candidate gene for Yr15. To validate the function of WTK1, we used EMS to mutagenize a set of Yr15 ILs and identified 2 out of 2112 tetraploid and 8 out of 1002 hexaploid M2 families that segregated for resistance to Pst (Supplementary Table 7). Sequencing of WTK1 in the 10 susceptible M2 plants confirmed the presence of independent missense mutations in each mutant. Five EMS mutants contained amino acid changes in the WTK1 kinase-like domain 1 (KinI) and five in kinase-like domain II (KinII) (Supplementary Table 7). Examples of these mutants (EMS4, having a mutation in KinI, and EMS6, with a mutation in KinII) are presented in Fig. 3a. In the F2 progenies derived from crosses between the resistant wild-type Avocet + Yr15 and two of the susceptible mutants (EMS4 and EMS6), homozygosity for the non-functional WTK1 mutations, designated as wtk1, co-segregated with susceptibility to Pst (Supplementary Table 8). These 10 independent mutations demonstrate that both KinI and KinII domains of WTK1 are necessary for the resistance conferred by Yr15.

To determine whether WTK1 is sufficient to confer resistance to Pst, we transformed the susceptible varieties T. aestivum cv. Fielder and Triticum durum cv. Kronos with a 9.1 kb genomic fragment harboring the complete WTK1 coding region and its regulatory elements. Expression of the WTK1 transgene was detected in 17 out of 22 T0 transgenic plants and their progenies. In subsequent generations when challenged with Pst (race PST-130 and isolate #5006), the WTK1 transgene co-segregated with the yellow rust resistance phenotype (Fig. 3b, c and Supplementary Fig. 5a–g), thereby validating that WTK1 is indeed Yr15.

Expression of Yr15 alternative transcript variants. Three alternative transcript variants—isoform 1 (IF1), isoform 2 (IF2), and isoform 3 (IF3)—were revealed by sequencing of 48 WTK1 cDNAs (Supplementary Fig. 6a). IF1, which encodes the complete WTK1 protein, was found to be the most common transcript (85.7% of tested WTK1 cDNAs), while the other isoforms are much less abundant (14.3% together) and encode truncated proteins. WTK1 is expressed in roots, stems, and leaves (Fig. 3d). Transcript levels of the three WTK1 isoforms were significantly lower in leaves of IL B9 inoculated with Pst, at all dpi (except 0 dpi), relative to mock-inoculated plants (Fig. 3e and Supplementary Fig. 6b). In addition, a significant reduction in expression was observed over time (1–9 dpi) for both Pst-inoculated (p < 8.0 × 10−8, F test, df1,df2 = 1,30) and mock-inoculated (p < 2.4 × 10−5, F test, df1,df2 = 1,30) plants for the WTK1 IF1 (Fig. 3e and Supplementary Fig. 6b). These findings suggest that expression of WTK1 is down-regulated by the presence of the pathogen.

Changes in reading frames of non-functional alleles of Yr15. A search for WTK1 sequences on chromosome arm 1BS of the CS and Zavitan, relative to G25 (Fig. 2e, f). In addition, orthologs of Yr15 were found also on chromosome arms 1AS and 1DS. The predicted protein encoded by the ortholog on 1A of CS shares 83% similarity with WTK1 on 1BS of G25, but is 21 amino acids longer due to intron retentions. In contrast, the predicted proteins encoded by the 1A ortholog in Zavitan and the 1D ortholog in CS are shorter than the G25 1BS WTK1 due to framen shifts resulting in stop codons; their sequence similarities to G25 1BS WTK1 are relatively low (39.9 and 52.7%, respectively).

Phylogeny of orthologs, paralogs, and homologs of WTK1. In addition to the orthologs of WTK1 that were found on all homoeologous group 1 chromosomes of WEW (AABB) and common (AABBD) wheats, we found WTK1 orthologs also in diploid wheats relative Triticum urartu (AA), Aegilops speltoides (SS), and Aegilops tauschii (DD), representing the ancestral A, B, and D genomes, respectively, and on chromosome 1R of rye (Secale cereale, RR). A phylogenetic analysis clustered all the predicted proteins into a single branch, consisting of sub-clusters corresponding to the 1A, 1B, and 1D copies (Supplementary Fig. 7). No WTK1 ortholog was detected on chromosome 1H of barley (Hordeum vulgare, HH). Paralogs of WTK1 were found on the homoeologous group 6 chromosomes of tetraploid and hexaploid wheat, their diploid relatives T. urartu and A. tauschii, and on chromosome 6H of barley. All group 6 copies clustered together phylogenetically. A homologous copy from chromosome Bd2 of B. distachyon showed an intermediate position between the two clusters, while the closest predicted protein from O. sativa was resolved as an outgroup. The presence of WTK1 paralogs on both group 1 and group 6 chromosomes of almost all diploid wheat species, as well as the relatively low nucleotide sequence identity between group 1 and group 6 paralogs of WTK1 (~80% in KinI and ~30% in KinII), suggests an old duplication event.

Distribution of Wtk1 allele in wild and cultivated wheats. We have evaluated 390 Triticum accessions for the presence of Wtk1 by gene-specific markers (Supplementary Tables 9–12). Wtk1 was detected in 18% of the accessions of the southern (Israel, Lebanon, Jordan, and Syria) WEW populations but was not detected in the northern (Turkey and Iran) ones (Supplementary Table 10). We screened 93 modern durum and common wheat accessions and found that the functional Wtk1 allele was present only in recently developed Yr15 ILs (Supplementary Table 12).

WTK1 TKP architecture. Both WTK1 domains belong to a protein kinase-like superfamily with a conserved catalytic domain, but contain neither membrane-targeting motifs nor known receptor sequences. The lack of transmembrane domains in WTK1 is consistent with its cytoplasmic localization (with some partition into the nucleus) in barley protoplasts transformed with a WTK1-green fluorescent protein (GFP) fusion (Supplementary Fig. 8a–i). The presence of key conserved residues in KinI and the lack of most of these residues in KinII implies that WTK1 possesses a putative kinase-pseudokinase structure (Supplementary Fig. 9). Moreover, KinI belongs to serine/threonine non-RD kinases, shown previously to be associated with immune receptors. WKS1 is another Pst resistance gene from WEW that harbors a non-RD kinase domain. However, it is attached to a lipid-binding START domain and does not exhibit high similarity to either domains of WTK1 (Supplementary Fig. 10). Only two plant resistance genes have a tandem
kinase structure similar to WTK1: barley RPG1, conferring resistance to stem rust\(^{30}\), and MLOC\_38442.1, which was proposed as a candidate for the barley true loose smut resistance gene \(Un^{8}\). Nevertheless, phylogenetic analysis shows that they are quite distant from WTK1 (Supplementary Fig. 10, 11a).

**The TKP protein family is found across the plant kingdom.**

The occurrence of functional resistance genes with a tandem kinase structure (WTKI and RPG1\(^{30}\)), as well as of many WTK1 orthologs and paralogs in wheat and its near relatives, motivated us to search for similar protein architectures across the plant kingdom. Altogether, we found 92 predicted proteins that are composed of putative kinase and pseudokinase domains in tandem; like WTK1, none had additional conserved domains. Most of the putative kinase domains share key conserved residues\(^{28}\) (Fig. 4a, b and Supplementary Table 13), while the putative pseudokinase domains are generally highly divergent in these positions, suggesting that they probably have no or impaired kinase activity\(^{32}\) (Fig. 4a and Supplementary Table 14). The least conserved residues in the putative pseudokinase domains were found in the catalytic loop (D\(^{166}\) and D\(^{184}\)) and the glycine-rich phosphate-binding loop (G\(^{28}\))\(^{33}\) (Fig. 4a). Analyses of motifs shared between putative kinase domains of TKP family members revealed additional conserved residues in sequences neighboring the core motifs (Fig. 4b). These results indicate that the putative kinase domains in this protein family may share a common structure.

The phylogenetic analysis showed that all 184 putative kinase and pseudokinase domains of the 92 predicted proteins could be sorted into 11 major clades and two singletons (Supplementary Fig. 11a). Homology-based comparisons of the individual kinase domains of the 92 TKPs with the annotated *Arabidopsis* kinome families\(^{34}\) revealed clear relationships between the 11 clades and specific *Arabidopsis* families or subfamilies (Supplementary Fig. 11a). Members of clades 1–10 and cluster 11.2 were found to correspond to five families of plant receptor-like kinases (RLKs): concanavalin A-like lectin protein kinases (L-LPK), leucine-rich-repeat-repeat receptor kinases (LRR), receptor-like cytoplasmic kinases (RLCK), other kinases with no published family (RK), and cell-wall-associated kinases (WAKs). Members of cluster 11.1 were not classified as RLKs. Instead, they were found to be associated with the following families of soluble kinases: cATP-dependent, cGTP-dependent, and phospholipid-dependent kinases, cyclin-dependent kinases, raf-like MAPK (mitogen-activated protein kinase) kinase kinase (MAP3K-rat), MAPK kinase kinase (MAP3K), and SNF1-related kinase 3. All 92 TKPs can be divided into two distinct groups based on the topology of the phylogenetic tree: (i) two kinase domains of the same gene that are clustered together in the same branch most likely resulted from a duplication event; (ii) two kinase domains belonging to the same gene but are positioned on separate branches represent a gene fusion event (Supplementary Fig. 11a). While four clades (2, 5, 7, and 11) included proteins that, by this criterion, resulted from either gene duplications or fusions, three clades (6, 9, and 10) included proteins that originated only from duplication events, and four clades (1, 3, 4, and 8) included proteins that resulted only from fusion events. A total of 52 TKPs were derived from duplication (occurring in kinase domains associated with 10 different families or subfamilies) and the other 40 proteins resulted from gene fusions (represented by 10 combinations of domains belonging to different families or subfamilies; Supplementary Fig. 11b). The most common domains present in TKPs were similar to WAK kinases (63 out of 184) and LRR\_8B (cysteine-rich kinases; 49 out of 184). WTK1 appears to be
derived from a fusion of WAK and RLCK_8 kinase domains. Interestingly, most fusion events were found in monocot TKPs. Notably, clade 6 includes proteins of distant species and was presumably originated from a single ancient duplication event of a kinase in the L-LPK family before the divergence of the monocots and eudicots.

Discussion

Although new R-genes are frequently discovered from various sources, the rapid evolution of Pst populations overcomes most of the commonly used Yr genes, rendering most of the world wheat production susceptible to devastating epidemics. The combination of marker-assisted selection and classical approaches can speed up the progressive incorporation of individual new R-genes into elite wheat varieties. However, it seems that the appearance of new virulent Pst races in the fields is faster than the incorporation of novel R-genes into elite wheat varieties.

Yr15 and Yr5 showed broad-spectrum resistance to Pst races from around the globe and are among the most promising genes for protecting wheat yield. Nevertheless, virulent races were detected against each one of them, separately, in Australia for Yr522 (isolate AU85569) and in Europe for Yr1519 (e.g., isolate DK92/02). Therefore, we were interested to know whether the pyramiding of Yr5 and Yr15 into one cultivar could provide a more durable resistance. Our results indeed show that Yr5 and Yr15 complement each other and provide full resistance to the virulent races. Pyramiding resistance genes by classical breeding methods may take many years. However, approaches based on genetic transformation of individual cloned R-genes or of R-cassettes that include several R-genes with complementary resistances to various races, preferably via different mechanisms, can be considerably faster and can also overcome the problem of individual genes segregating in backcross populations. Yr5, Yr15, Yr16, Yr36, and Yr46 are obvious candidates to be included in such a cassette, since they represent different gene architectures (NLK, kinase-pseudokinase, ABC transporter, Kinase-START, and hexose transporter, respectively) and most likely confer resistance via different molecular mechanisms.

The comparison of fungal development in susceptible vs. resistant NILs presented here provides indications for the possible mechanism of resistance conferred by Yr15. The macroscopic and microscopic observations, together with the quantification of fungal biomass in the resistant genotype that carries Yr15, show clear development of HR after fungal penetration, establishment of fungal feeding structures within the plant cells and restriction of fungal growth. The microscopic observations showed that a strong PCD response took place as part of the resistance mechanism, resulting in localized cell death that arrested further colonization of host cells by Pst. We showed that the size of chlorotic spots as a macroscopic indication of HR varied between the different genetic backgrounds of the Yr15 ILs. The HR/PCD response can be activated via different cellular pathways; further experiments are needed to identify the proteins interacting with WTK1 and the exact pathway that activates HR, in order to provide a better understanding of the mechanism of action and the host–parasite co-evolutionary processes behind it.

Two out of the six Yr genes cloned in wheat are derived from WEW (Wks1 and WTK1) and provide broad-spectrum resistance against Pst. Both are present only in the southern distribution range of WEW populations in the Fertile Crescent (Israel, Lebanon, Jordan, and Syria) and were not detected in the northern populations (Turkey and Iran). Since wheat was probably domesticated from wild emmer in southeast Turkey, both Yr15 and Yr36 have, until recently, not found their way into the genepool of cultivated wheat germplasm. While the WKS1 sequence is completely absent from all three genomes of hexaploid wheat, in the current study we have shown that non-functional orthologs and paralogs of WTK1 are present on all chromosomes of groups 1 and 6 of the tetraploid (AABB) and hexaploid (AABBDD) wheat genomes. Despite the relatively low similarity of orthologous and paralogous copies of WTK1 proteins in wheat that are associated with the susceptibility phenotype (Zavitan and CS), they still may carry resistance to some pathogens in other genetic backgrounds (such as diploid species) or may have other functional alleles in natural WEW populations. Furthermore, the presence of orthologs, paralogs, and homologs of WTK1 in all ancestral diploid genomes of wheat, and even in the basal Pooid Brachypodium, indicates that WAK-RLCK kinase-pseudokinase fusion represents an old event that occurred prior to the divergence of the wheat and Brachypodium lineages, about 35–40 million years ago. In contrast, the kinase-START protein architecture of WKS1 has not been found in other species outside of the Triticaceae tribe and is likely a recent fusion of existing kinase and START domains. For current breeding purposes, both Yr15 and Yr36 constitute un tapped resources to improve resistance to Pst in domesticated wheat, since they are absent in all tested modern durum and common wheat varieties, except for recent introgressions.

The kinase protein superfamily represents 1–2% of the functional genes across plant and animal genomes, as reflected by the wide involvement of kinases in almost all cellular processes. The RLK clade is relatively expanded in flowering plants and plays an important role in plant immune responses. The current phylogenetic analysis of the TKP protein family revealed that 175 out of 184 kinase domains were associated with RLKs, suggesting that TKPs are involved in plant defense mechanisms. Moreover, broad-spectrum resistance to biotrophic fungal pathogens (e.g., rusts and smut) was already demonstrated for WTK1 and RPG1 and proposed for Un8, raising the question of whether the widespread TKPs serve generally as a family of resistance proteins. Most RLKs in Arabidopsis share multi-domain architecture with receptor, transmembrane, and kinase domains, but 24% of them contain no extracellular or transmembrane domains, which are likewise lacking from TKP family members. Moreover, both RPG1 and WTK1 are localized mainly in the cytoplasm.

The evolution of multi-domain proteins, such as TKPs, presumably monophyletic or polyphyletic origins, respectively, from single or multiple independent evolutionary events that can be explained by single or multiple birth models for protein families. Our phylogenetic analysis indicates that TKP family members originated from either gene duplication or gene fusion, suggesting a polyphyletic origin of the TKPs. Moreover, even divergent members of the TKP family, such as RPG1 and WTK1, which evolved, respectively, by duplication and fusion, share not only similar gene architecture but also analogous function (e.g., resistance to pathogens). Most of the TKP proteins (52 out of the considered 92) were formed by tandem duplication, earlier shown to be more frequent than fusion during the evolution of multi-domain proteins. In addition, we have seen indications on lineage-specific origin of some TKP members, that is, domains related to WAK family that were found only in monocots.

After gene duplication or fusion, one of the copies can gain a new function, a process known as neofunctionalization. Interestingly, all members of the TKP family discussed here possess putative kinase-pseudokinase domain architecture and probably fit this neofunctionalization model. This structure may be related to the function of TKPs since some pseudokinases were shown to play an important role as baits for pathogen effectors. Moreover, 20% of the RLK family members are pseudokinases, a higher
percentage than the average in the Arabidopsis kinase (13%)36. The WTK1 pseudokinase belongs to the RCLK family. Members of this family were shown to serve as targets for inhibition by pathogen effectors (e.g., BIK1 in Arabidopsis)43. Moreover, the duplication event itself was suggested as a source for the decay model44. Thus, the decay role can be proposed as one of the potential mechanisms of function of the TKP family members in immune response. However, further studies are required in phylogenetically representative plant species in order to elucidate the mechanism of resistance conferred by this unique protein family.

Furthermore, members of the animal Janus kinase (JAK) family contain a similar structure of kinase and pseudokinase domains, a unique feature of JAKs compared to other animal protein kinases45. The participation of members of both JAK46 and TKP (e.g., WTK1) families in innate immune responses, particularly in PCD, calls attention to not only structural but also possibly functional similarities between these proteins, suggesting convergent molecular evolution of proteins involved in immunity in both the plant and the animal kingdom.

Methods

Mapping populations and yellow rust assays: The yellow rust resistance gene Yr15 was introgressed from the resistant WES accession G25 (G25) into the susceptible durum wheat (T. turgidum ssp. durum) accession D447 (LD939/2*Langdon ND58-322) to develop the resistant BC3F9 (B9) or BC3F10 (B10) ILs21. A large mapping population, consisting of 8573 F2 plants, was developed by crossing D447 with these ILs (B9 or B10), which carry Yr15 within a segment of chromosome 1BS and 1BS. Phenotyping of recombinant, mutated, and transgenic lines under growth chamber conditions was carried out with Israeli Pst isolate #5006 (race 3BE134) and US Pst race PST-130 using a standard protocol23. Plants were inoculated either at the two-leaf to four-leaf stage ("seedling inoculation") or at the stem elongation stage ("adult plant inoculation"). The yellow rust response variation was evaluated 14 to 18 dpi using a 0 to 9 scale of 

Microscopy of Pst-wheat interactions within infected leaves. Fluorescence microscopy of P. striiformis was performed using wheat germ agglutinin (WGA; a lectin that binds specifically to 0 (1-4)-N-acetyl-g-glucosamine, i.e., chitin) conjugated with a fluorescent dye46. Leaf segments (second leaf, 10 cm long) from the NILs, Avocet, S, and Avocet × Yr15, each of which had been inoculated with ureidiospores of Pst isolate #5006, were sampled from 1 until 11 dpi (10:00 to 10:30 a.m., every 2 days), when visible sporulation developed on the susceptible plants at each time point, with eight biological replicates collected for each line at each time point. Leaf samples were fixed in ethanol:chloroform (3:1, v/v) + 0.15% (w/v) trichloroacetic acid solution for at least 24 h. Segments were then washed twice in 50% ethanol (10 min) followed by clearing in 0.5% M NaOH (30 min). After washing with deionized water (DI) and 0.1 M Tris-HCl buffer (pH 5.8), samples were stained for 10 min in Tris-HCl buffer containing 0.1% Uvitex 2B (Polysciences Inc.) (w/v). Specimens were washed four times in DI followed by one wash in 25% glycerol and were left overnight in DI to remove access stain. Specimens were stored in 50% glycerol until further use. The microscopic investigations of the leaf samples were mounted on glass slides in 75% glycerol. Fungal colonies and host cell autofluorescence were analyzed using an Olympus FV1200 confocal laser scanning microscope. Fungal structures and leaf tissues were excited with 405 and 515 nm lasers, and the emitted light was detected in two channels with filter settings 486–530 and 554–654 nm, respectively. Z-stacks were collected with 1 µm separation. The images from the two detection channels were merged and 3D projections were performed with IMARIS’ bitplane 6.2. The resulting 2D images were adjusted for brightness and contrast. This included small adjustments in individual color channels to obtain an equal representation of the background signal from the healthy plant tissue on all images. All adjustments were applied equally across a whole image using the free software paint.net (https://www.getpaint.net/).

Development of a high-density genetic map. We developed Yr15 to deletion bin Sat0.31 of chromosome arm 1BS by developing and mapping dominant marker markers uw250, as well as cleaved amplified polymorphic sequences markers uw252 and uw254, based on expressed sequence tags assigned to Sat0.31 (Supplementary Table 2). Using the GenoMapZapper approach45, which identifies JAKs with high similarity to their orthologous regions of B. distachyon, O. sativa, and S. bicolor (Supplementary Table 3), eight additional markers were developed and the Yr15-containing region was reduced to a 0.3 CM interval from flanking markers uw264 and uw259 (Supplementary Fig. 4a).

All-nucleotide polymorphism (SNP) markers RAC875_c826_839 and BSO0022902_51 derived from the 15K wheat SNP array (Trait Genetics GmbH, Gatersleben, Germany) were converted to Kompetitive Allele-Specific PCR (KASP) markers using PolyMarker (http://polymarker.igac.ac.cn/) (Supplementary Table 4).

The screening of F2 plants from the D447 × B9 and D447 × B10 crosses for recombination events in the region carrying the Yr15 gene was carried out with the following marker sets (Supplementary Fig. 4b): (i) wmc060 and gwmc273 (6); (ii) RAC875_c826_839 and BSO0022902_51; (iii) uw264 and uw259. A total of 94 non-homologous recombinant inbred lines (RILs), corresponding to 13 independent recombinant events, were developed for use for the cloning of Yr15 for chromosome walking. These RILs were evaluated for resistance to isolate #5006.

Physical maps of the Yr15 region on chromosome arm 1BS. Markers uw264 and uw259, flanking Yr15, were used to screen a gridded CS BAC library53. Three positive BAC clones (TaaCsp364011, TaaCsp103282, and TaaCsp115820) were detected with marker uw264 and two (TaaCsp729H14 and TaaCsp141G12) with marker uw259 (Supplementary Fig. 4b-d). The BAC-end sequence (BES) of clone TaaCsp103282 was used to develop marker uw267. Screening of the 15K wheat BAC library with uw267 yielded clone TaaCsp99177. The BES of clone TaaCsp729H14 was used to develop marker uw268, which was mapped distal to uw259 (Supplementary Fig. 4c). These efforts generated two BAC contigs, flanking Yr15, with a gap between them, which was bridged using a BES physical map constructed from a CS BAC library53. The 4647 BACs selected from 49,412 BACs were converted to Kompetitive Allele-Specific PCR (KASP) markers using PolyMarker (http://polymarker.igac.ac.cn/) (Supplementary Table 4).

Screening of the F2 plants from the D447 × B9 and D447 × B10 crosses for recombination events in the region carrying the Yr15 gene was carried out with the following marker sets (Supplementary Fig. 4b): (i) wmc060 and gwmc273 (6); (ii) RAC875_c826_839 and BSO0022902_51; (iii) uw264 and uw259. A total of 94 non-homologous recombinant inbred lines (RILs), corresponding to 13 independent recombinant events, were developed for use for the cloning of Yr15 for chromosome walking. These RILs were evaluated for resistance to isolate #5006.

Quantification of fungal biomass within infected leaves. After inoculation with isolate #5006, leaf segments (middle, second leaf) of the resistant Avocet + Yr15 and susceptible Avocet × Yr15 NILs were cut every 2 days (1–14 dpi) from different plants at each time point, with eight biological replicates collected for each line at each time point, and used for quantification of fungal biomass by qPCR measurement46. Three technical replicates were made for each tissue sample. Fluorescence was measured on a SpectraMax M2e Microplate Reader (Molecular Devices, Sunnyvale, CA, USA), using 485-nm excitation and 535-nm emission wavelengths, a 1.0 s measurement time, and a cross pattern of well-scanning, yielding an average measurement per well. Statistical analyses of parameters associated with the accumulation of fungal biomass were determined by the maximum likelihood estimation method (SPSS) following the logistic population growth model. The null hypothesis (no growth) was tested by the likelihood ratio test.

Confocal microscopy. The second leaf of 16-day-old plants of the susceptible Avocet S and resistant Avocet × Yr15 NILs were point inoculated with fresh spores of the Pst isolate UK75/30. One leaf segment per plant was sampled at 4 and 8 dpi. Leaf samples were fixed in ethanol:chloroform (3:1, v/v) + 0.15% (w/v) trichloroacetic acid solution for at least 24 h. Segments were then washed twice in 50% ethanol (10 min) followed by clearing in 0.5% M NaOH (30 min). After washing with deionized water (DI) and 0.1 M Tris-HCl buffer (pH 5.8), samples were stained for 10 min in Tris-HCl buffer containing 0.1% Uvitex 2B (Polysciences Inc.) (w/v). Specimens were washed four times in DI followed by one wash in 25% glycerol and were left overnight in DI to remove access stain. Specimens were stored in 50% glycerol until further use. The microscopic investigations of the leaf samples were mounted on glass slides in 75% glycerol. Fungal colonies and host cell autofluorescence were analyzed using an Olympus FV1200 confocal laser scanning microscope. Fungal structures and leaf tissues were excited with 405 and 515 nm lasers, and the emitted light was detected in two channels with filter settings 486–530 and 554–654 nm, respectively. Z-stacks were collected with 1 µm separation. The images from the two detection channels were merged and 3D projections were performed with IMARIS’ bitplane 6.2. The resulting 2D images were adjusted for brightness and contrast. This included small adjustments in individual color channels to obtain an equal representation of the background signal from the healthy plant tissue on all images. All adjustments were applied equally across a whole image using the free software paint.net (https://www.getpaint.net/).
proximal markers uwh288, uwh289, uwh287, uwh285, uwh286, and uwh281 (Supplementary Table 5). Further screening of the G25 RAC library with the closest distal markers uwh297 and uwh292, the co-segregating marker uwh296, and the proximal marker uwh273 yielded six BAC clones (G25–33, G25–86, G25–40, G25–141, G25–35, and G25–25; Fig. 2). DNA samples of these BACs were extracted using a Qiagen Plasmid Midi Kit or Qiagen Large-Construct Kit (Qiagen, Hilden, Germany). Contiguous sequences were generated on a Pabos R5 II (Pacific Biosciences, Menlo Park, CA, USA) at the Institute of Biotechnology, University of Helsinki (Helsinki, Finland). BAC clones were sequenced and assembled (HGA3 implemented in SMRT portal 2.3) separately. The GAP4 program (Staden package) was used to edit and join the assembled BAC contigs into two contigs (G25c26, 27 and G25c13, 17, and G25c26, 17 bp upstream of G25r1 gene region; Fig. 2). One Yr15 distal marker (uwh300) and two co-segregating markers (uwh302 and uwh301) were developed by comparing of G25 sequences to the 1BS pseudomolecule of Zavitan25 (Supplementary Table 6).

G25 contig annotation and identification of candidate genes. Repetitive ele-
ments were masked using the TRitez Repeat Sequence Database (TREPT, http://
botvers2.uzh.ch/keldatta/trep-db/index.html). Non-repetitive sequences were ana-
yzed for genes by BLASTN search against the GenBank (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and the TIGR Wheat Genome (http://tigrblast.tigr.org/euk-blast/index.cgi?project=taei) databases, using GenScan (http://genes.mit.edu/
genescan.html) and FGENESH (http://softberry.com/berry.phtml). A
BLASTN search was carried out with the G25c26 and G25c26 contigs against the high-confidence (HC) gene models of the 1BS pseudomolecule of Zavitan25 was used to reveal the presence of putative HC candidate genes.

Isolation and sequencing of the full-length WTK1 cDNA. Total RNA was isolated with the Trizol reagent (Invitrogen, Thermofisher Scientific, Waltham, MA, USA). Poly(A)+ RNA was purified from the total RNA with a Qigioxet
mRNA Midi Kit (Qiagen). First-strand cDNA was synthesized with SuperScript II
(Invitrogen) using primer E1820 (Supplementary Table 9). Nested PCR was carried out first with E1820 and the 5′-untranslated region (UTR) primer Y15F2 and then with E2146 (matching part of E1820; Supplementary Table 9). The PCR products were purified, cloned, and sequenced.

To determine the start site of the WTK1 transcript, we used 5′ rapid amplification of cDNA ends (RACE)26. In brief, PCR (Poly(A)+ RNA was treated with tobacco acid pyrophosphatase (Epicentre, Madison, WI, USA), and then amplified from genomic DNA with PstI and SacI. The PCR products were puri-
digested by the restriction enzyme pairs SpeI–XhoI and XbaI–MscI and then transferred to the endogenous control. The genotype of the 15 independent T0 plants in Kronos (Tyr15-K1 to Tyr15-K15) and 7 T0 plants in Fielder (Tyr15-F1 to Tyr15-F7) were obtained. Three primer pairs (HpyFf1R1, R1, Tyr15F1/F1R1, and Y15K1F2/Y15S2P2; Supplementary Table 9) were used to validate the presence of WTK1 in the transgenic plants. In addition, we extracted GRNA from all T0 plants and estimated the transcript levels of the three WTK1 isoforms (isoform 1, IF1; isoform 2, IF2; isoform 3, IF3) by quantitative real-time PCR (qRT-PCR) using Ubiquitin as the endogenous control.

We germinated 10–25 T1 seeds from each transgenic event that expressed WTK1 and inoculated the plants with Pst race PST-130, which is virulent on Kronos and Fielder. All 12 positive T1 families in Kronos and the 5 positive T1 families in Fielder showed resistance to this Pst race. We genotyped all tested T1 plants and confirmed co-segregation of the resistance with the presence of the transgene. Furthermore, 5 to 10 independent plants from each of the T2 families from transgenic events Tyr15-F1, Tyr15-F4, Tyr15-F5, Tyr15-F6, Tyr15-K7, Tyr15-K8, Tyr15-K10, Tyr15-K12, and Tyr15-K14 were also tested with Pst isolate #5006, which is likewise virulent on Kronos and Fielder.

Gene expression analysis by quantitative RT-PCR. Expression analysis was conducted with total RNA isolated from various plant tissues (leaves, roots, and stems) using a RNeasy Plant Mini Kit (Qiagen). First-strand cDNA was generated using the qscript TM Flex cDNA Synthesis Kit (Quanta Biosciences, Beverly, MA, USA). The qRT-PCR was performed on a StepOne Plus Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) using the following program: 95 °C for 20 s; 40 cycles of 95 °C for 3 s and 60 °C for 30 s. The qRT-PCR reaction mixture contained the following components in a total volume of 10 µl: 5 µl Fast SYBR green master mix (Applied Biosystems); 2.5 µl diluted cDNA; 300 nM of each primer (Table 1); and 7.5 µl of a mixture of 0.25 mM of each of the five forward primers of the duplicated fivefold dilutions (1.1, 1.5, 1.25, and 1.125) in triplicates. Amplitudes of efficiencies of all primers were higher than 95%. Transcript levels are expressed as linearized fold-Ubiquitin levels calculated by the formula 2^(-ΔΔCT) ± standard error of the mean (SEM).

Assessment of alternative splicing of WTK1 mRNA. Alternative splicing var-
iants (IF1, IF2, IF3; Supplementary Fig. 6a) were revealed by sequencing of 48 B9 cDNA clones, which were amplified with a poly(T) primer and WTK1-specific
primers, or with WTK1-specific primers alone. To determine the relative transcript levels of these three splicing variants, we designed PCR primers for each WTK1 isoform (Supplementary Table 9). Leaves of B9 were collected after the following treatments: (i) before inoculation (0 h), as the control; (ii) after inoculation with spores of Pst isolate #5006 suspended in Solotrol® 170 light oil; (iii) after sprayer with Solotrol® 170 light oil lacking Pst spores, as a mock control. Samples of treatments (ii) and (iii) were collected at 1, 3, 5, 7, and 9 dpi, with six biological replicates for each treatment at each time point. All data were subjected to sta-
tistical analysis using the general linear model in SPSS.

Structure of wtk1 from CS and Zavitan. A search for WTK1 sequences in the chromosomal arm 1BS of CS28 and in the Zavitan25 whole-genome assemblies revealed the presence of non-functional alleles in both of these susceptible lines. The presence of three TEs: (i) a non-autonomous TRIM LTR retrotransposon of the Veju family (RLX_Taes_Veju) in intron 1 of CS and Zavitan; (ii) a DNA transposon of the DTH_Harbinger family (RLX_Taes_Harbinger) in intron 3 of CS and Zavitan; (iii) a non-autonomous DNA transposon of the DTM_MITE type (RP1764-1) in intron 3 of CS, was identified by a search in two TE databases—mipsRE-
data_9.3p_Poacea_TEs and trep-db_complete_Rel-16.

Candidate gene validation by transgenic complementation. Agrobacterium

tumefaciens-mediated transformation of susceptible durum and common wheat varieties, respectively Kronos and Fielder, served to further verify WTK1 function by complementation. Phusion® High-Fidelity DNA Polymerase (New England BioLabs, Ipswich, MA, USA) was used to amplify the WTK1 genomic region from Durum Wheat Genome Database (Morex pseudomolecules)29, respectively, to primers Yr15F1/R1 and Yr15F2/R2 (Supplementary Table 9) to enable cloning into the pLC41Hm transformation vector. Two overlapping PCR products were digested by the restriction enzyme pairs Xhol–Msci and Msci–AvrII, purified from agarose gel bands, and cloned into Mol-Sol linearized pLC41Hm. A 9116-bp genomic fragment that contains the full-length WTK1 coding region, as well as 3248 bp upstream of the start codon and 1031 bp downstream of the stop codon, was cloned for transformation. Sanger sequencing confirmed the conformation of the construct. Agrobacterium tumefaciens-mediated transformation26 of Kronos and Common wheat with this construct was carried out at the University of California Plant Transformation Facility (Davis, CA, USA) (http://ucdptf.ucdavis.edu/).

In total, 15 independent T1 plants in Kronos (Tyr15-K1 to Tyr15-K15) and 7 T0 plants in Fielder (Tyr15-F1 to Tyr15-F7) were obtained. Three primer pairs (HpyFf1R1, R1, Tyr15F1/F1R1, and Y15K1F2/Y15S2P2; Supplementary Table 9) were developed by comparing of G25 sequences to the 1BS pseudomolecule of Zavitan25 (Supplementary Table 6).
proline sequences from nine different species. Two of these proteins were used for further phylogenetic analysis (R. distichyum BRAD12G38370.1 and O. sativa LOC_Os01g20880.1). The analysis included 21 protein sequences. All sequence positions that contained gaps, as compared with WTK1-1B from G25, were eliminated from the analysis. The procedure was conducted using 26 amino acid residues out of the 665 of the full-length WTK1-1B from the DNA sequence of clone G25-141. The evolutionary history of WTK1 was inferred using the neighbor-joining algorithm based on evolutionary distances computed by the Poisson correction method. The quality of the derived phylogeny of WTK1 was assayed using a bootstrap test with 10,000 replicates. The analysis was performed with MEGA7 (https://www.megasoftware.net/).

### Distribution of WTK1 among various Tristece species

A set of two gene-specific, diagnostic markers, one for Kin1 and one for Kin2 (Supplementary Table 9), was developed and then used to test the distribution of WTK1 among various Tristece species. These dominant PCR markers were designed to identify the presence of the WTK1 functional allele, amplifying a 1 kb PCR product for Kin1 and 2 kb for Kin2.

### Subcellular localization of the WTK1 protein

Barley protoplasts (8 x 10^5) were isolated from H. vulgare cv. Boni plants and electroporated with 20–40 μg DNA at 300 V/cm essentially as described before, diluted into Gamborg B-5 Basal Medium with minimal organic acids that contained 10% glucose, and then transferred into glass-bottomed 35 mm microwell dishes (MatTek Corporation, Ashland, MA, USA) for microscopy. Electroporated protoplasts were cultured overnight at 22 °C before imaging under an inverted confocal laser scanning microscope (Leica TCS SP5 II; Leica Microsystems, Wetzlar, Germany) with a ×63 water immersion objective. The following light ranges were used: GFP (Ex 488 nm, Em 500–543 nm); 4',6-diamidino-2-phenylindole (Ex 405 nm, Em 430–550 nm); chlorophyll autofluorescence (Ex 488 nm, Em 673–725 nm). A WTK1 cDNA clone was prepared as described under isolation and sequencing. All expression vectors were created with the Multisite Gateway Three-Fragment Vector Construction Kit (Invitrogen). The 5’-end of an entry clone was prepared by the BP reaction between pDONR221 and a gene fragment amplified from pB221 (Genbank accession #A5021281) using primers attB4F35S and attB1R35S (Supplementary Table 9). The entry clone for the full-length WTK1 was created by carrying out the BP reaction between pDONR221 and the PCR product amplified from the WTK1 cDNA clone with primers attB1FKinase and attB2Ry15 (Supplementary Table 9). The entry clone for the N-terminal and C-terminal kinase-like domains of the kinase was similarly prepared using primers attB1FKinase and attB2RyKinaise and attB1Freg and attB2Ry15 (Supplementary Table 9), respectively. The 3’-end entry clone was prepared by carrying out the BP reaction with pDONRP2R3 and a GFPnos fragment amplified from pVEC8_GFP (GenBank: FJ949107.1) by PCR using primers attB2FGFPnos and attB3RGFPnos (Supplementary Table 9). The final clones used for transient transformation and expression in protoplasts were created by the LR reaction between the respective entry clones and the destination vector pDEST4-R3 according to the instructions of the Gateway Kit.

### Analysis of WTK1 kinase-like domains

A BLASTP search of the NCBI non-redundant protein database was used to assign the WTK1 kinase-like domains to a specific kinase superfamily and to search the proteins similar to the WTK1. Multiple alignments of WTK1 kinase domains with 23 different plant kinase domains were performed using Clustal Omega with default parameters. In this analysis, we included the 4 closest kinases identified according to the instructions of the Gateway Kit.

### Analysis of the TKP family in plants

A search for predicted proteins with a TKP structure was conducted in whole-genome assemblies of various plant taxa: O. sativa (http://rapdb.dna.affrc.go.jp/download/IRgsp1.html), Zea mays (http://ensembl.grameen.org/Zea_mays/), H. vulgare (http://webblast.ipk-gatersleben.de/barley_ibsc/), Secale cereale (http://webblast.ipk-gatersleben.de/ryeselect/), T. aestivum7, S. bicolor (http://phytozone.jgi.doe.gov/pz/portal.html?info=alias-Organ_Sbicolor), Arabidopsis thaliana (https://www.arabidopsis.org/download/index-auto.jsp?download=files/Proteins), Physcomitrella patens (https://plantsensembl.org/Physcomitrella_patens/), Solanum tuberosum (https://plants.ensembl.org/Solanum_tuberosum), Brassica napus (https://www.genoscope.cns.fr/brassicaneุs/data/), and Populus trichocarpa (http://phytozone.jgi.doe.gov/pz/portal.html?info=alias-Organ_Ptrichocarpa). Kinomes were extracted among various plant species. These dominant PCR markers were designed to identify the presence of the WTK1 functional allele, amplifying a 1 kb PCR product for Kin1 and 2 kb for Kin2.

### Data availability

Sequences have been deposited in GenBank under accession numbers MG649384 and MG674157. The authors declare that all other data supporting the findings of this study are available within the manuscript and its supplementary files or are available from the corresponding author upon request.

Received: 7 March 2018 Accepted: 26 July 2018

Published online: 03 October 2018

### References

1. Nevo, E., Korol, A. B., Beiles, A. & Fahima, T. Evolution of Wild Emmer and Wheat Improvement: Population Genetics, Genetic Resources, and Genome Organization of Wheats Progenitor, Triticum dicoccoides (Springer, Berlin, 2002).
2. Shewry, P. R. & Hey, S. J. The contribution of wheat to human diet and health. Food Energy Secur. 4, 178–202 (2015).
3. Wellings, C. R. Global status of stripe rust: a review of historical and current threats. Euphytica 179, 129–141 (2011).
4. Beddow, J. M. et al. Research investment implications of shifts in the global geography of wheat stripe rust. Nat. Plants 1, 15132 (2015).
5. Greenberg, J. T. & Yao, N. The role and regulation of programmed cell death in plant–pathogen interactions. Cell Microbiol. 6, 201–211 (2004).
6. Jones, J. D. J. & Dangl, J. L. The plant immune system. Nature 444, 323–329 (2006).
7. Krattinger, S. G. & Keller, B. Molecular genetics and evolution of disease resistance in cereals. New Phytol. 212, 320–332 (2016).
8. Chen, X. & Kang, Z. Stripe Rust (Springer Science+Business Media B.V., Berlin, Germany, 2017).
9. Marchal, C. et al. BED-domain-containing immune receptors confer diverse resistance spectra to yellow rust. Nat. Plants https://doi.org/10.1038/s41477-018-0236-4 (2018).
10. Fu, D. et al. A kinase–START gene confers temperature-dependent resistance to wheat stripe rust. Science 323, 1357–1360 (2009).
11. Gou, J.-Y. et al. Wheat stripe rust resistance protein WKS1 reduces the ability of the thylakoid-associated ascorbate peroxidase to detoxify reactive oxygen species. Plant Cell 27, 1755–1770 (2015).
12. Moore, J. W. et al. A recently evolved hexose transporter variant confers resistance to multiple pathogens in wheat. Nat. Genet. 47, 1494–1498 (2015).
13. Krattinger, S. G. et al. A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. Science 323, 1360–1363 (2009).
14. Hung, L. et al. Evolution and adaptation of wild emmer wheat populations to biotic and abiotic stresses. Annu. Rev. Phytopathol. 54, 276–301 (2016).
15. Gerechter- Amitai, Z. K., van Silfhout, C. H., Grama, A. & Kleitman, F. Yr15—a new gene for resistance to Puccinia strigosum in Triticum dicoccoides sel. G25. Euphytica 43, 187–199 (1989).
16. Ali, S. et al. Yellow rust epidemics worldwide were caused by pathogen races divergent from genomic lineages. Front. Plant Sci. 8, 1057 (2017).
17. Liu, T., Wan, A., Liu, D. & Chen, X. Changes of races and virulence genes in Puccinia striiformis f. sp. triticci, the wheat stripe rust pathogen, in the United States from 1968 to 2009. Plant Dis. 101, 1522–1532 (2017).
18. Sharma-Poudyal, D. et al. Virulence characterization of international isolates of the stripe rust pathogen, Puccinia striiformis f. sp. tritici. Plant Dis. 97, 379–386 (2013).
20. Sun, G. L. et al. Identification of molecular markers linked to the Yr15 stripe rust resistance gene of wheat in cultivated wild emmer, Triticum dicoccoides. Theor. Appl. Genet. 95, 622–628 (1997).

21. Yaniv, E. et al. Evaluation of marker-assisted selection for the stripe rust resistance gene Yr15, introgressed from wild emmer wheat. Mol. Breed. 35, 43 (2015).

22. Wellings, C. R. Puccinia striiformis in Australia: a review of the incursion, evolution, and adaptation of stripe rust in the period 1979–2006. Aust. J. Agric. Res. 58, 567–75 (2007).

23. Campbell, C. L. in The Epidemiology of Plant Diseases (ed. Jones, D. G.) 181–206 (Springer Science+Business Media, Singapore, 1998).

24. Raats, D. et al. The physical map of wheat chromosome 1BS provides insights into its gene space organization and evolution. Genome Biol. 14, R138 (2013).

25. Avni, R. et al. Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. Science 357, 93–97 (2017).

26. International Wheat Genome Sequencing Consortium (IWGSC). Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science 361, eaar7191 (2018).

27. Hanks, S. K., Quinn, A. M. & Hunter, T. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science 241, 42–52 (1988).

28. Kobilan, N., Taylor, S. S., Zhai, Y., Venter, J. C. & Manning, G. Structural and functional diversity of the microbial kinase. PLoS Biol. 5, 467–478 (2007).

29. Dardick, C., Schewssinger, B. & Ronald, P. Non-arginine-aspartate (non-RD) kinases are associated with innate immune receptors that recognize conserved microbial signatures. Curr. Opin. Plant Biol. 15, 358–366 (2012).

30. Brueggerman, R. et al. The barley stem rust-resistance gene Yr15 is a novel disease-resistance gene with homology to receptor kinases. Proc. Natl. Acad. Sci. USA 99, 9328–9333 (2002).

31. Zang, W. et al. Fine mapping and identification of a candidate gene for the barley Yr8 true loose smut resistance gene. Theor. Appl. Genet. 128, 1343–1357 (2015).

32. Manning, G., Whyte, D. B., Martínez, R., Hunter, T. & Sadowsam, S. The protein kinase complement of the human genome. Science 298, 1912–1934 (2002).

33. Huse, M. & Kuriyan, J. The conformational plasticity of protein kinases. Cell 109, 275–282 (2002).

34. Zulawski, M. et al. The Arabidopsis Kinome: phylogeny and evolutionary insights into functional diversification. BMC Genom. 15, 548 (2014).

35. Bosselini, E., Wicker, T., Knobel, P. A. & Keller, B. Comparison of orthologous loci from small grass genomes Brachypodium and rice: implications for wheat genomics and grass genome annotation. Plant J. 49, 704–717 (2007).

36. Lehti-Shiu, M. D. & Shiu, S. H. Diversity, classification and function of the plant protein kinase superfamily. Philos. Trans. R. Soc. Lond. Ser. B 367, 2619–2639 (2012).

37. Couto, D. & Zipfel, C. Regulation of pattern recognition receptor signalling in plants. Nat. Rev. Immunol. 16, 537–552 (2016).

38. Shiu, S.-H. & Bleecker, A. B. Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in Arabidopsis. Genetics 168, 1625–1633 (2004).

39. Aloussi, S. et al. Construction and characterization of a hexaploid wheat (Triticum aestivum L.) BAC library from the reference germplasm “Chinese Spring”. Cereal Res. Commun. 31, 331–338 (2003).

40. Usuy, C. et al. A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat. BMC Plant Biol. 9, 115 (2009).

41. Ishida, Y., Tsunashima, M., Hiei, Y. & Komari, T. in Methods in Molecular Biology, Agrobacterium Protocols 3rd edn (ed. Wang, K.) 189–198 (Springer, New York, 2014).

42. Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425 (1987).

43. Zuckerkandl, E. & Pauling, L. in Evolving Genes and Proteins (eds Bryanson., V. & Vogel, H. J.) 97–166 (Academic Press, New York, 1965).

44. Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791 (1985).

45. Suomi, A., Narvanto, A. & Schulman, A. H. The BARE-1 retrotransposon is transcribed in barley from an LTR promoter active in transient assays. Plant Mol. Biol. 31, 295–306 (1996).

46. Altschul, S. F. et al. Protein database searches using compositionally adjusted substitution matrices. FEBS J. 272, 5101–5109 (2005).

47. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Biol. Evol. 31, 1889–1902 (2014).

48. Stamatakis, A. RAXML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313 (2014).

49. Jones, D. T., Taylor, W. R. & Thornton, J. M. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8, 275–282 (1992).

Acknowledgements
This research was supported by the European Community Framework projects BioExploit (CT-2005-51399) and TriticaceaeGenome (FP7-212019), the COST Actions TriticiGen (FA0604), SUSTAIN (FA1208), and Impact of Nuclear Domains on Gene Expression and Plant Traits (INDEPTH, CA1612), the GSC TT Center for Science (Finland), the Israel Science Foundation (1719/08), and the US-Israel Binational Agricultural Research and Development Fund (IS-4628-13, US-4916-16). We thank the International Wheat Genome Sequencing Consortium (IWGSC) for providing access to unpublished wheat genome sequences and the Global Yellow Rust Gene Bank for providing Pst races. We also thank P. Auvinen, P.C. Bailey, S. Barinova, H. Belcram, A. Dahan, Z. Gerechter-Amiat, A. Graama, M. Iso-Kokkila, T. Kost-Papo, D. Lewinsohn, U.S. Lönnqvist, A. Malik, Y. Manisterski, I. Manov, H. Miah, A.-M. Narvanto, M.L. Neufeld, K. Neufeld, Y. Ronin, O. Rybak, J. Shanir, N. Sher, O.-P. Smolander, and J. Tanskanen.

Author contributions
E.Y., L.H., D.R., V.K., L.F., A.F., Z.F., H.S., and A.D. developed markers. E.Y., L.H., D.R., V.K., and L.F. developed recombinants, conducted Pst phenotyping, and constructed genetic maps. D.R., Z.F., T.P., and A.B.K. constructed the CS 1BS physical map and identified the scaffolding spanning the Yr15 gene region. E.Y., L.H., and B.C. constructed and phenotyped PiG25 non-gridded BAC library. L.P. and P.L. sequenced the G25 BAC clones and assembled the G25 contigs. L.H., T.F., and A.F. identified the candidate genes. L.H., L.F., G.L., T.K., H.B., and T.F. developed EMS mutants. S.C. and J.D. performed the transgenic experiments. L.H. and W.C. isolated and sequenced full-length cDNA. W.C. and M.J. conducted the cellular localization experiment. V.K., T.K., S.C., and L.H. conducted expression analysis. V.K. carried out microscopic observations and quantification of fungal biomass upon Pst inoculation. K.S. performed confocal microscopy observations. A.F., C.S., and D.R. performed bioinformatics analyses of genomic and RNA-Seq data. A.F. and V.K. analyzed the predicted sequences of plant TKP family. V.K.
and L.H. conducted the WTK1 distribution study. C.K.S. and M.S.H. phenotyped Yr15/Yr5 ILs with a set of Pst isolates. T.F. conceived the project; T.F., A.H.S. and A.B.K. provided project leadership and planned the positional cloning strategy. V.K., T.F., and L.H. co-wrote the manuscript with input from all co-authors.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-06138-9.

Competing interests: The University of Haifa and The Natural Resources Institute Finland (Luke) have filed two patent applications (numbers 62575688, 62575670) on the use of the sequences of the stripe rust resistance gene and the use of diagnostic markers that are based on the results described here, in which T.F., A.B.K., I.H., D.R., V.K., Z.P., A.F., A.H.S., and E.Y. are listed as inventors. The remaining authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018