Published online April 25, 2019

Accessing a Russian Wheat Aphid Resistance Gene in Bread Wheat by Long-Read Technologies

Zuzana Tulpová, Helena Toegelová, Nora L. V. Lapitan, Frank B. Peairs, Jiří Macas, Petr Novák, Adam J. Lukaszewski, David Kopecký, Mira Mazáčová, Jan Vrána, Kateřina Holušová, Philippe Leroy, Jaroslav Doležel, and Hana Šimková*

Z. Tulpová, H. Toegelová, D. Kopecký, M. Mazáčová, J. Vrána, K. Holušová, J. Doležel, H. Šimková, Inst. of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 31, 78371 Olomouc, Czech Republic; N.L.V. Lapitan, Bureau for Food Security, United States Agency for International Development, Washington, DC 20004; F.B. Peairs, Dep. of Bioagricultural Sciences and Pest Management, Colorado State Univ., Fort Collins, CO 80523; J. Macas, P. Novák, Biology Centre, Czech Academy of Sciences, Institute of Plant Molecular Biology, Branišovská 31, 37005 České Budějovice, Czech Republic; A.J. Lukaszewski, Dep. of Botany and Plant Sciences, Univ. of California, Riverside, CA 92521; P. Leroy, Genetics, Diversity and Ecophysiology of Cereals, Inst. National de la Recherche Agronomique, Univ. Clermont Auvergne, 5 chemin de Beaulieu, 63039 Clermont-Ferrand, France.

ABSTRACT Russian wheat aphid (RWA) (*Diuraphis noxia* Kurdjumov) is a serious invasive pest of small-grain cereals and many grass species. An efficient strategy to defy aphid attacks is to identify sources of natural resistance and transfer resistance genes into susceptible crop cultivars. Revealing the genes helps understand plant defense mechanisms and engineer plants with durable resistance to the pest. To date, more than 15 RWA resistance genes have been identified in wheat (*Triticum aestivum* L.) but none of them has been cloned. Previously, we genetically mapped the RWA resistance gene *Dn2401* into an interval of 0.83 cM on the short arm of chromosome 7D and spanned it with five bacterial artificial chromosome (BAC) clones. Here, we used a targeted strategy combining traditional approaches toward gene cloning (genetic mapping and sequencing of BAC clones) with novel technologies, including optical mapping and long-read nanopore sequencing. The latter, with reads spanning the entire length of a BAC insert, enabled us to assemble the whole region, a task that was not achievable with short reads. Long-read optical mapping validated the DNA sequence in the interval and revealed a difference in the locus organization between resistant and susceptible genotypes. The complete and accurate sequence of the *Dn2401* region facilitated the identification of new markers and precise annotation of the interval, revealing six high-confidence genes. Identification of *Epoxide hydrolase 2* as the most likely *Dn2401* candidate opens an avenue for its validation through functional genomics approaches.

Abbreviations: 2OG-Fe(II) oxygenase, 2-oxoglutarate and Fe(II)-dependent oxygenase; 7DS, short arm of wheat chromosome 7D; BAC, bacterial artificial chromosome; BNG map, Bionano genome map; CS, cultivar Chinese Spring; EH, epoxide hydrolase; *EH2*, *epoxide hydrolase 2*; IWGSC, International Wheat Genome Sequencing Consortium; PCR, polymerase chain reaction; RWA, Russian wheat aphid; SNP, single nucleotide polymorphism; UTR, untranslated region.

CORE IDEAS

- An accurate sequence of the *Dn2401* region was generated from Illumina and nanopore reads.
- Structural variation in the region was analyzed by optical mapping and resequencing.
- New markers located *Dn2401* within a 133.2-kb interval with six high-confidence genes.
- *Epoxide hydrolase 2* was identified as the most likely resistance gene candidate.

Russian wheat aphid (RWA) was first reported in 1978 as a pest of small-grain cereals in South Africa (Walters et al., 1980). Since then, the aphid has spread around the world, most recently to Australia, and has become a serious invasive pest not only of wheat and barley (*Hordeum vulgare* L.) but also of many other plants from 43 genera, including over 140 species of cultivated and wild grasses (Yazdani et al., 2017). During feeding, aphids remove plant photoassimilates, which results in chlorosis, longitudinal streaking around the main leaf vein, head trapping, a substantial reduction in biomass, and, in severe cases, plant death (Burd and Burton, 1992).

Citation: Tulpová, Z., H. Toegelová, N.L.V. Lapitan, F.B. Peairs, J. Macas, P. Novák, A. J. Lukaszewski, D. Kopecký, M. Mazáčová, J. Vrána, K. Holušová, P. Leroy, J. Doležel, and H. Šimková. 2019. Accessing a Russian wheat aphid resistance gene in bread wheat by long-read technologies. Plant Genome 12:180065. doi:10.3835/plantgenome2018.09.0065

Received 10 Sept. 2018. Accepted 10 Dec. 2018.
*Corresponding author (simkovah@ueb.cas.cz).

Another undesirable effect accompanying aphid feeding is leaf rolling, which that serves as a shelter against natural predators of the aphid and also against insecticide spraying. Thus the most efficient strategy to defy RWA attacks lies in identifying sources of natural resistance and introducing them into susceptible cultivars.

Plant defenses against insect damage include three mechanisms that are frequently combined: antibiosis (mechanical and chemical defensive factors impacting insect biology, including fertility), antixenosis (nonpreference of a particular plant as a host), and tolerance (the ability of a plant to withstand insect damage) (Botha et al., 2005). To date, more than 15 RWA resistance genes, mostly underlying tolerance and antibiosis, have been described in wheat. Many of these originated from *Aegilops tauschii* Coss. as inferred from their map location in the D genome (Du Toit, 1987, 1989; Du Toit et al., 1995; Fazel-Najafabadi et al., 2015; Liu et al., 2001; Ma et al., 1998; Miller et al., 2001; Peng et al., 2007; Valdez et al., 2012; Voothuluru et al., 2006). At least one resistance gene, *Dn7*, originating from rye (*Secale cereale* L.) has been reported (Anderson et al., 2003; Marais et al., 1994; Lapitan et al., 2007). New highly virulent RWA biotypes have rendered previously resistant cultivars susceptible, which has stimulated an intensive search for novel sources of resistance as well as efforts to unravel the mechanisms underlying the trait. Although several components of the plant defense pathways have been proposed (Anderson et al., 2014; Van Eck et al., 2014), to date, no aphid resistance gene has been cloned in wheat. Nine of the hitherto mapped RWA resistance genes, namely *Dn1*, *Dn2*, *Dn5* (Du Toit, 1987), *Dn6* (Saidi and Quick, 1996), *Dn8*, (Liu et al., 2001), *Dnx* (Harvey and Martin, 1990), *Dn2401* (Voothuluru et al., 2006; Fazel-Najafabadi et al. 2015), *Dn626580* (Valdez et al., 2012), and *Dn100695* (Aykut Tonk et al., 2016) are located on the short arm of wheat chromosome 7D (7DS). Except for *Dn8*, which is located at the terminal part of the arm, the remaining genes were mapped to the interstitial part of 7DS and most of them were found linked to marker *Xgwm111* (Liu et al. 2002; Fazel-Najafabadi et al. 2015). It has not yet been resolved if these genes are allelic or tightly linked.

Here, we focused on the *Dn2401* gene identified in line 'CI2401', which originated from Tajikistan. Previously, Staňková et al. (2015) mapped the *Dn2401* gene into a 0.83-cM interval on chromosome arm 7DS, delimited by the markers *Xowm705* and *Xowm711*. The interval was spanned by five overlapping BAC clones from a 7DS-specific BAC library of the cultivar Chinese Spring (CS) (Šimková et al., 2011). Chinese Spring, a reference genome of bread wheat, is susceptible to RWA (Peng et al., 2009). In the current work, we sequenced, assembled, and annotated the five CS BAC clones, proposing a list of genes located in the interval. The study of Staňková et al., (2015) pointed to a striking decrease (a minimum of eightfold) in the physical/genetic map distance ratio in the *Dn2401* interval compared with the neighboring regions. We speculated that this decrease is likely to be caused by a higher recombination rate in the region. However, an alternative explanation might be a structural variation between CS, in which we measured the physical distance, and CI2401 and Glupro, which were used to construct the genetic map. In the case of the larger deletion in CS, the list of candidate genes deduced from the CS reference sequence might not be complete. Considering the size of the region (~300 kb) and a highly repetitive nature of the wheat genome, it was unrealistic to analyze the entire interval in CI2401 by resequencing. Thus an easier approach to a fast and affordable comparative analysis of the region was chosen: a long-read optical mapping on the nanochannel platform of Bionano Genomics (San Diego, CA) (Lam et al. 2012), which generates maps of short sequence motif that are hundreds to thousands of kb long and can be used to support or validate sequence scaffolding and to perform large-scale comparative analyses. Thanks to a previously constructed Bionano genome (BNG) map of the 7DS arm from CS (Staňková et al., 2016), a newly prepared BNG map of 7DS from CI2401 facilitated a straightforward comparison of the entire region between the two accessions.

An accurate and complete sequence of the interval delimited by the flanking markers is a prerequisite for successful gene cloning. We approached the region through low complexity sequencing of individual BAC clones on the short-read Illumina (San Diego, CA) platform, which did not produce a continuous sequence of the entire interval. Recently, two advanced whole-genome assemblies of bread wheat were released: Triticum 3.1 (Zimin et al., 2017), combining short Illumina and long Pacific Biosciences (Menlo Park, CA) reads with the fully annotated IWGSC RefSeq version 1.0 (International Wheat Genome Sequencing Consortium, 2018), which was based on short reads only and coupled the whole-genome assembly with assemblies of physical map-ordered BAC clones. Triticum 3.1 and IWGSC RefSeq version 1.0 cover 97 and 92% of the estimated wheat genome size of 15.76 Gb (International Wheat Genome Sequencing Consortium, 2018), respectively. Such reference genome assemblies promise to be an excellent tool for accelerating gene cloning. Unfortunately, not even these assemblies allowed us to close the gap in the BAC assembly of the region of interest. To finally resolve the *Dn2401* region, we adopted an alternative long-read technology: nanopore sequencing on the MinION platform of Oxford Nanopore Technologies, which produces reads reaching tens or even hundreds of kb (Deschamps et al., 2018), thus having the potential to span long transposable elements– the main obstacle in assembling complex genomes. Coupling nanopore and short reads of the BAC clones allowed us to complete the sequence of the *Dn2401* region, which enabled precise annotation of the region, identification of new markers, and identification and resequencing of candidate genes.

Table 1. Predicted high-confidence coding sequences in the wider *Dn2401* region of Chinese Spring wheat.

Gene No.	IWGSC RefSeq ID†	Functional analysis‡	Domain ID
1	*TraesCS7D01G224100.1*	HVA22	PF03134, IPR004345
	TraesCS7D01G224100.2		
2	***TraesCS7D01G224200.1***	**Epoxide hydrolase 2**	**PF00561, IPR000073**
3	***TraesCS7D01G224300.1***	**Aspartic peptidase family**	**PF14543, PF14541, IPR001461**
4	***TraesCS7D01G224400.1***	**Bacterial trigger factor protein**	**PF05697, PF05698, IPR005215**
5	***TraesCS7D01G224500.1***	**2-oxoglutarate and Fe(II)-dependent oxygenase superfamily protein**	**PF14226, PF03171, IPR005123**
6	***TraesCS7D01G224600.1***	**Nuclear pore complex protein Nup98–Nup96**	**–**
7	***TraesCS7D01G224700.1***	**Dual specificity protein phosphatase, putative**	**PF00782, IPR000340**
8	*TraesCS7D01G224800.1*	Microtubule associated family protein	PF03999
9	*TraesCS7D01G224900.1*	RNA recognition motif	IPR000504
10	*TraesCS7D01G225000.1*	Membrane protein	PF09991, IPR018710
11	*TraesCS7D01G225100.1*	Acyl-coenzyme A oxidase	PF02770, PF00441, PF01756, IPR002655
12	*TraesCS7D01G225200.1*	Vacuolar protein sorting-associated protein VTA1	PF04652, IPR023175
13	*TraesCS7D01G225300.1*	PHD finger protein	PF00628, PF01426, IPR001025

† International Wheat Genome Sequencing Consortium RefSeq ID as in International Wheat Genome Sequencing Consortium (2018).

‡ The wider region was delimited by the markers *Xowm705* and *Xowm711*. Coding sequences from the reduced interval delimited by markers *Xowm716* and *Xowm713* are in bold.

MATERIALS AND METHODS

Plant Material

Genetic mapping was performed in an F_2 population derived from a cross between a RWA-resistant winter wheat line CI2401 and the susceptible cultivar Glupro. The current population (designated CI2401 × Glupro F_2) consisted of 333 plants and combined 184 individuals from a previous study (CI2401 × Glupro F_2–1)(Staňková et al., 2015) with a new set of 149 F_2 plants (CI2401 × Glupro F_2–2) being included in this work.

The short arms of homeologous group 7 chromosomes of CS were flow-sorted from the respective double ditelosomic lines. Seed samples of these lines were kindly provided by Prof. Bikram Gill (Kansas State Univ., Manhattan, KS). To enable flow-sorting of the 7DS arm from line CI2401, a 7DS ditelosomic line was generated by misdivision of univalent $7D_{CI2401}$ in a hybrid between CS nullisomic 7D (CS N7D) × CI2401. Among 280 progeny screened, three telocentric $7DS_{CI2401}$ chromosomes were identified and the respective plants were self-pollinated. Ditelosomic 7DS were selected from their progeny and propagated.

Sequencing and Assembly of BAC Clones

Five overlapping BAC clones from CS 7DS arm-specific BAC library (Šimková et al., 2011) spanning the *Dn2401* region (*TaaCsp7DS088E19*, *TaaCsp7DS010E01*, *TaaCsp7DS112N11*, *TaaCsp7DS044E09*, and *TaaCsp7DS086H04*) were purified with a NucleoSpin 96 Flash kit (Macherey-Nagel, Düren, Germany) and individually pair-end sequenced on the Illumina MiSeq platform. The 2 × 250 bp reads obtained were assembled with RAY (Boisvert et al., 2010) separately for each clone. The resulting 90 contigs, which were mostly doubled because of BAC overlaps, were manually merged in Geneious version 7.1.2 (http://www.geneious.com, accessed 14 Feb. 2019)(Kearse et al., 2012), with support from BLAST and the 7DS BNG map of CS (Staňková et al., 2016). A larger gap in the clone *TaaCsp7DS088E19* was closed with scaffolds assembled by SASSY from the 2 × 150 bp reads obtained from pooled 7DS BAC clones (Tulpová et al., 2019). The SASSY scaffolds did not close the 14.4-kb gap identified in *TaaCsp7DS086H04*.

Additional data for the BAC clone *TaaCsp7DS086H04* were obtained by long-read nanopore sequencing. To maximize the read length, BAC DNA was extracted manually by the alkaline lysis method followed by phenol-chloroform extraction and ethanol precipitation. As a final step, the DNA was purified by a 5-min incubation with 1:1 AMPure XP beads (Beckman Coulter, Miami, FL) and eluted into 30 μL of 10 mM tris(hydroxymethyl)-aminomethane (pH 8.5). The sequencing library was prepared using the Rapid Sequencing Kit (SQK-RAD003, Oxford Nanopore Technologies, Oxford, UK) and the sample was run on the MinION platform (Oxford Nanopore Technologies). Size-selected reads (352 in total) ranging from 10 to 131 kb and representing 55-fold coverage of the BAC sequence were combined with the 2 × 250 bp Illumina reads for hybrid assembly with MaSuRCA (Zimin et al., 2013). The resulting assembly was manually verified by comparison with the longest nanopore reads spanning the entire (or a majority of) the insert length to confirm correct assembly of structurally complex repeated regions.

Sequence Analysis

Annotation of the *Dn2401* region was performed with the TriAnnot (online version 4.3.1) pipeline (Leroy et al., 2012) and predictions were manually curated with GenomeView 2250 software (genomeview.org, accessed 14 Feb. 2019)(Abeel et al., 2012). This manual annotation was compared with the IWGSC RefSeq Annotation version 1.0 (International Wheat Genome Sequencing Consortium, 2018), once it was released (Table 1). To resequence candidate genes from CI2401 and Glupro, primers were designed to cover all of the coding sequences as well as the upstream and downstream regulatory regions. Primers, polymerase chain reaction (PCR) conditions and amplicon sizes are provided in the Supplemental Table S3.

Table 2. Markers mapped in the *Dn2401* region of wheat.

Marker ID	Polymorphism†	Primers	Ta	Amplicon size
			°C	bp
Xowm713	SNP G/A	Forward: CGTGCATGATCCTCGACTATGAT Reverse: TTGCCTATTTTAACAATGCTCGT	67	467
Xowm714	SNP A/G	Forward: TCTGTAATGTGGAATGTTGTCTTAGT Reverse: GGCTGAAACAAAGAATCCATC	64	574
Xowm715	PAV	Forward: CCCTCGATACGAGCTGGA Reverse: GAGGGAGGGAGGTTGTCA	62	232
Xowm716	SNP C/T	Forward: GGGCAAAATGGTCTTTTCAC Reverse: AAGAAATTCGACTGAAATGAGGA	62	579

† PAV, presence–absence variation; SNP, single nucleotide polymorphism; Ta, annealing temperature

For templates with a high GC content, 5% dimethyl sulfoxide was added into a PCR reaction. Amplification products were cleaned up and Sanger-sequenced on an ABI 3730xl DNA analyzer (Applied Biosystems, Waltham, MA). Polymorphisms between parental lines were identified by sequence alignment in Geneious version 7.1.2.

Bionano Genome Map and Analysis of Structural Variation

To investigate the overall sequence organization of the *Dn2401* region in CI2401 and to detect a possible structural variation between CI2401 and the reference genome of CS, we constructed a new BNG map of the chromosome arm 7DS originating from CI2401 and compared it with the previously assembled 7DS BNG map of CS (Staňková et al., 2016). The 7DS chromosome arm was flow-sorted from a 7DS ditelosomic line of CI2401 and the BNG map was constructed as described in Staňková et al. (2016) with minor modifications. High molecular weight DNA molecules labeled at Nt.BspQI sites (GCTCTTC) with Alexa546-dUTP fluorochromes were analyzed on the Irys platform (Bionano Genomics). In total, 79 Gb of raw data greater than 150 kb, corresponding to 207× of the 7DS arm (Šafář et al., 2010), were collected from a single IrysChip (Supplemental Table S1). De novo assembly of the optical map was done through a pairwise comparison of all single molecules and graph building in IrysView software (Bionano Genomics). A *p*-value threshold of $1e^{-10}$ was used during the pairwise assembly, $1e^{-11}$ for the extension and refinement steps, and $1e^{-15}$ for the final refinement. Sequence-to-map and map-to-map alignments were done with IrysView 2.1.1 software. To align the sequences, *cmaps* were generated from fasta files of the BAC sequence assemblies and 7D pseudomolecules of wheat whole-genome assemblies IWGSC RefSeq version 1.0 (International Wheat Genome Sequencing Consortium, 2018) or Triticum 3.1 (Zimin et al., 2017). Query-to-anchor comparisons were performed with default parameters and a *p*-value threshold of $1e^{-10}$.

Narrowing Down the *Dn2401* Interval

To identify new markers within the interval delimited by the flanking markers *Xowm705* and *Xowm711*, we adopted a strategy proposed in our previous study (Staňková et al., 2015) with some modifications. Briefly, to design D-genome-specific primers, assembled sequences of five BAC clones spanning the *Dn2401* region in 7DS arm of CS were compared with homeologous sequence contigs from the short arms of CS chromosome 7A and 7B (Berkman et al., 2013). The specificity of the proposed primers was tested and PCR was optimized with DNA from the short arms of chromosomes 7A, 7B, and 7D, which were flow-sorted from respective ditelosomic lines of CS. New polymorphisms between CI2401 and Glupro were identified by amplicon sequencing, with DNA of flow-sorted chromosome arm 7DS (CI2401) and chromosome 7D (Glupro) as templates. The 7D chromosome could be distinguished from other chromosomes of Glupro after FISHIS labeling of GAA repeats (Giorgi et al., 2013). Sequences of the PCR products from both parents were compared with Geneious version 7.1.2 and their specificity to the *Dn2401* region was verified by alignment to sequences of the assembled BAC clones and the CS RefSeq version 1.0 (International Wheat Genome Sequencing Consortium, 2018).

Genetic mapping of the newly identified single nucleotide polymorphism (SNP) markers *Xowm713*, *714*, and *716* was carried out by Sanger-sequencing the respective amplicons from plants of the extended CI2401 × Glupro F_2 mapping population. The presence–absence variation of the marker *Xowm715* was assayed by PCR and electrophoresis. Primers and PCR conditions for all markers are given in Table 2. The 149 newly added individuals of the mapping population (CI2401 × Glupro F_2–2) were also assayed for the original flanking markers *Xowm705* and *Xowm711*. Scoring of the RWA response was done at the Colorado State University Insectary in the same fashion as described for the first part of the mapping population (Staňková et al., 2015). Genotype and phenotype data of all 333 individuals from the extended CI2401 × Glupro F_2 mapping population were processed with JoinMap version 4.0 software (Van Ooijen and Voorrips, 2001).

Data Availability

The BAC hybrid assembly was submitted to NCBI (GenBank accession number MH806875). The optical map of CS 7DS is available at https://urgi.versailles.inra.fr/download/iwgsc/IWGSC_RefSeq_Annotations/v1.0/iwgsc_refseqv1.0_optical_maps_group7.zip (accessed 14 Feb. 2019). The optical map of CI2401 7DS is available on request.

RESULTS

Sequencing of Candidate BAC Clones

Five candidate BAC clones, spanning the *Dn2401* interval as delimited by Staňková et al., (2015), were Illumina sequenced and assembled, resulting in two scaffolds of 70.8 and 367 kb (Supplemental Fig. S1). To evaluate the completeness and correctness of the assembly, the scaffolds were aligned to the previously constructed BNG map of the CS 7DS arm (Staňková et al., 2016), which showed a 14.4-kb gap in the region corresponding to the BAC clone *TaaCsp7DS086H04*. We attempted to close

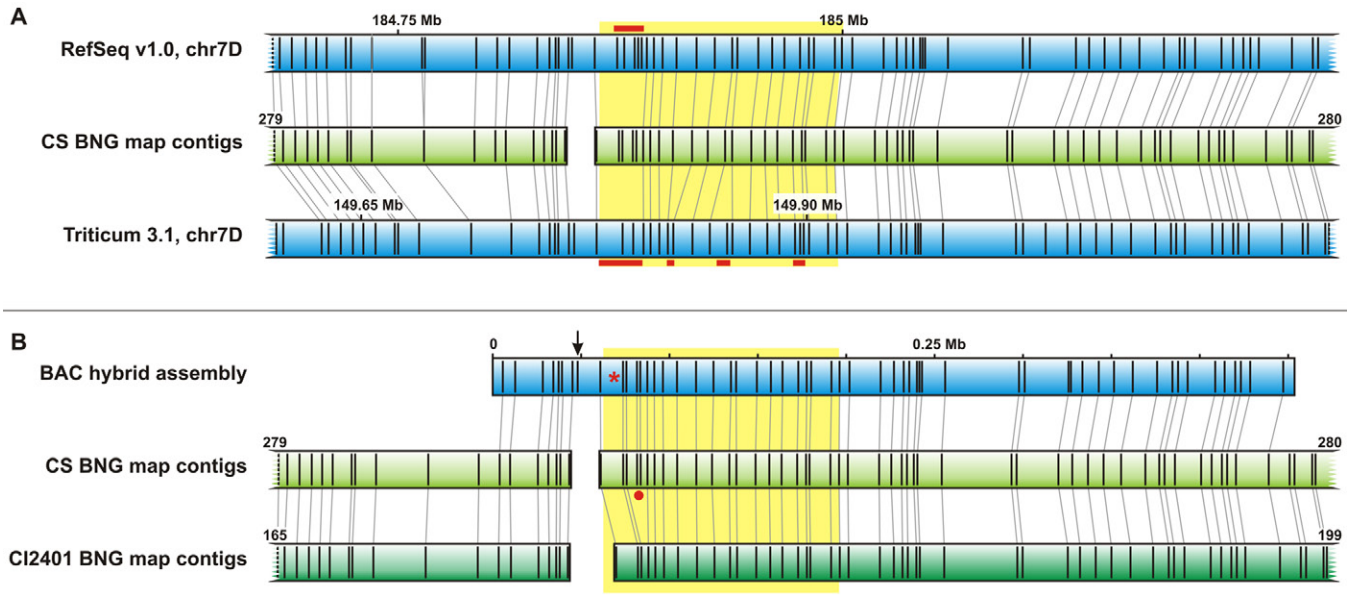

Fig. 1. Assembly and structural variation of the *Dn2401* region in wheat. (A) Alignments of the Bionano genome (BNG) map of the short arm of wheat chromosome 7D (7DS) of Chinese Spring (CS) (light green bar) to in silico Nt.BspQI digests of whole-genome wheat assemblies (blue bars) showed misassembled regions (red lines) both in the IWGCS RefSeq version 1.0 (International Wheat Genome Sequencing Consortium, 2018) and Triticum 3.1 (Zimin et al., 2017) assemblies. Numbers above the blue bars indicate positions in the 7D pseudomolecules. (B) The hybrid assembly of CS bacterial artificial chromosome (BAC) clones (blue bars) is in agreement with the CS BNG map. The 15-kb gap between BNG map contigs 279 and 280 (CS) and between 165 and 199 (CI2401) is a consequence of a fragile site in the DNA molecule (arrow). Alignment of 7DS BNG maps from susceptible CS and resistant CI2401 indicated structural variation (red dot) upstream of the *EH2* gene (red asterisk). The yellow box highlights the *Dn2401* region as delimited by *Xowm716* and *Xowm713*.

that gap with two recently published whole-genome CS sequence assemblies (Zimin et al., 2017; International Wheat Genome Sequencing Consortium, 2018) but the BNG map indicated misassemblies in the region of interest for both of them (Fig. 1A). To complete the sequence of the region, long-read nanopore data were generated for the BAC clone *TaaCsp7DS086H04* and were combined with the BAC short-read data (Supplemental File S1), which resulted in a hybrid BAC assembly of 452,442 bp that spanned the entire interval between *Xowm705* and *Xowm711* by a continuous sequence of 304,109 bp (no internal gaps) and was in full agreement with the BNG map of CS 7DS (Fig. 1B, Supplemental Fig. S1).

Candidate Gene Identification and Comparison

Gene modeling on the assembled sequence of CS BAC clones spanning the *Dn2401* interval led to the prediction of 13 coding loci (Table 1, Supplemental Fig. S2), including the gene *HVA22*, whose coding sequence lies 182 bp distal to the interval but, because of its close proximity, the downstream regulatory region of the gene might potentially fall within the interval.

Narrowing down the *Dn2401* interval by mapping the new marker *Xowm713* excluded six proximally located genes (Table 1). Three of the remaining genes, *HVA22*, *Epoxide hydrolase 2* (*EH2*), and a gene coding for a 2-oxoglutarate and Fe(II)-dependent oxygenase [2OG-Fe(II) oxygenase] superfamily protein, were considered potential candidates for the RWA resistance gene *Dn2401*. *HVA22*, initially identified in barley (Shen et al., 1993), is a gene induced by abscisic acid, which is known to mediate developmental and physiological processes such as seed development and stress responses, including plant responses to aphid feeding (Smith and Boyko, 2007). *EH2* was found previously to be one of the genes showing increased expression after aphid feeding in a RWA-resistant line, PI220127 (Boyko et al., 2006). The 2OG-Fe(II) oxygenase appears to be involved in plant response to biotic stress because the underlying gene was found to be overexpressed after inoculating wheat with fungal pathogens (Zhang et al., 2014) (www.wheat-expression.com, accessed 14 Feb. 2019).

To compare the *HVA22* locus between the parents of the mapping population (the RWA-resistant line CI2401 and the RWA-susceptible Glupro), we sequenced in both of them a continuous region starting 3.2 kb upstream and ending 765 bp downstream of the *HVA22* coding sequence. The gene consisted of five exons that did not differ between the parents, and four introns, the second of which bore one SNP [A/G]. The A variant of the SNP is shared by the resistant line CI2401 and the susceptible CS. In the promoter region of both CI2401 and Glupro, we observed abscisic acid response elements ABRE2 and ABRE3 as well as three TATAA domains localized 1105, 717, and 163 bp upstream of the gene. In the 3' untranslated region (UTR), in both accessions, we found two AUAAA domains and in the downstream region, we precisely positioned a 43-bp deletion that was specific for Glupro, which was earlier identified as the marker *Xowm711* (Staňková et al., 2015). The sequence spanning this indel was also inspected in CS and it was confirmed that the

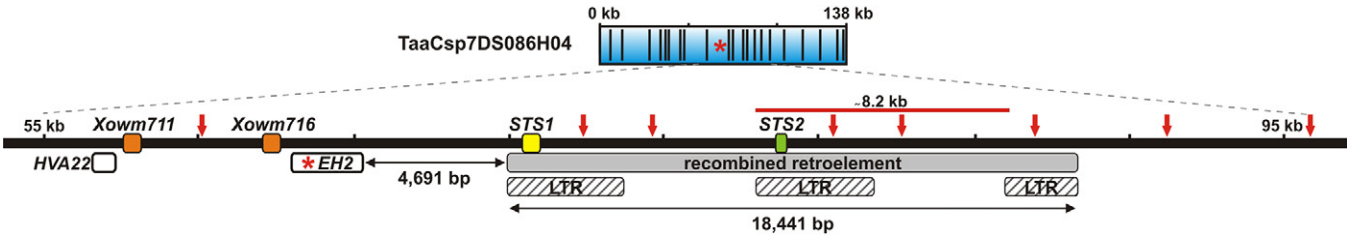

Fig. 2. Annotation and comparative analysis of the *HVA22–EH2* region in wheat. The analysis was performed on a hybrid assembly of the bacterial artificial chromosome (BAC) clone *TaaCsp086H04*, shown as an Nt.BspQI in silico digest (blue bar). Nt.BspQI recognition sites are depicted as black vertical lines (top) or red arrows (bottom). White boxes stand for coding sequences; the orange, yellow and green boxes correspond to markers applied in the study. The red line indicates the approximate position and size of the variable region missing in CI2401, as deduced from a comparison of Bionano genome (BNG) maps and a polymerase chain reaction assay for the *STS2* marker.

susceptible CS had the same haplotype as the resistant CI2401 in the *HVA22* locus. Hence this polymorphism was not associated with the resistant phenotype.

To assess the *EH2* locus, we compared sequences starting 2871 bp upstream and ending 963 bp downstream of the *EH2* coding sequence that consisted of three exons and two introns. The total length of the coding sequence was 2275 bp. In the 5' flanking region of the gene, we found a *Ty1/Copia* retroelement located 1550 bp upstream of the coding sequence (Supplemental Fig. S3, Supplemental File S1). Alignment of the CI2401 and Glupro sequences revealed no polymorphism in the entire *EH2* region except for a single SNP [C/T] located in the 3'UTR, 290 bp downstream from the coding sequence. The T variant was shared by the susceptible cultivars CS and Glupro.

We also resequenced and compared a continuous DNA segment of 8470 bp covering the entire *2OG-Fe(II) oxygenase* locus as well as a neighboring gene coding for a bacterial trigger factor protein, including their up- and downstream regulatory regions (Supplemental Fig. S4). The sequence comparison revealed a single polymorphism corresponding to the marker *Xowm714*, located in the second exon of the gene for the bacterial trigger factor protein. This enabled a haplotype analysis of the locus, revealing that the A variant of the SNP was shared by the resistant line CI2401 and the susceptible CS. The data obtained by resequencing the two loci did not support them as the true *Dn2401* candidates. In the search for new markers in the interval, we also analyzed the entire gene for aspartic peptidase, including the adjacent regions and the majority of the gene for dual specificity protein phosphatase (Supplemental Table S3) without detecting any polymorphisms.

Bionano Genome Mapping and Structural Variations in the *Dn2401* Region

With the aim of inspecting the overall structure in the *Dn2401* interval in the CI2401 line and to compare it with the reference genome of CS, we investigated the region by optical mapping. By using the Bionano Genomics Irys platform and the approach applied previously for constructing the 7DS BNG map of CS, we generated a BNG map of the 7DS chromosome arm of CI2401 that consisted of 468 BNG map contigs with an average length of 765 kb and for which the sequence

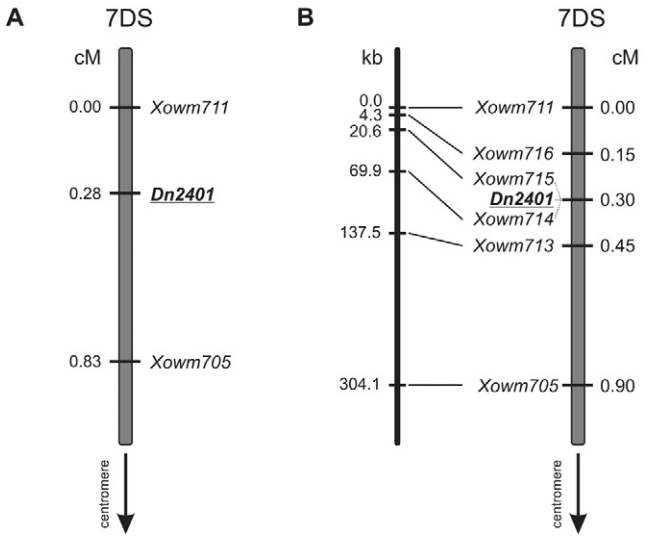

Fig. 3. Genetic and physical maps of the *Dn2401* region in wheat. (A) The *Dn2401* interval, as delimited in a previous study (Staňková et al., 2015). (B) The genetic map generated in the current study with an extended mapping population and new markers. The physical distances of the markers are shown on the left.

length of the shortest contig at 50% of the total genome length was 1.36 Mb. This novel 7DS-CI2401 BNG map had a total length of 358 Mb and covered 94% of the estimated arm length (Supplemental Table S1).

Alignments of the two 7DS BNG maps to the BAC assembly of the region led to the identification of two BNG map contigs spanning the *Dn2401* region in each of the accessions: contigs 279 and 280 for CS and contigs 165 and 199 for CI2401 (Fig. 1B). The reason for the BNG map split was the presence of a "fragile site" that arose from the occurrence of proximally located Nt.BspQI nicking sites on opposite DNA strands. This introduced a gap of ~15 kb in both BNG maps. Alignment of the two maps indicated an 8.2-kb indel at the distal end of the BNG map contigs 280 and 199 of CS and CI2401, respectively (Fig. 1B). Detailed analysis of the variable region in CS positioned the CS-specific insert ~12.7 kb upstream of the *EH2* gene and identified it as a recombined copy of a long terminal repeat retroelement from the *Ty3/gypsy* superfamily (Fig. 2), which was probably present as the native single-copy element in CI2401. The absence of the

duplicated retroelement in CI2401 was validated by PCR with primers designed for diagnostic SNPs in long terminal repeats of the retroelements (Supplemental Table S2), which confirmed the presence of this duplicated recombined element in Glupro. Apart from this rearrangement, the alignment of the optical maps did not reveal other structural variations, indicating that the list of genes obtained from the CS sequence was most likely to be complete.

Narrowing Down the *Dn2401* Interval

In a previous study on the F_2 population of 184 individuals, we identified *Xowm705* and *Xowm711* as markers flanking the *Dn2401* gene and delimiting an interval of 0.83 cM (Staňková et al., 2015). To further increase the resolution of the genetic map, we expanded the mapping population by an additional 149 F_2 individuals and identified three new recombinants for the interval delimited by *Xowm705* and *Xowm711*. The new data extended the interval to 0.90 cM and localized *Dn2401* as 0.60 cM distal to *Xowm705* and 0.30 cM proximal to *Xowm711* (Fig. 3).

To further narrow down the *Dn2401* region, 46 new pairs of primers were designed and, after optimization, 44 primer pairs (96%) provided a single 7DS-specific PCR product. Sequencing of 41 amplicons from both parents did not detect any polymorphism, whereas one amplicon provided a mixed sequence, probably caused by sequence duplication within the 7DS. Only two amplicons yielded new SNPs, which were mapped in the complete CI2401 × Glupro F_2 population. The SNPs, named *Xowm713* and *Xowm714* (Table 2), were mapped into the *Dn2401* interval. *Xowm713* became a new proximal flanking marker, reducing the interval to 0.45 cM and 137.5 kb, whereas *Xowm714* was found to cosegregate with *Dn2401* (Fig. 3). The structural variation revealed in the retroelement upstream of the *EH2* gene with *STS2* primers (Fig. 2) could be mapped as a presence–absence polymorphism in the mapping population. This new marker, labeled *Xowm715*, was found to cosegregate with *Xowm714* and *Dn2401*. Finally, we mapped the SNP identified in the 3'UTR of the *EH2* gene between *Xowm711* and *Xowm715*. This novel distal flanking marker, named *Xowm716* (Table 2), delimited a reduced *Dn2401* interval of 0.3 cM and 133.2 kb (Fig. 3).

DISCUSSION

Sequencing and Analysis of the *Dn2401* Region

The *Dn2401* region proved to be highly challenging for several sequencing technologies and approaches applied. The biggest obstacle to complete and correct assembly was the duplicated recombined retroelement (Fig. 2), which was found to be misassembled in the whole-genome assembly generated from short-read Illumina data (International Wheat Genome Sequencing Consortium, 2018) and could not be resolved even after reducing sample complexity to a single BAC clone, despite applying the several assemblers including SASSY and RAY, which, in combination, succeeded in assembling the remaining part of the interval. We expected that the duplication might be resolved in the hybrid assembly of Zimin et al. (2017), which combined short-read Illumina data with long-read data obtained through PacBio sequencing. Surprisingly, optical mapping indicated even more discrepancies in the Triticum 3.1 assembly of the *Dn2401* region. The entire *Dn2401* interval could be resolved only after adding long nanopore reads, which enabled a reliable scaffolding of precise Illumina reads.

Bacterial artificial chromosome libraries, used previously for generating reference genomes, are available for many important crops, including maize (*Zea mays* L.) (Schnable et al. 2009), barley (IBGSC, 2012), and wheat (International Wheat Genome Sequencing Consortium, 2018). Chromosome- or chromosome-arm-specific BAC libraries have been constructed for all chromosomes of CS bread wheat (Šafář et al. 2010; International Wheat Genome Sequencing Consortium, 2018) and can be screened easily through pools of BAC clones (http://cnrgv.toulouse.inra.fr/en/library/wheat, accessed 14 Feb. 2019). Moreover, short-read assemblies of BAC clones have been generated for several chromosomes of CS and are publicly accessible (International Wheat Genome Sequencing Consortium, 2018; Tulpová et al. 2019). The availability of the BAC resources makes the precision resequencing of a region of interest (as used in the present work) widely applicable. In previous wheat cloning studies, the BAC-based approach was coupled with the PacBio technique rather than nanopore sequencing (Schweiger et al., 2016; Roselli et al., 2017). This alternative is handicapped by shorter reads [e.g., Zimin et al., (2017) reported that the the sequence length of the shortest contig at 50% of the total genome length was ~10 kb] that hardly span the entire length of a BAC insert, typically exceeding 100 kb. Schweiger et al. (2016) sequenced 10 wheat BAC clones spanning their region of interest with Roche 454 and PacBio technologies, but they were unable to produce a continuous assembly because of two gaps of several kb around transposable elements. The MinION sequencer is highly affordable (https://nanoporetech.com/products/minion, accessed 14 Feb. 2019) and relatively undemanding in terms of DNA input (less than 1µg per sequencing library), which makes it a favorable platform for small-scale in-house sequencing of specific BAC clones.

Optical BNG maps proved to be a useful and cost-efficient tool for overall sequence validation and the identification of structural variations in the region of interest. The minimum size of indels that can be resolved by optical mapping ranges from 1 kbp (Zhu et al., 2018) to as little as 500 bp (https://bionanogenomics.com/technology/dls-technology/, accessed 14 Feb. 2019) for the Bionano Genomics Irys and Saphyr platforms, respectively. To generate BNG maps, we took advantage of complexity reduction by flow-sorting the 7DS arm from wheat ditelosomic lines carrying 7DS arms as telocentric chromosomes. Preparation of a particular telosomic line from a cultivar of interest is technically simple but requires an unpredictable amount of cytology, as the misdivision frequencies

of individual chromosomes vary and cannot be gauged in advance. In several applications of the flow-sorted chromosomes, the use of telosomics has been circumvented by applying FISHIS (Giorgi et al. 2013), which labels DNA at trinucleotide microsatellites, typically GAA. This enables the flow-sorting of specific wheat chromosomes from the majority of wheat accessions with high purity (Vrána et al., 2016). In a pilot experiment, FISHIS labeling turned out to be problematic for BNG mapping with the Nt.BspQI nickase (unpublished data, 2016), perhaps because of an overlap of GAA with the Nt.BspQI recognition site. It is yet to be determined if the FISHIS procedure interferes with the Direct Label and Stain chemistry, introduced recently by Bionano Genomics as an alternative to nickase-based labeling (Deschamps et al., 2018). Whole-genome BNG maps are available for reference genomes of all major cereal crops, including rice (*Oryza sativa* L.) (Chen et al., 2017), maize (Jiao et al., 2017), barley (Mascher et al., 2017), and wheat (Luo et al. 2017); all of them have been generated on the Irys platform of Bionano Genomics. The new Bionano platform Saphyr, which has improved throughput and image resolution, mitigates the need for the complexity reduction and allows analyses of Gbp-sized genomes of interest in their entirety in a relatively short time and at a reasonable cost.

The *Dn2401* Candidate

Annotation of the *Dn2401* region delimited by the markers *Xowm711* and *Xowm713* indicated the presence of seven high confidence genes (Table 1; Supplemental Fig. S2); of these, *HVA22* (*TraesCS7D01G224100*), *EH2* (*TraesCS7D01G224200*), and *2OG-Fe(II) oxygenase* (*TraesCS7D01G224500*) were considered as potential resistance gene candidates on the basis of their preliminary functional analysis. Mapping of a new marker, *Xowm716*, into the region shifted *HVA22* further away from the *Dn2401* interval. Furthermore, we revealed that the resistant CI2401 and the susceptible CS shared the same haplotype in the *HVA22* region, which excluded *HVA22* as the *Dn2401* candidate. In contrast, the haplotype in the *EH2* region is common to the susceptible cultivars and is different from that of the resistant line CI2401. *EH2* was among the genes showing increased expression after aphid feeding in the RWA-resistant line PI220127, which bears the *Dnx* resistance gene in chromosome arm 7DS (Boyko et al., 2006). Both *Dn2401* and *Dnx* belong to a group of RWA resistance genes linked with the marker *Xgwm111* (Liu et al. 2001; Fazel-Najafabadi et al., 2015), which might be allelic (Liu et al. 2002). This makes *EH2* a robust candidate for the *Dn2401* gene.

Epoxide hydrolases (EHs) are enzymes present in all living organisms; they transform epoxide-containing lipids through the addition of water. Plants contain multiple soluble EH isoforms, which can be both constitutive and infection-induced (Newman et al., 2005). The transcription of the inducible enzymes can be increased by exogenous exposure to hormones, including the host-defense regulator methyl jasmonate (Stapleton et al., 1994). The substrate specificity and regulatory behavior of soluble plant EHs argue for a primary function of this enzyme in host defense and growth. The defensive functions can be related to both passive (cutin biosynthesis) and active (antifungal chemical synthesis) roles (Newman et al. 2005). The effect of the cuticle lipid (wax and cutin) composition has been considered in plant defense against both pathogens and pests. Studies on wheat's response to feeding by Hessian fly (*Mayetiola destructor*) larvae suggested cuticle composition and integrity are important components of host resistance (Kosma et al. 2010; Khajuria et al. 2013). The involvement of EHs in cutin biosynthesis matches well with the tolerance type of RWA resistance identified in CI2401 by Dong et al. (1997) and Voothuluru et al. (2006) and assayed in the present work.

We were also intrigued by a role of EHs in insects, whose juvenile hormone EH irreversibly degrades juvenile hormones, which is essential for insect metamorphosis. The gene for juvenile hormone EH was thus proposed as a promising target for hemipteran pest management (Tusun et al., 2017). In aphids, with their reproductive polyphenism (i.e., alternating reproductive modes from parthenogenesis to sexual reproduction in response to short photoperiods), the juvenile hormone degradation pathway appears to be involved in regulating the transition from asexual to sexual reproduction (Ishikawa et al., 2012). Interestingly, the EHs of cress and potato were shown to efficiently hydrolyze insect juvenile hormones (Morisseau et al., 2000), which lets us speculate about a possible pest–host interaction that could explain antibiosis (reduced populations) observed in the CI2401 line (Voothuluru et al., 2006).

During the analysis of the *EH2* locus, we resequenced a total of 6109 bp (Supplemental Fig. S3) from the resistant CI2401 and the susceptible Glupro and compared them with the reference genome of the susceptible CS. The only polymorphism discovered was a single SNP located in the 3'UTR, 290 bp downstream of the gene. A larger structural variation (a 8.2-kb indel) between the resistant line CI2401 and the susceptible cultivars CS and Glupro was identified 12.7 kb upstream of the gene. Both polymorphisms were found to cosegregate with the trait. It is yet to be determined by functional analysis if the SNP in the 3'UTR or the indel in the upstream region affect the expression of *EH2* and thus contribute to the resistance of CI2401.

The gene coding for the 2OG-Fe(II) oxygenase (*TraesCS7D01G224500*) was selected as a potential candidate because of its reported overexpression in response to a biotic stress, specifically, wheat infection by powdery mildew (*Blumeria graminis* f. sp. *tritici*) and stripe rust (*Puccinia striiformis* f. sp. *tritici*) (Zhang et al., 2014). On the other hand, no pathogen-dependent expression patterns were observed after inoculation of a resistant wheat line with *Fusarium* head blight (Schweiger et al., 2016), which suggested that 2OG-Fe(II) oxygenase was not a universal component of plant defense pathways. This, together with the absence of polymorphisms in the coding

and regulatory regions of the gene, reduce the chance of *2OG-Fe(II) oxygenase* being the true *Dn2401* candidate. As to the remaining genes in the interval, we have not found any published evidence of their changed expression in response to aphid feeding (Boyko et al., 2006; Smith and Boyko, 2007; Botha et al., 2014). The comprehensive wheat transcriptomics database built on the RNA-seq data (http://www.wheat-expression.com, accessed 14 Feb. 2019) (Ramírez-González et al., 2018) does not currently include any pest-response related studies. However, it comprises data from multiple projects on host–pathogen interactions, which did not indicate any significant pathogen-related response for the genes *TraesCS7D01G224300* (aspartic peptidase), *TraesCS7D01G224400* (a bacterial trigger factor protein), *TraesCS7D01G224600* (nuclear pore complex protein), and *TraesCS7D01G224700* (dual specificity protein phosphatase) located in the *Dn2401* interval. Previous transcriptomic studies on wheat's response to RWA infestation were based on subtractive hybridization, amplified fragment length polymorphism based transcript profiling (cDNA-AFLP), and real-time quantitative PCR on a selected gene set, and microarray hybridization, which may provide limited information. We expect that real-time quantitative PCR analysis of all genes from the critical interval will provide more comprehensive data to ultimately resolve the true *Dn2401* candidate.

Author Contributions

ZT, HT, KH, JM, and PN were involved in sequencing and assembling the *Dn2401* region. The annotation was done by PL, JM, and ZT. The 7DS ditelosomic line from CI2401 was prepared by AJL and DK and the optical map was generated by HT. Flow-sorting of particular chromosome arms was done by JV. Analysis of structural variation was performed by ZT. Marker development and genetic mapping were done by ZT and MM. Phenotyping was conducted by NLVL and FBP. The manuscript was drafted by ZT and HŠ, who conceived and coordinated the project. JD participated in project organization and manuscript writing.

Supplemental Information

Supplemental Table S1: BNG map of the 7DS arm from CI2401: preparation and statistics.

Supplemental Table S2: Primers and PCR conditions applied for interrogating local structural variations in the *Dn2401* region.

Supplemental Table S3: Primers and PCR conditions applied for resequencing of the candidate genes from parental lines.

Supplemental Fig. S1: Comparison of a short-read and a hybrid BAC assembly of the *Dn2401* region. In silico Nt.BspQI digests of a short-read and a hybrid assembly were aligned to the 7DS BNG map of CS. A 14.4-kb gap identified in the short-read assembly has been closed in the hybrid assembly. The position of the *EH2* gene is indicated by a red asterisk.

Supplemental Fig. S2: Distribution of coding sequences in the wider *Dn2401* region in CS. The wider region, localized at ~185 Mb of the IWGSC RefSeq version 1.0 7D pseudomolecule (grey bars), was delimited by the markers *Xowm711* and *Xowm705* and comprised 13 high-confidence coding sequences (dark and light green arrowheads), listed in Table 1. Mapping of new flanking markers, *Xowm713* and *Xowm716*, reduced the number of coding loci to six (dark green). Colors assigned to particular markers indicate the identity of the CS allele with that of CI2401 (blue) or Glupro (red).

Supplemental Fig. S3: Analysis of the DNA sequence in the *EH2* region. Sequences of the CS BAC *TaaCSp-7DS086H04*, CI2401 and Glupro were compared and a single SNP 290 bp downstream from the gene was identified. Susceptible cultivars CS and Glupro share the same SNP variant.

Supplemental Fig. S4: Resequencing of loci for 2OG-Fe(II) oxygenase and bacterial trigger factor protein. Green boxes represent particular amplicons with designations corresponding to primers stated in Supplemental Table S3. The blue box depicts the SNP marker *Xowm714*, the single polymorphism in the sequenced region.

Supplemental File S1: Hybrid assembly of the BAC clone *TaaCSp7DS086H04* and annotation of the *EH2* region.

Conflict of Interest Disclosure

The authors declare that there is no conflict of interest.

ACKNOWLEDGMENTS

We thank Prof. B.S. Gill for providing seeds of the wheat ditelosomic lines. We acknowledge the excellent assistance of Romana Šperková, Eva Jahnová, Radka Tušková, and Marie Seifertová in chromosome sorting, preparation of chromosomal DNA, and BAC isolation. We also acknowledge the help of Hana Vanžurová with construction of the genetic map. The authors thank the IWGSC for prepublication access to IWGSC RefSeq version 1.0. This work was supported by the Czech Science Foundation (award No. P501/12/2554); the Czech Ministry of Education, Youth and Sports (National Program of Sustainability I, grant LO1204); and the Czech Academy of Sciences (RVO60077344).

REFERENCES

Abeel, T., T. Van Parys, Y. Saeys, J. Galagan, and Y. Van de Peer. 2012. GenomeView: A next-generation genome browser. Nucleic Acids Res. 40. e12. doi:10.1093/nar/gkr995

Anderson, G., D. Papa, J. Peng, M. Tahir, and N.L.V. Lapitan. 2003. Genetic mapping of *Dn7*, a rye gene conferring resistance to the Russian wheat aphid in wheat. Theor. Appl. Genet. 107:1297–1303. doi:10.1007/s00122-003-1358-1

Anderson, V.A., S.D. Haley, F.B. Peairs, L. van Eck, J.E. Leach, and N.L.V. Lapitan. 2014. Virus-induced gene silencing suggests that *(1,3;1,4)-b-glucanase* is a susceptibility factor in the compatible Russian wheat aphid– wheat interaction. Mol. Plant Microbe Interact. 27:913–922. doi:10.1094/MPMI-05-13-0141-R

Aykut Tonk, F.A., D. İştipliler, M. Tosun, F. Turanli, H. İlbi, and M. Çakir. 2016. Genetic mapping and inheritance of Russian wheat aphid resistance gene in accession IG 100695. Plant Breed. 135:21–25. doi:10.1111/pbr.12339

Berkman, P.J., P. Visendi, H.C. Lee, J. Stiller, S. Manoli, M.T. Lorenc, et al. 2013. Dispersion and domestication shaped the genome of bread wheat. Plant Biotechnol. J. 11:564–571. doi:10.1111/pbi.12044

Boisvert, S., F. Laviolette, and J. Corbeil. 2010. Ray: Simultaneous assembly of reads from a mix of high-throughput sequencing technologies. J. Comput. Biol. 17:1519–1533. doi:10.1089/cmb.2009.0238

Botha, A.M., Y. Li, and N.L.V. Lapitan. 2005. Cereal host interactions with Russian wheat aphid: A review. J. Plant Interact. 1:211–222. doi:10.1080/17429140601073035

Botha, A.M., L. van Eck, N.F.V. Burger, and Z.H. Swanevelder. 2014. Near-isogenic lines of *Triticum aestivum* with distinct modes of resistance exhibit dissimilar transcriptional regulation during *Diuraphis noxia* feeding. Biol. Open 3:1116–1126. doi:10.1242/bio.201410280

Boyko, E.V., C.M. Smith, V.K. Thara, J.M. Bruno, Y. Deng, S.R. Starkey, et al. 2006. Molecular basis of plant gene expression during aphid invasion: Wheat *Pto*- and *Pti*-like sequences are involved in interactions between wheat and Russian wheat aphid (Homoptera: Aphididae). J. Econ. Entomol. 99:1430–1445. doi:10.1093/jee/99.4.1430

Burd, J.D., and R.L. Burton. 1992. Characterization of plant damage caused by Russian wheat aphid (Homoptera: Aphididae). J. Econ. Entomol. 85:2017–2022. doi:10.1093/jee/85.5.2017

Chen, P., X. Jing, B. Liao, Y. Zhu, J. Xu, R. Liu, et al. 2017. BioNano genome map resource for *Oryza sativa* ssp. *japonica* and *indica* and its application in rice genome sequence correction and gap filling. Mol. Plant 10:895–898. doi:10.1016/j.molp.2017.02.003

Deschamps, S., Y. Zhang, V. Llaca, L. Ye, G. May, and H. Lin. 2018. A chromosome-scale assembly of the sorghum genome using nanopore sequencing and optical mapping. Nat. Commun. 9: 4844.

Dong, H., J.S. Quick, and Y. Zhang. 1997. Inheritance and allelism of Russian wheat aphid resistence in several wheat lines. Plant Breed. 116:449–453. doi:10.1111/j.1439-0523.1997.tb01029.x

Du Toit, F. 1987. Resistance in wheat (*Triticum aestivum*) to *Diuraphis noxia* (Homoptera: Aphididae). Cereal Res. Commun. 15:175–179.

Du Toit, F. 1989. Inheritance of resistance in two *Triticum aestivum* lines to Russian wheat aphid (Homoptera: Aphididae). J. Econ. Entomol. 82:1251–1253. doi:10.1093/jee/82.4.1251

Du Toit, F., W.G. Wessel, and G.F. Marais. 1995. The chromosome arm location of the Russian wheat aphid resistance gene, *Dn5*. Cereal Res. Commun. 23:15–17.

Fazel-Najafabadi, M., J. Peng, F.B. Peairs, H. Šimková, A. Kilian, and N.L.V. Lapitan. 2015. Genetic mapping of resistance to *Diuraphis noxia* (Kordjumov) biotype 2 in wheat (*Triticum aestivum* L.) accession CI2401. Euphytica 203:607–614. doi:10.1007/s10681-014-1284-0

Giorgi, D., A. Farina, V. Grosso, A. Gennaro, C. Ceoloni, and S. Lucretti. 2013. FISHIS: Fluorescence in situ hybridization in suspension and chromosome flow sorting made easy. PLoS One. doi:10.1371/journal.pone.0057994

Harvey, T.L., and T.J. Martin. 1990. Resistance to Russian wheat aphid, *Diuraphis noxia*, in wheat (*Triticum aestivum*). Cereal Res. Commun. 18:127–129.

International Barley Genome Sequencing Consortium. 2012. A physical, genetic and functional sequence assembly of the barley genome. Nature 491:711–716. doi:10.1038/nature11543

Ishikawa, A., K. Ogawa, H. Gotoh, T.K. Walsh, D. Tagu, J.A. Brisson, et al. 2012. Juvenile hormone titre and related gene expression during the change of reproductive modes in the pea aphid. Insect Mol. Biol. 21:49–60. doi:10.1111/j.1365-2583.2011.01111.x

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

Jiao, Y., P. Peluso, J. Shi, T. Liang, M.C. Stitzer, B. Wang, et al. 2017. Improved maize reference genome with single-molecule technologies. Nature 546:524–527.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, et al. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28:1647–1649. doi:10.1093/bioinformatics/bts199

Khajuria, C., H.Y. Wang, X.M. Liu, S. Wheeler, J.C. Reese, M. El Bouhssini, et al. 2013. Mobilization of lipids and fortification of cell wall and cuticle are important in host defense against Hessian fly. BMC Genomics 14:423. doi:10.1186/1471-2164-14-423

Kosma, D.K., J.A. Nemacheck, M.A. Jenks, and C.E. Wiliams. 2010. Changes in properties of wheat leaf cuticle during interaction with Hessian fly. Plant J. 63:31–43.

Lam, E.T., A. Hastie, C. Lin, D. Ehrlich, S.K. Das, M.D. Austin, et al. 2012. Genome mapping on nanochannel arrays for structural variation analysis and sequence assembly. Nat. Biotechnol. 30:771–776. doi:10.1038/nbt.2303

Lapitan, N.L.V., J. Peng, and V. Sharma. 2007. A high-density map and PCR markers for Russian wheat aphid resistance gene *Dn7* on chromosome 1RS/1BL. Crop Sci. 47:811–818. doi:10.2135/cropsci2006.08.0529

Leroy, P., N. Guilhot, H. Sakai, A. Bernard, F. Choulet, S. Theil, et al. 2012. TriAnnot: A versatile and high performance pipeline for the automated annotation of plant genomes. Front. Plant Sci. 3:5. doi:10.3389/fpls.2012.00005

Liu, X.M., C.M. Smith, B.S. Gill, and V. Tolmay. 2001. Microsatellite markers linked to six Russian wheat aphid resistance genes in wheat. Theor. Appl. Genet. 102:504–510. doi:10.1007/s001220051674

Liu, X.M., C.M. Smith, and B.S. Gill. 2002. Identification of microsatellite markers linked to Russian wheat aphid resistance genes *Dn4* and *Dn6*. Theor. Appl. Genet. 104:1042–1048. doi:10.1007/s00122-001-0831-y

Luo, M.C., Y.Q. Gu, D. Puiu, H. Wang, S.O. Twardziok, K.R. Deal, et al. 2017. Genome sequence of the progenitor of the wheat D genome *Aegilops tauschii*. Nature 551:498–502.

Ma, Z.Q., A. Saidi, J.S. Quick, and N.L.V. Lapitan. 1998. Genetic mapping of Russian wheat aphid resistance genes *Dn2* and *Dn4* in wheat. Genome 41:303–306. doi:10.1139/g98-013

Marais, G.F., M. Horn, and F. Du Toit. 1994. Intergeneric transfer (rye to wheat) of gene(s) for Russian wheat aphid resistance. Plant Breed. 113:265–271. doi:10.1111/j.1439-0523.1994.tb00735.x

Mascher, M., H. Gundlach, A. Himmelbach, S. Beier, S.O. Twardziok, T. Wicker, et al. 2017. A chromosome conformation capture ordered sequence of the barley genome. Nature 544:427–433. doi:10.1038/nature22043

Miller, C.A., A. Altinkut, and N.L.V. Lapitan. 2001. A microsatellite marker for tagging a wheat gene conferring resistance to the Russian wheat aphid. Crop Sci. 41:1584–1589. doi:10.2135/cropsci2001.4151584x

Morisseau, C., J.K. Beetham, F. Pinot, S. Debernard, J.W. Newman, and B.D. Hammock. 2000. Cress and potato soluble epoxide hydrolases: Purification, biochemical characterization, and comparison to mammalian enzymes. Arch. Biochem. Biophys. 378:321–332. doi:10.1006/abbi.2000.1810

Newman, J.W., C. Morisseau, and B.D. Hammock. 2005. Epoxide hydrolases: Their roles and interactions with lipid metabolism. Prog. Lipid Res. 44:1–51. doi:10.1016/j.plipres.2004.10.001

Peng, J., H. Wang, S.D. Haley, F.B. Peairs, and N.L.V. Lapitan. 2007. Molecular mapping of the Russian wheat aphid resistance gene *Dn2414* in wheat. Crop Sci. 47:2418–2429. doi:10.2135/cropsci2007.03.0137

Peng, J.H., Y. Bai, S.D. Haley, and N.L.V. Lapitan. 2009. Microsatellite-based molecular diversity of bread wheat germplasm and association mapping of wheat resistance to the Russian wheat aphid. Genetica (The Hague) 135:95–122.

Ramírez-González, R. H., Borrill, P., D. Lang, S. A. Harrington, J. Brinton, L. Venturini, et al. 2018. The transcriptional landscape of polyploid wheat. Science 361: eaar6089. doi:10.1126/science.aar6089

Roselli, S., A. Olry, S. Vautrin, O. Coriton, D. Ritchie, G. Galati, et al. 2017. A bacterial artificial chromosome (BAC) genomic approach reveals partial clustering of the furanocoumarin pathway genes in parsnip. Plant J. 89:1119–1132. doi:10.1111/tpj.13450

Šafář, J., H. Šimková, M. Kubaláková, J. Číhalíková, P. Suchánková, J. Bartoš, et al. 2010. Development of chromosome-specific BAC resources for genomics of bread wheat. Cytogenet. Genome Res. 129:211–223. doi:10.1159/000313072

Saidi, A., and J.S. Quick. 1996. Inheritance and allelic relationships among Russian wheat aphid resistance genes in winter wheat. Crop Sci. 36:256–258. doi:10.2135/cropsci1996.0011183X003600020007x

Schnable, P.S., D. Ware, R.S. Fulton, J.C. Stein, F. Wei, P. Pasternak, et al. 2009. The B73 maize genome: Complexity, diversity, and dynamics. Science 326:1112–1115. doi:10.1126/science.1178534

Schweiger, W., B. Steiner, S. Vautrin, T. Nussbaumer, G. Siegwart, M. Zamini, et al. 2016. Suppressed recombination and unique candidate genes in the divergent haplotype encoding *Fhb1*, a major *Fusarium* head blight resistance locus in wheat. Theor. Appl. Genet. 129:1607–1623. doi:10.1007/s00122-016-2727-x

Shen, Q., J.U. Scot, and T.H.D. Ho. 1993. Hormone response complex in a novel abscisic acid and cycloheximide-inducible barley gene. J. Biol. Chem. 268:23652–23660.

Šimková, H., J. Šafář, M. Kubaláková, P. Suchánková, J. Číhalíková, H. Robert-Quatre, et al. 2011. BAC libraries from wheat chromosome 7D: Efficient tool for positional cloning of aphid resistance genes. J. Biomed. Biotechnol. 2011:302543. doi:10.1155/2011/302543

Smith, C.M., and E.V. Boyko. 2007. The molecular bases of plant resistance and defense responses to aphid feeding: Current status. Entomol. Exp. Appl. 122:1–16. doi:10.1111/j.1570-7458.2006.00503.x

Staňková, H., M. Valárik, N.L.V. Lapitan, P.J. Berkman, J. Batley, D. Edwards, et al. 2015. Chromosomal genomics facilitates fine mapping of a Russian wheat aphid resistance gene. Theor. Appl. Genet. 128:1373–1383. doi:10.1007/s00122-015-2512-2

Staňková, H., R. Hastie, S. Chan, J. Vrána, Z. Tulpová, M. Kubaláková, et al. 2016. BioNano genome mapping of individual chromosomes supports physical mapping and sequence assembly in complex plant genomes. Plant Biotechnol. J. 14:1523–1531. doi:10.1111/pbi.12513

Stapleton, A., J.K. Beetham, F. Pinot, J.E. Garbarino, D.R. Rockhold, M. Friedman, et al. 1994. Cloning and expression of soluble epoxide hydrolase from potato. Plant J. 6:251–258. doi:10.1046/j.1365-313X.1994.6020251.x

Tulpová, Z., M.C. Luo, H. Toegelová, P. Visendi, S. Hayashi, P. Vojta, et al. 2019. Integrated physical map of bread wheat chromosome arm 7DS to facilitate gene cloning and comparative studies. N. Biotechnol. 48:12–19. doi:10.1016/j.nbt.2018.03.003

Tusun, A., M. Li, X.Z. Liang, T. Yang, B. Yang, and G.R. Wang. 2017. Juvenile hormone epoxide hydrolase: A promising target for hemipteran pest management. Sci. Rep. 7:789. doi:10.1038/s41598-017-00907-0 [erratum: 8: 6246]

Valdez, V.A., P.F. Byrne, N.L.V. Lapitan, F.B. Peairs, A. Bernardo, G. Bai, et al. 2012. Inheritance and genetic mapping of Russian wheat aphid resistance in Iranian wheat landrace accession PI626580. Crop Sci. 52:676–682. doi:10.2135/cropsci2011.06.0331

Van Eck, L., R.M. Davidson, S. Wu, B.Y. Zhao, A.M. Botha, J.E. Leach, et al. 2014. The transcriptional network of *WRKY53* in cereals links oxidative responses to biotic and abiotic stress inputs. Funct. Integr. Genomics 14:351–362. doi:10.1007/s10142-014-0374-3

Van Ooijen, J.W., and R.E. Voorrips. 2001. JoinMap 3.0, software for the calculation of genetic linkage maps. Plant J. 3:739–744.

Voothuluru, P., J. Meng, C. Khajuria, J. Louis, L. Zhu, S. Starkey, et al. 2006. Categories and inheritance of resistance to Russian wheat aphid (Homoptera: Aphididade) biotype 2 in a selection from wheat cereal introduction 2401. J. Econ. Entomol. 99:1854–1861. doi:10.1093/jee/99.5.1854

Vrána, J., P. Cápal, H. Šimková, M. Karafiátová, J. Čížková, and J. Doležel. 2016. Flow analysis and sorting of plant chromosomes. Curr. Protoc. Cytom. 78:5.3.1–5.3.43. doi:10.1002/cpcy.9

Walters, M.C., F. Penn, F. Du Toit, T.C. Botha, K. Aalberrsberg, P.H. Mewitt, and S.W. Broodryk. 1980. Russian wheat aphid. Farming in South Africa. Leaflet Series. Wheat G3: 1–6.

Yazdani, M., G. Baker, H. DeGraaf, K. Henry, K. Hill, B. Kimber, et al. 2017. First detection of Russian wheat aphid *Diuraphis noxia* Kordjumov (Hemiptera: Aphididae) in Australia: A major threat to cereal production. Aust. Entomol. 57:410–417. doi:10.1111/aen.12292

Zhang, H., Y. Yang, C. Wang, M. Liu, H. Li, Y. Fu, et al. 2014. Large-scale transcriptome comparison reveals distinct gene activations in wheat responding to stripe rust and powdery mildew. BMC Genomics 15:898. doi:10.1186/1471-2164-15-898

Zhu, T., Z. Hu, J.C. Rodriguez, K.R. Deal, J. Dvorak, J.P. Vogel, Z. Liu, and M.C. Luo. 2018. Analysis of *Brachypodium* genome with genome-wide optical maps. Genome 61(8):559–565.

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

Zimin, A.V., D. Puiu, R. Hall, S. Kingan, B.J. Clavijo, and S.L. Salzberg. 2017. The first near-complete assembly of the hexaploid bread wheat genome, *Triticum aestivum*. GigaSci. 6: gix097. doi:10.1093/gigascience/gix097