Theor Appl Genet (2015) 128:1343–1357
DOI 10.1007/s00122-015-2510-4

ORIGINAL PAPER

Fine mapping and identification of a candidate gene for the barley *Un8* true loose smut resistance gene

Wen Zang[1] · Peter E. Eckstein[1] · Mark Colin[1] · Doug Voth[1] · Axel Himmelbach[3] · Sebastian Beier[3] · Nils Stein[3] · Graham J. Scoles[1] · Aaron D. Beattie[2]

Received: 16 June 2014 / Accepted: 27 March 2015 / Published online: 17 April 2015

Abstract
Key message **The candidate gene for the barley *Un8* true loose smut resistance gene encodes a deduced protein containing two tandem protein kinase domains.**
Abstract In North America, durable resistance against all known isolates of barley true loose smut, caused by the basidiomycete pathogen *Ustilago nuda* (Jens.) Rostr. (*U. nuda*), is under the control of the *Un8* resistance gene. Previous genetic studies mapped *Un8* to the long arm of chromosome 5 (1HL). Here, a population of 4625 lines segregating for *Un8* was used to delimit the *Un8* gene to a 0.108 cM interval on chromosome arm 1HL, and assign it to fingerprinted contig 546 of the barley physical map. The minimal tilling path was identified for the *Un8* locus using two flanking markers and consisted of two overlapping bacterial artificial chromosomes. One gene located close to a marker co-segregating with *Un8* showed high sequence identity to a disease resistance gene containing two kinase domains. Sequence of the candidate gene from the parents of the segregating population, and in an additional 19 barley lines representing a broader spectrum of diversity, showed there was no intron in alleles present in either resistant or susceptible lines, and fifteen amino acid variations unique to the deduced protein sequence in resistant lines differentiated it from the deduced protein sequences in susceptible lines. Some of these variations were present within putative functional domains which may cause a loss of function in the deduced protein sequences within susceptible lines.

Communicated by D. E. Mather.

✉ Aaron D. Beattie
aaron.beattie@usask.ca

[1] Department of Plant Sciences, University of Saskatchewan, Saskatoon, SK, Canada

[2] Crop Development Centre, University of Saskatchewan, Saskatoon, SK, Canada

[3] Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany

Introduction

True loose smut of barley (*Hordeum vulgare* L.) is caused by the basidiomycete pathogen *Ustilago nuda* (Jens.) Rostr. (*U. nuda*). This seed-borne disease has been reported in 50–70 % of fields in the prairie provinces of Western Canada (Tekauz 2003) and is also common in the United States (Menzies et al. 2010). Yield reduction due to *U. nuda* infection is commonly less than 1 %, however over 10 % yield loss has been reported (Mathre 1997; Popovic et al. 1998). After colonization of the florets, *U. nuda* can overwinter in the embryo of mature seeds as dormant mycelium. Upon seed germination, the pathogen will colonize tissue behind the growing point of the barley host and eventually infect the flower where the florets are replaced with the distinctive black teliospore masses which serve as the next source of inoculum (Mathre 1997).

Among the common disease control strategies, plant resistance is the most economical and effective strategy. The first resistance gene, *Un*, was reported in the cv. 'Trebi' in the 1940s (Livingston 1942; Robertson et al. 1947). Since then, a total of 15 resistance genes (Livingston 1942; Robertson et al. 1947; Schaller 1949; Skoropad and Johnson 1952; Andrews 1956; Metcalfe and Johnston 1963; Metcalfe 1966) associated with true loose smut resistance have since been identified, with *Un11*, *Un12*, *Un13*, and *Un15* being the most recent additions (Mueller 2006). Among these resistance genes, *Un8*, which was found in

the cv. 'PR28' derived from the winter barley line C.I. 4966 (Metcalfe and Johnston 1963; Metcalfe 1966), is the most effective and long-lived resistance, effective against all known true loose smut isolates in Western Canada (Thomas and Menzies 1997).

Breeding for resistance to true loose smut involves individual hand inoculation of florets at early anthesis and evaluating the phenotype at heading in the following growing season. This process is both labour-intensive and time-consuming and moreover, the occurrence of false negatives (escapes) resulting from the failure of artificial inoculation, necessitates several rounds of screening to ensure the presence of resistance. Molecular marker-assisted selection (MMAS) for true loose smut resistance is one of the best examples of how markers can improve selection since significant increases in efficiency and accuracy are achievable.

Un8 was initially mapped onto the long arm of barley chromosome 5 (1HL) in linkage with the ABC 261 RFLP marker (Eckstein et al. 1993). Subsequently, microsatellite (Li et al. 2001) and sequence characterized amplified region (SCAR) (Eckstein et al. 2002) markers were developed for *Un8*.

When dense unigene-based single nucleotide polymorphism (SNP) maps became available for barley (Close et al. 2009) and the accompanying information deposited into databases (e.g. HarvEST:Barley), it was possible to develop additional markers for *Un8*. Barley unigenes 4245, 16527 and 14722 (HarvEST:Barley v. 1.83, assembly 35) were used to create three TaqMan® assays which defined a region of approximately 6.2 cM around *Un8* (P.E. Eckstein, personal communication). After several years of MMAS in the Crop Development Centre barley breeding programme in which over 20,000 lines were evaluated, it became apparent that despite the close linkage of the *Un8* markers initially identified, there was a higher recombination frequency in this region of the barley genome than indicated by the smaller populations used originally to define the *Un8* region (Eckstein et al. 2002). As a result, the usefulness of the markers was reduced.

Recent advances towards understanding the barley genome have provided a number of avenues to identify molecular markers in tighter linkage to the *Un8* gene. The existence and defining of micro-colinearity between barley and other model species, such as rice (*Oryza sativa* L.) and *Brachypodium distachyon* L. Beauv (Brachypodium) (Mayer et al. 2011), permit the use of genomic sequence available in the syntenic regions for additional molecular marker development. This strategy has been exploited in barley to fine map the *sdw3* semi-dwarfing gene (Vu et al. 2010), *dsp* spike density gene (Shahinnia et al. 2012), two novel QTL (Silvar et al. 2012) and *Ror1* (Acevedo-Garcia et al. 2013) conferring powdery mildew resistance, the BaMMV/BaYMV resistance gene *rym11* (Lüpken et al. 2013), *HvNax3* (Shavrukov et al. 2013) and *HvNax4* (Rivandi et al. 2011) which limit Na^+ accumulation, and *Ryd3* controlling tolerance to barley yellow dwarf virus (Lüpken et al. 2014). With respect to the barley *Un8* region, the syntenic regions in Brachypodium and rice are chromosomes 2 and 5, respectively (Mayer et al. 2011). Assembly of the 5.1 Gb barley genome which integrates physical and genetic information together with gene expression and bacterial artificial chromosome (BAC) clones (The International Barley Genome Sequencing Consortium 2012) provides a valuable tool for not only marker development, but also for the positional cloning of the *Un8* gene.

The objective of this study was to enrich the 6.2 cM interval harbouring the *Un8* true loose smut resistance gene using a variety of strategies (EST data, SNP genotyping arrays, synteny, and BAC and whole-genome sequence data) which take advantage of the array of genomic tools available in barley, and to identify the candidate gene(s) for *Un8*. This would allow us to accomplish two goals: (1) develop perfect markers diagnostic for the presence of the *Un8* gene which would assist our MMAS programme and (2) characterize the durable *Un8* resistance gene candidate.

Materials and methods

Plant materials and mapping populations

An F_4 recombinant inbred line (RIL) population (4625 lines) derived from the cross TR09398 × TR07728 was used for genetic mapping. After the initial cross the population was advanced from the F_1 to F_4 generations using the bulk breeding method. The F_4 RILs used in this study were randomly selected individual seeds from the larger F_4 bulk seed sample (comprised of ~150,000 seeds). Information about the parents of this population is provided in Table 1. The F_4 population was screened with two co-dominant flanking TaqMan® markers, *Un8* SNP1 and *Un8* SNP6, using the Applied Biosystems® StepOnePlus™ Real-Time PCR System. Lines were selected if they showed recombination between *Un8* SNP1 and *Un8* SNP6 and were homozygous for both markers. Genomic DNA from 122 such lines was isolated from leaves using the modified cetyltrimethylammonium bromide (CTAB) method (Procunier et al. 1991). As well, seeds from each line were harvested separately to evaluate their reaction to true loose smut. The phenotypic and genotypic information of the 'Harrington' (susceptible) × TR306 (resistant) doubled-haploid (DH) population (149 lines; Eckstein et al. 2002) was also used in this study to help position newly developed markers. An additional 19 barley lines of diverse of origins (Table 1) were sequenced for the *Un8* candidate gene and phenotyped for their reaction to true loose smut.

Table 1 Origin, true loose smut reaction and *Un8* candidate gene allele carried by the F_4 recombinant inbred line mapping population parents (TR09398 and TR07728) and 19 barley lines of diverse origin

Line	Pedigree	Origin	True loose smut reaction[a]	*Un8* candidate gene allele group[b]
TR09398	TR238/Wpg8412-9-2-1//'Baronesse'/TR336	Canada	R	Group I
TR11698	'Ponoka'/H93102002	Canada	R	
TR12135	TR253/BM9216-4//SM04261	Canada	R	
HB11316	'CDC Rattan'/SH041242	Canada	R	
'AC Metcalfe'	'AC Oxbow'/'Manley'	Canada	R	
TR07728	'Salute'/'Xena'	USA	S	Group II
'Bowman'	'Klages'//'Fergus'/'Nordic'/3/ND1156/4/'Hector'	USA	S	
'OAC 21'[c]	Selection from manchurian introduction	Canada	S	
'Morex'[c]	'Cree'/'Bonanza'	USA	S	
TR12737	'Xena'/'Sebastian'	USA	S	
'Calcule'	97-7207-484/'Zenobia'	Germany	S	
'Streif'	'Pasadena'/'Aspen'	Germany	S	
'Barke'	'Libelle'/'Alexis'	Germany	S	Group III
'Baudin'	'Stirling'/'Franklin'	Australia	S	
'Quilmes Carisma'	'Femina'/O6306//L5184/'Prisma'	South America	S	
'Champlain'[c]	'Moore'/'Montcalm'	Canada	S	
'Montcalm'[c]	Michigan 31604/Common 6-Rowed 4307 MC// Mandscheuri 1807 MC	Canada	S	
'Jet'	Ethiopian introduction	Ethiopia	S	
'Quilmes Ayelén'	G6066/'Quilmes Alfa'	South America	S	Group IV
'CDC Austenson'	TR128//TR236/WM862-6/3/94Ab12271	Canada	S	
'Optic'	'Corniche'/'Force'//'Chad'	UK	S	

R indicates resistance to true loose smut, *S* indicates susceptibility to true loose smut

[a] Phenotypic reaction to true loose smut evaluated by artificial inoculation as described by Eckstein et al. (2002)

[b] The Group I allele is present in barley lines resistant to true loose smut while the Group II–IV alleles are present in barley lines susceptible to true loose smut

[c] Six-row barley type. All other lines are two-row barley types

This was done to identify polymorphisms which consistently differentiated alleles present in resistant lines from susceptible lines to provide further evidence for the candidate gene being *Un8*.

Evaluation of true loose smut reaction

Reaction to true loose smut was tested at flowering, as previously described by Eckstein et al. (2002), in the greenhouse and field by artificial inoculation using a mixture of true loose smut pathotypes. All lines were evaluated at the North Seed Farm disease nursery (Saskatoon, SK, Canada) where lines were planted as hill plots (15 seeds/hill) (with susceptible checks throughout the nursery). In the greenhouse, three seeds of each line were sown in a pot. The cultivar 'CDC Austenson' was used as a susceptible control in the greenhouse experiments. For both field and greenhouse inoculations, 6–8 heads were inoculated at anthesis using a 3-ml syringe and at least 15 inoculated seeds were tested for disease reaction in the following generation. If a line showed susceptibility to the disease (i.e. smutted heads were observed) no further testing was done. If a line showed resistance (i.e. no smutted heads were observed), two additional inoculations were conducted to confirm the resistance. A goodness of fit to a 1:1 ratio (resistant to susceptible) was tested using the Chi-squared test (χ^2).

Marker development and genotyping

Because *Un8* was previously assigned to chromosome arm 1HL between markers *Un8* SNP1 and *Un8* SNP6 (P. Eckstein, personal communication) which were designed based on EST unigene sequences 4245 and 14722, respectively, (HarvEST:Barley version 1.83, assembly 35), other unigenes located on the barley integrated map (HarvEST:Barley) within the interval flanked by these two markers were explored for marker development. EST unigene sequences were extracted from HarvEST:Barley,

formatted as FASTA files and used to query the barley cv. 'Morex' whole genome assembly using the BLASTN basic search programme within the ViroBLAST interface tool (http://www.webblast.ipk-gatersleben.de/barley/viroblast.php). Genomic DNA contig sequences identified through these queries were used to assist marker development.

The 9K Barley iSELECT Infinium SNP Assay was also used to develop markers in the *Un8* target region. Three true loose smut resistance resources ('CDC Meredith', TR306 and TR09398) and four susceptible sources (TR07728, TR09397, 'Harrington' and 'CDC Kindersley') were genotyped with the 9K assay. Available information for sequence surrounding the SNPs which differentiated resistant from susceptible lines was used to identify additional sequence information contained in 'Morex' BACs for marker development.

The syntenic relationship of barley with model species was exploited to identify markers in the *Un8* interval. Based on information within HarvEST:Barley, the putative orthologous genes to the genes from which *Un8* SNP1 and *Un8* SNP6 were developed were identified in rice (Os05g48422 and Os05g49030, respectively) and Brachypodium (Bradi2g16930 and Bradi2g16430, respectively). Once this region was defined in both reference genomes, all gene sequences located in the syntenic region were extracted from rice (http://www.ricemap.org/) and Brachypodium (http://www.brachypodium.org/g-mod/genomic/contigs), respectively. The gene sequences were queried against the barley EST database in HarvEST:Barley using the BLASTN function (E value $\leq e^{-10}$ and identity ≥ 80 %) to find the putative orthologous unigenes in barley. Barley unigenes identified in this manner were also queried against the barley cv. 'Morex' whole genome assembly using the ViroBLAST interface tool (as mentioned above) to identify genomic DNA contig sequence for further marker development (and to confirm the 1HL chromosome arm location).

After initial analysis of the phenotypic and genotypic data generated from the TR09398 × TR07728 population, it was determined that *Un8* was close to the *Un8* SNP4 marker. Based on that information the HarvEST:Barley database was queried to identify BAC clones spanning the *Un8* SNP4 marker. BAC clones HVVMRXALLhA0751D06 and HVVMRXALLhA0772N02 were found to co-locate with *Un8* SNP4, while BAC HVVMRXALLhA0498L15 was located only 0.7 cM away from *Un8* SNP4 on the barley integrated map (HarvEST:Barley). BAC clone sequences were downloaded from HarvEST:Web (http://www.harvest-web.org/hweb/pickassy.wc) for marker development.

Once genomic DNA sequence was obtained using the above strategies, PCR primers were designed using Primer Premier 5.0 (PREMIER Biosoft International, Paulo Alto, CA, USA) to amplify a fragment of the genomic DNA to identify polymorphisms between TR09398 and TR07728. Standard PCR amplifications were performed in a 25 µl volume containing 1× Lucigen® PCR buffer, dNTPs (100 µM each), primers (0.2 µM each), 50 ng genomic DNA, and 1 U Taq DNA polymerase. PCR conditions were: 5 min at 94 °C for initial denaturation, followed by 35 cycles of 94 °C for 45 s, 55–65 °C for 45 s, 72 °C for 1 min, and a final extension step at 72 °C for 5 min. The amplification products were separated on 1 % agarose gels and 5–6 clones of each amplicon were cloned into the TOPO® TA® Cloning Vector, Sanger Sequenced at the National Research Council (Saskatoon, SK, Canada) and aligned using DNAMAN v. 7 (Lynnon Biosoft, San Ramon, CA, USA) to ensure the consistency of the sequence data. All allele-specific, amplicon size shift and cleaved amplified polymorphic sequence (CAPS) markers developed were run under the standard PCR conditions listed above. Amplicon size shift markers resulted from the fortuitous design of the original PCR primers. Allele-specific markers were created by designing new PCR primers targeted against SNP sites identified in the originally sequenced PCR products. The allele-specific primers were designed with an additional mismatch nucleotide introduced into the third nucleotide position from the SNP site at the 3′-end of the primer, according to the method described by Liu et al. (2012), to increase the SNP detection efficiency. For CAPS markers, PCR products were subsequently digested with 2 U of restriction endonuclease (NEB) corresponding to the SNP site identified in the originally sequenced PCR product. PCR products for all markers were separated on 1.5 % agarose gels.

TaqMan® assays were developed to target SNP sites identified in the originally sequenced PCR products. TaqMan® SNP genotyping was performed with the ABI StepOnePlus™ Real-Time PCR System in a 10 µl volume which included 1× ABI TaqMan® GTXpressTM Master Mix, 0.36× ABI TaqMan® SNP Genotyping assay and 25 ng genomic DNA. PCR conditions were: 30 s at 60 °C for pre-PCR read and 10 min at 94 °C for hot-start activation, followed by 40 cycles at 94 °C for 15 s, 60 °C for 30 s and 60 °C for 30 s for post-PCR read.

Linkage analysis

Linkage analysis was carried out by screening all the newly developed markers on the 122 F_4 lines derived from the TR09398 × TR07728 population which had been preselected for recombination between *Un8* SNP1 and *Un8* SNP6. Genetic distance was estimated according to the Kosambi mapping function (Kosambi 1944) based on a population size of 4625 lines. This was the number of lines remaining after removing 211 lines which displayed a recombination between *Un8* SNP1 and *Un8* SNP6 in a

Table 2 Information on PCR primers used to amplify the *Un8* candidate gene from 21 barley lines of diverse origin differing in their reaction to true loose smut

Name	Forward primer sequence (5′–3′)	Reverse primer sequence (5′–3′)	Target region[a]
0751D06_1	GATCCTCGTGTACTCAGGTCTC	AAATGCTACTCCAGGCTACG	−292 to +1039
0751D06_2	ACAAACAGGTTGGCGATTC	TGACATCTTTGTGGCATTACC	+957 to +2103

[a] Indicates the region of the *Un8* candidate gene. Positive numbers refer to the nucleotide position relative to the adenosine nucleotide of the translation start codon (position 1). Position numbers are specific for the resistant allele in which the coding region was 2037 nucleotides in length

heterozygous state (i.e. only one homologous chromosome was recombinant), from the original population of 4836 lines evaluated for recombination between the two markers. The 211 lines were used for a second calculation of total genetic distance between *Un8* SNP1 and *Un8* SNP6, but since these lines were discarded after the initial screening with *Un8* SNP1 and *Un8* SNP6, they could not be evaluated with the subsequent markers developed and thus they did not contribute to the linkage map created. The linkage map was constructed with JoinMap 4.0 (Kyazma B.V., Wageningen, The Netherlands).

Physical map construction, BAC sequencing and assembly

Two markers, *Un8* SNP4 and 0498L15 F3/R3, were utilized to identify the fingerprinted contigs (FPC) in the physical map of barley (The International Barley Genome Sequencing Consortium 2012; Ariyadasa et al. 2014) which encompassed the *Un8* gene. The minimal tilling path (MTP) of the targeted FPC was then identified.

Shotgun sequencing of DNA from 'Morex' BACs HVVMRXALLmA0180J17 and HVVMRXALLeA0154F16, which composed the MTP of the targeted FPC, was performed using the Illumina HiSeq 2000 (2 × 100 cycles) device essentially as described (Meyer and Kircher 2010). Individual assemblies for the targeted BACs were produced with clc Assembly Cell version 4.0.6 beta.

Nextera mate pair sequencing libraries with insert sizes ranging between 3 and 10 kb were prepared following the instructions of the manufacturer (Illumina) and sequenced using the Illumina MiSeq (2 × 250 cycles) and HiSeq2000 (2 × 100 cycles) devices. Shotgun assemblies were scaffolded with mate pairs using SSPACE PREMIUM version 2.3.

Gene prediction and annotation and protein domain annotation

Identification of all putative gene sequences within 'Morex' BAC clones comprising the MTP was accomplished using several methods. GeneMark (http://www.opal.biology.gatech.edu/GeneMark/eukhmm.cgi) and GENSCAN (http://www.genes.mit.edu/GENSCAN.html) were used to locate possible genes. In addition, predicted genes from the recently released barley genome assembly (http://www.barleyflc.dna.affrc.go.jp/hvdb/index.html) and HarvEST Barley (v. 1.98, assembly 37) were identified and compared with predictions from GeneMark and GENSCAN. Predicted genes were annotated using the BLASTP tool to query the NCBI and iTAK (plant transcription factor and protein kinase identifier and classifier) databases. Domain annotation of the deduced protein sequence of the *Un8* candidate gene was carried out by SMART (http://www.smart.embl-heidelberg.de/), PROSITE (http://www.prosite.expasy.org/), and the conserved domain database (CDD) (http://www.ncbi.nlm.nih.gov/cdd).

Un8 candidate gene sequencing

The *Un8* candidate gene was sequenced from genomic DNA isolated from the 21 barley lines listed in Table 1. DNA was extracted using the modified CTAB method (Procunier et al. 1991). The candidate gene was PCR amplified using the standard PCR conditions listed previously with two sets of primers which produced overlapping amplicons (Table 2). Amplicons were cloned, sequenced and aligned as described previously to ensure the consistency of the sequence data obtained. Overlapping amplicons for each barley line were then aligned and a consensus sequence for the candidate gene created and used for translation to the putative protein using DNAMAN v. 7.

Results

Evaluation of true loose smut reactions

After development of a F_4 population derived from the cross TR09398 (resistant, carries *Un8*) × TR07728 (susceptible, lacks *Un8*), a total of 4836 lines from this population were screened with the *Un8* SNP1 and *Un8* SNP6 TaqMan markers with 122 recombinant lines identified. After inoculation with the mixture of true loose smut pathotypes, 57 lines showed resistance and 65 lines showed

susceptibility which was consistent with a single gene mode of resistance ($\chi^2 = 0.525$, $P = 0.4689$).

Marker enrichment and fine genetic mapping of the *Un8* interval

Un8 was initially confined to a 6.2 cM region on chromosome arm 1HL by the flanking markers *Un8* SNP1 (unigene 4245) and *Un8* SNP6 (unigene 14722) (P. Eckstein, personal communication). To enrich the *Un8* region, four different methods were used. First, sequences from 12 barley unigenes located between unigenes 4245 and 14722 were identified and used for marker development. After these unigene sequences were queried against the barley cv. 'Morex' whole genome assembly using the ViroBLAST interface tool to obtain larger genomic DNA sequence reads, polymorphisms between the mapping population parents, TR09398 and TR07728, were identified in five of these unigenes which became the basis of markers 8487, 1406, 0498L15 F3/R3, 3602 and 13742 (Table 3). Second, genotyping data obtained from three true loose smut resistance sources ('CDC Meredith', TR306 and TR09398) and four susceptible sources (TR07728, TR09397, 'Harrington' and 'CDC Kindersley') using the 9K Barley iSELECT Infinium SNP Assay identified 21 SNP markers in the *Un8* interval. This resulted in the placement of two additional markers (48060 and 10924) in the *Un8* region (Table 3). Third, colinearity between barley chromosome arm 1HL (location of *Un8*) and the syntenic regions on the long arm of rice chromosome 5 and Brachypodium chromosome 2 were exploited to find additional markers. Barley unigenes 4245 (*Un8* SNP1) and 14722 (*Un8* SNP6), which bracket the *Un8* gene, were used to define the orthologous regions in rice and Brachypodium. All rice and Brachypodium genes contained in the orthologous interval were identified and queried against the HarvEST:Barley (assembly 35) database and the barley cv. 'Morex' whole genome assembly (using the ViroBLAST interface tool) to identify previously unidentified or unmapped barley unigenes in the *Un8* region. This produced two additional markers (17452 and 21217) in the *Un8* region (Table 3). Finally, based on available BAC sequences in the *Un8* region, four markers were developed, 0751D06 F6/R6 from BAC HVVMRXALLhA0751D06, 0498L15 F8/R8 from BAC HVVMRXALLhA0498L15, and *Un8* SNP7 and HI1406 from BAC HVVMRXALLhA0772N02 (Table 3).

Thirteen new markers were developed for the *Un8* interval using the 4625 RILs derived from TR09398 × TR07728 which, along with the *Un8* SNP1, *Un8* SNP4 and *Un8* SNP6 markers, defined a 2.853-cM region (Fig. 1). Among the 122 lines used to create the linkage map spanning the *Un8* locus, a total of 132 recombinations were observed within the *Un8* SNP1 to *Un8* SNP6 interval because several of the lines contained multiple recombinations. When the additional 211 lines which contained a single recombinant chromosome in the *Un8* region were included in the calculation of genetic distance between *Un8* SNP1 and *Un8* SNP6, a value of 4.911 cM was obtained. After including the true loose smut reaction data, two markers, 0751D06 F6/R6 and *Un8* SNP4, showed complete linkage with *Un8* in the TR09398 × TR07728 population (Fig. 1).

It was previously known that one line from the DH population derived from the cross 'Harrington' × TR306 (Eckstein et al. 2002) showed a recombination between the *Un8* gene and the *Un8* SNP4 marker. To determine if the 0751D06 F6/R6 marker co-segregating with the *Un8* gene in the TR09398 × TR07728 population was closer to *Un8* than the *Un8* SNP4 marker, this marker was screened on the 'Harrington' × TR306 population. The 0751D06 F6/R6 marker showed no recombination with *Un8*, indicating that it was the most closely linked marker to the *Un8* gene (Fig. 2a). Ultimately, the *Un8* gene was determined to be within a genetic interval flanked by markers *Un8* SNP4 and 0498L15 F8/R8 (Fig. 2a).

Synteny between barley, rice and Brachypodium

The syntenic regions between barley chromosome arm 1HL, rice chromosome 5 and Brachypodium chromosome 2 around the *Un8* gene were delimited by markers *Un8* SNP1 and *Un8* SNP6 and very few rearrangements of marker order were observed (Fig. 1). No orthologous sequences for the most closely linked distal marker, 0498L15 F8/R8, could be identified in either rice or Brachypodium (Figs. 1, 2a). Therefore, we used the 0498L15 F3/R3 and *Un8* SNP4 markers, which were 0.346 cM apart (Fig. 1), to calculate the physical distance of the syntenic regions in rice and Brachypodium. In rice, the orthologous region spanned approximately 24,000 bp and contained three genes, while the same region was less than 3000 bp in Brachypodium and no genes were present (Fig. 1).

Physical mapping and candidate gene prediction for *Un8*

Fingerprinted contig 546, part of the genome-wide physical map of barley (The International Barley Genome Sequencing Consortium 2012; Ariyadasa et al. 2014), was anchored to the *Un8*-targeted genetic map generated in this study by markers *Un8* SNP4 and 0498L15 F3/R3 and the MTP for FPC 546 was defined. The MTP was composed of two overlapping BACs, HVVMRXALLmA0180J17 (~160 kb) which contained marker *Un8* SNP4 and HVVMRXALLeA0154F16 (~150 kb) which contained 0498L15 F3/R3 (Fig. 2b). Subsequently, the 0498L15 F8/R8 marker was identified within BAC HVVMRXALLeA0154F16 and the 0751D06 F6/R6 marker was

Table 3 Detailed information on markers identified in the *Un8* interval between *Un8* SNP1 and *Un8* SNP6 in the TR09398 × TR07728 population

Name	Origin	Primer sequences (5′–3′)		Type	Unigene ID[a]	BAC ID[b]	Morex contig
Un8 SNP1[c]	HarvEST:Barley	F	CTTGTCAGTTGAATGCCAATCTCTT	TaqMan	4245	HVVMRX83KhA0023N07_c1	Contig_158214
		R	GTTCGACAATGATTGCATCTCACA				
		Reporter 1	TGCAGCTTGGTCTCAAT				
		Reporter 2	AGCCTGGCCTCAAT				
8487	HarvEST:Barley	F1	AGTACCCTCACCCTTACAAATTC	Allele-specific PCR	8487	HVVMRXALLhA0368G13_v37_c5	Contig_1569595
		F2	GTACCCTCACCCTTACAAATTG				
		R	CACCTCTACATTTGGGTCCTTG				
48060	iSELECT Assay	F	GTCCATTTCTCAGGCTCAGTG	Allele-specific PCR	48060	HVVMRXALLhB0089C19	Contig_38392
		R1	CCGACGGTCAAGGTCTCA				
		R2	CGACGGTCAAGGTCTCG				
Un8 SNP7	Morex BAC	F	GGCGAAATTCTCCTTGAAAACATGT	TaqMan	1406	HVVMRXALLhA0772N02_IPK_NODE_0001	Contig_5603
		R	CATGTTTGCCCCAATAAAAATGTCTAAC				
		Reporter 1	CACTTGTCTTGCCACTTC				
		Reporter 2	TTCACTTGTCTTACCACTTC				
HI 1406	Morex BAC	F	AAGTGCGACGACGGAATACA	CAPS (cut by *Msc* I)	–	HVVMRXALLhA0772N02_IPK_NODE_0001	–
		R	ATAGGACGCAAACCGACAAG				
1406	HarvEST:Barley	F1	TGTCCTTAGTTCACTTGTCCTG	Allele-specific PCR	1406	HVVMRXALLhA0772N02_IPK_NODE_0001	Contig_5603
		F2	CATGTCCTTAGTTCACTTGTCCTA				
		R	GCTACTACTGACTATCGCCACAT				
Un8 SNP4[c]	HarvEST:Barley	F	CTACATCTGCTCCTGCGACTT	TaqMan	16527	HVVMRXALLeA0355N04	Contig_171284
		R	TCAAAATCGAGCTTCCCATCACAAT				
		Reporter 1	CACAGTATATCGCACGAGAA				
		Reporter 2	TCACAGTATATCACACGAGAA				
0751D06 F6/R6	Morex BAC	F	TCAGAGATGGCTGTGAGGATG	Allele-specific PCR	–	HVVMRXALLmA0180J17_sc1	Contig_93215
		R	CTGCGTTTACGAAGATGGATGT				
0498L15 F8/R8	Morex BAC	F	TCGGTGTTCAGTCCCAAGTC	Allele-specific PCR	–	HVVMRXALLhA0498L15_NODE_0053.1	–
		R	TGACCCTGCTGGTAGGTAGAGT				
0498L15 F3/R3	HarvEST:Barley	F	TGGCTGCAATATCATGGTCAT	Amplicon size	15283	HVVMRXALLeA0154F16_sc2	Contig_46703
		R	AATCATTTGCCAGGTCAGAAG				
3602	HarvEST:Barley	F	TTGCTGTTTGGTCTGGTCTTG	CAPS (cut by *Taq*I)	3602	HVVMRXALLeA0217G19	Contig_43658
		R	GCACCTTCAGCAATCTCAATCT				
10924	iSELECT assay	F	GCTTCTGTTCACGCCACTGT	Allele-specific PCR	10924	HVVMRXALLeA0217G19	Contig_43658
		R1	CGCTTACTCCATGGTGCTAAT				
		R2	CGCTTACTCCATGGTGCTAAG				

Table 3 continued

Name	Origin	Primer sequences (5′–3′)		Type	Unigene ID[a]	BAC ID[b]	Morex contig
17452	Synteny	F	CCTGGTGGTGATGCAGAAGAT	Amplicon size	17452	HVVMRX83KhA0046D01_c1	Contig_162350
		R	CACCTGATAGGCAGAGGAGTACAC				
13742	HarvEST:Barley	F	AAGACCATCACGCTGGAG	CAPS (cut by *Sac*II)	13742	HVVMRXALLhA0568K13_c1	Contig_275666
		R	CAATGACAAACGACAGGG				
21217	Synteny	F	GGTCGGGAGTATGACTTAGGAAT	Allele-specific PCR	21217	HVVMRXALLhA0568K13_v11_c1	Contig_159925
		R	CAGGAAGCAAGGAATACTGGAAT				
Un8 SNP6[c]	HarvEST:Barley	F	GGCAACCCACGGAAACAC	TaqMan	14722	HVVMRXALLhB0144C24	Contig_39431
		R	CCACGCTGATCTTTATCTATGGCTAA				
		Reporter 1	TGACGACCAAACGATACAT				
		Reporter 2	TGACGACCAAAATATAG				

[a] Unigene ID determined from HarvEST:Barley v. 1.83, assembly 35

[b] The 'sc' or 'c' suffix indicates the scaffold or contig, respectively, on which the predicted gene resides within the associated BAC

[c] Previously developed by P. Eckstein

located within both BAC clones (Fig. 2b). Complete sequence for each BAC was contained in multiple scaffolds and contigs of varying size (Fig. 2c). BAC HVVMRXALLmA0180J17 was composed of one very large scaffold (J17_sc1; ~151 kb) and three small contigs (Fig. 2c), while BAC HVVMRXALLeA0154F16 was covered by two large scaffolds (F16_sc1 and F16_sc2; ~52 and ~45 kb) and seven smaller scaffolds and contigs (Fig. 2c). DNA sequences of both BACs were analyzed for putative genes and a total of 17 were identified (Fig. 2c; Table 4). Only scaffolds and contigs containing putative genes are shown in Fig. 2c. Among this group, only two resistance-associated genes were identified using GeneMark and GENSCAN. One was a cell wall invertase (β-fructofuranosidase, Fig. 2c; Table 4), but a CAPS marker designed for this gene identified one recombination between it and *Un8* within the 'Harrington' × TR306 population. The second predicted resistance-associated gene was a protein kinase containing two tandem kinase catalytic domains. It was co-located within the same two BAC clone scaffolds as the *Un8* co-segregating marker 0751D06 F6/R6 at a distance of ~3000 bp (Fig. 2c; Table 4). The gene was also identified in HarvEST:Barley (v. 1.98, assembly 37), denoted as MLOC_38442 (The International Barley Genome Sequencing Consortium 2012), and in the barley whole genome assembly as a RNA-Seq gene (denoted XLOC_040148; The International Barley Genome Sequencing Consortium 2012). The predicted function of this gene and its identification from multiple sources make it a good candidate for the *Un8* gene. According to the iTAK (plant transcription factor and protein kinase identifier and classifier) database, the best BLASTP hit for the *Un8* candidate gene was to a wall-associated protein kinase in *Oryza sativa* (BLAST E-score $2e^{-112}$, 41 % identity (275/672) at the amino acid level with 55 gaps). There were two additional predicted proteins present within the same BAC scaffold as the *Un8* co-segregating marker 0751D06 F6/R6 (Fig. 2c; Table 4) which cannot be fully disregarded as possible candidate genes. However, the lack of an annotated function for both and the absence of a corresponding RNA sequence (The International Barley Genome Sequencing Consortium 2012) for one of the predicted genes make them weaker candidates. No putative orthologous barley genes corresponding to the three rice genes which were located within the syntenic region identified by the *Un8* SNP1 and *Un8* SNP6 markers were present within the two BACs spanning the *Un8* locus.

Un8 candidate gene alleles and deduced protein domain annotation

Sequencing of the candidate gene from the resistant and susceptible parents indicated a length of 2037 and 2031 nucleotides, respectively, and no intron in either. The translated amino acid sequences contained 26 substitutions or small indels which differentiated them (Group I

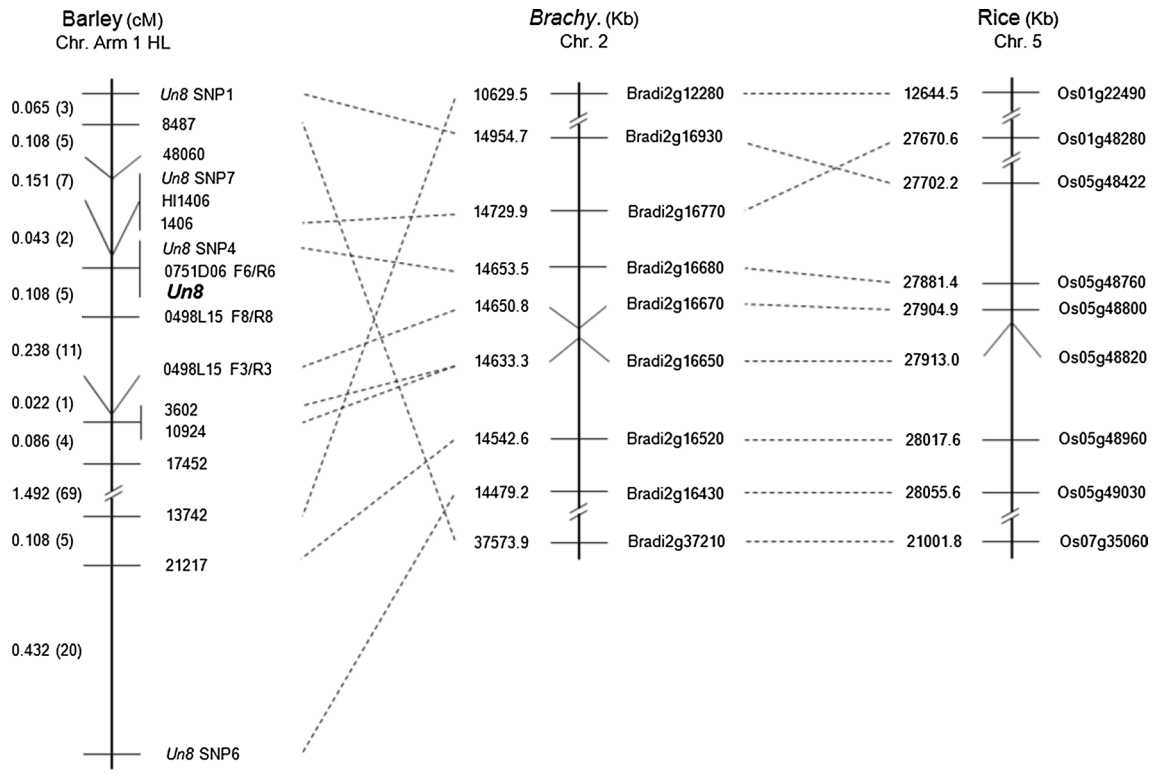

Fig. 1 Genetic map of the *Un8* region on barley chromosome arm 1HL created using 4625 recombinant inbred lines derived from TR09398 × TR07728, and comparison of this interval with the physical maps of *Brachypodium distachyon* chromosome 2 and rice chromosome 5. *Dashed lines* connect putative orthologous genes. Marker names (barley) and gene names (*Brachypodium* and rice) are indicated to the *right* of each map while distance (cM and recombination events (in *brackets*) in barley, kb from the top of chromosome 2 in *Brachypodium* and the top of chromosome 5 in rice) are denoted to the *left* of each linkage group. The *Un8* gene is denoted in *bold text*. *Double slashes* within each linkage group represent a large interval, or alternate chromosome in the case of rice, not in scale with the remainder of the linkage group

versus Group II sequences in Fig. 3). To better understand the amino acid variations present in the deduced protein sequence of the *Un8* candidate gene, 19 lines (4 resistant and 15 susceptible, Table 1) were sequenced and translated into the corresponding amino acid residues additionally (Fig. 3). The deduced amino acid sequence from the resistant parent (TR09398) and four additional resistant lines (TR11698, TR12135, HB11316, and 'AC Metcalfe') showed no sequence difference among them (Fig. 3). In contrast, the deduced amino acid sequence of the susceptible lines resulted in three groupings (Fig. 3): TR07728, 'Bowman', 'Morex', 'OAC 21', TR12737, 'Calcule', and 'Streif' (Group II); 'Barke', 'Baudin','Quilmes Carisma', 'Champlain', 'Jet', and 'Montcalm' (Group III);'Quilmes Ayelén', 'CDC Austenson', and 'Optic' (Group IV). The deduced amino acid sequence obtained for Groups II–IV displayed a variety of substitutions or small indels in comparison to that obtained from the resistant lines. In comparison to the 678 amino acids translated from the candidate gene of resistant lines, the Group II and III alleles from susceptible lines produced a deduced protein of 676 amino acids while the Group IV allele from susceptible lines was translated into a slightly longer deduced protein of 681 amino acids (Fig. 3).

Among the 21 resistant and susceptible barley lines which were sequenced, 15 amino acids substitutions unique to the deduced protein sequence present in resistant lines differentiated it from all deduced protein sequences identified in susceptible lines. Most of the substitutions (11 of 15) existed within the kinase II domain (Fig. 3). Three of the sequence variations were notable for their potential impacts on function (and thus resistance), as they occurred within predicted functional regions of the deduced protein. The amino acid substitution at position 532 (serine in resistant lines and threonine in susceptible lines) was present within the activation loop of the kinase II domain. Similarly, the substitution at position 530 (glutamine in resistant lines and leucine in susceptible lines) was also present in the activation loop of the kinase II domain, but it had additional predicted functions as an active site and substrate binding site. The third substitution at position 513 (glutamine in resistant lines and histidine in susceptible lines) within the

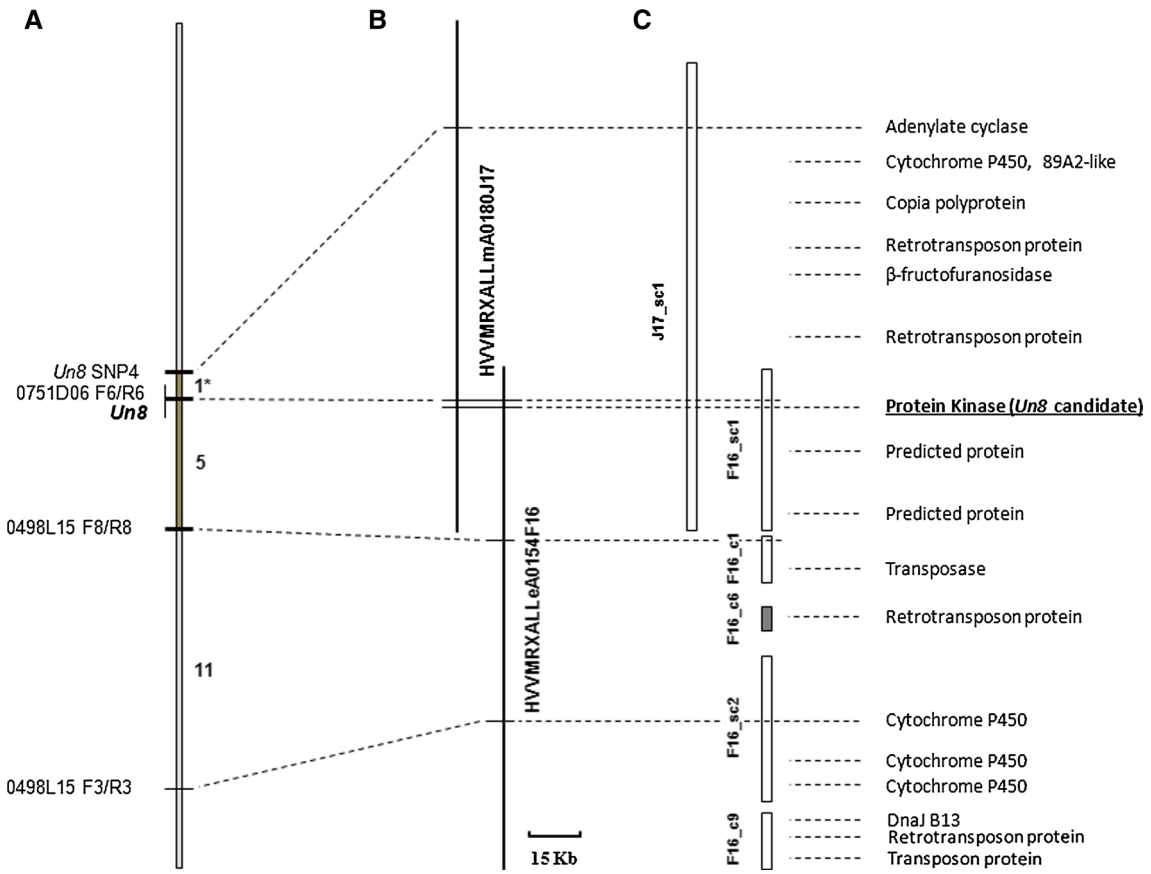

Fig. 2 Fine-scale orientation of the *Un8* region on barley chromosome arm 1HL indicating flanking markers, BAC clones spanning the *Un8* locus and location of all predicted genes within the BAC clones. **a** Genetic map displaying flanking markers which encompass the *Un8* locus (*shaded region*) with marker names to the *left* and recombination events observed in both the TR09398 × TR07728 and 'Harrington' × TR306 mapping populations indicated to the *right*. The *asterisked number* indicates a recombination between *Un8* and the *Un8* SNP4 marker observed in the 'Harrington' × TR306 population (Eckstein et al. 2002). Marker 0751D06 F6/R6 co-segregated with the *Un8* gene in both mapping populations. **b** Physical map of the *Un8* region. The *Un8* locus is spanned by two BACs (HVVMRXALLmA0180J17 and HVVMRXALLeA0154F16) with an overlapping region of ~50 kb in which the *Un8* locus resides. A *scale bar* for the physical map is provided at the bottom. **c** Predicted genes in the *Un8* region. Gene annotations are presented on the *right* and BAC scaffolds ('sc') and contigs ('c') are indicated to the *left*. J17 and F16 denote the corresponding BAC clone with which the scaffold or contig is associated. Only BAC scaffolds and contigs containing predicted genes are shown. The exact location of the F16_c6 contig (denoted with a *shaded box*) relative to F16_sc2 and F16_c1 has not been determined. The correct orientation of F16_c1, F16_c6, F16_sc2 and has not been resolved

kinase II domain was predicted to be important as an active site, ATP binding site and substrate binding site.

Discussion

Over the past 70 years, a minimum of 15 resistance loci conferring true loose smut resistance were reported (Mueller 2006) and *Un8*, which was identified more than half of a century ago, is still the most effective (Metcalfe and Johnston 1963; Metcalfe 1966; Thomas and Menzies 1997). Barley lines harbouring *Un8* are resistant to all known true loose smut isolates in Western Canada, making it the most valuable resource for true loose smut resistance breeding. However, the search for new sources of resistance continues, for example the identification of resistance on chromosome 3H (Menzies et al. 2010), which could provide alternative resistance should *Un8* resistance become ineffective with the evolution of new virulent pathotypes.

Fine mapping the barley *Un8* locus

Developing markers to *Un8* resistance has been instrumental to allow for MMAS of *Un8*-based resistance, but it also provided the initial tools towards map-based cloning of the underlying gene. Building on the 20 years of effort to genetically map the *Un8* true loose smut resistance gene (Eckstein et al. 1993, 2002; Li et al. 2001; P. Eckstein,

Table 4 Annotations, BLASTP ID and E-scores, gene ID, BAC ID and 'Morex' contig associated with all predicted genes identified in BACs HVVMRXALLmA0180J17 and HVVMRXALLeA0154F16 which comprise the minimum tilling path spanning the *Un8* region

Predicted gene annotation[a]	BLASTP ID[b]	E-score[b]	Barley gene ID[c]	BAC ID[d]	'Morex' Contig
Adenylate cyclase	XP_006655586	0	MLOC_25774	HVVMRXALLmA0180J17_sc1	contig_171284
Cytochrome P450, 89A2-like	XP_005335357	$7e^{-111}$	–	HVVMRXALLmA0180J17_sc1	contig_58606
Copia polyprotein	AF466199	0	–	HVVMRXALLmA0180J17_sc1	–
Retrotransposon protein	ABB46931	0	–	HVVMRXALLmA0180J17_sc1	contig_84931
Beta-fructofuranosidase	XP_003575078	0	MLOC_5612	HVVMRXALLmA0180J17_sc1	contig_136454
Retrotransposon protein	BAA22288	0	–	HVVMRXALLmA0180J17_sc1	–
Protein kinase (*Un8* candidate)	EEE50557 (AAL25177)	$3e^{-179}$ ($2e^{-112}$)	MLOC_38442	HVVMRXALLmA0180J17_sc1 HVVMRXALLeA0154F16_sc1	contig_2550456
Predicted protein	AF427791	0	–	HVVMRXALLmA0180J17_sc1 HVVMRXALLeA0154F16_sc1	contig_1590184, contig_1973319, contig_243921
Predicted protein	BAK06775	0	MLOC_65367	HVVMRXALLmA0180J17_sc1 HVVMRXALLeA0154F16_sc1	contig_50087
Transposase	EMT30676	0	–	HVVMRXALLeA0154F16_c1	–
Tetrotransposon protein	BAA22288	0	–	HVVMRXALLeA0154F16_c6	–
Cytochrome P450	BAJ96841	0	MLOC_17557	HVVMRXALLeA0154F16_sc2	contig_1577063
Cytochrome P450	BAJ96841	0	MLOC_17557	HVVMRXALLeA0154F16_sc2	contig_1577063
Cytochrome P450	BAJ96841	0	MLOC_17557	HVVMRXALLeA0154F16_sc2	contig_1577063
DnaJ homolog subfamily B member 13	EMT30186	$5e^{-130}$	MLOC_23473	HVVMRXALLeA0154F16_c9	contig_140197, contig_162350
Retrotransposon protein	BAH79979	0	–	HVVMRXALLeA0154F16_c9	contig_223045
Transposon protein	AAP53844	$4e^{-74}$	–	HVVMRXALLeA0154F16_c9	–

[a] Predicted genes are presented in the same order as in Fig. 2c

[b] BLASTP ID and E-scores were determined using the NCBI database for all predicted genes. The BLASTP ID and E-score indicated in brackets for the *Un8* candidate gene were determined with the iTAK database

[c] MLOC gene identifiers obtained from the IPK barley blast server (http://www.webblast.ipk-gatersleben.de/barley/viroblast.php)

[d] The 'sc' or 'c' suffix indicates the scaffold or contig, respectively, on which the predicted gene resides within the associated BAC

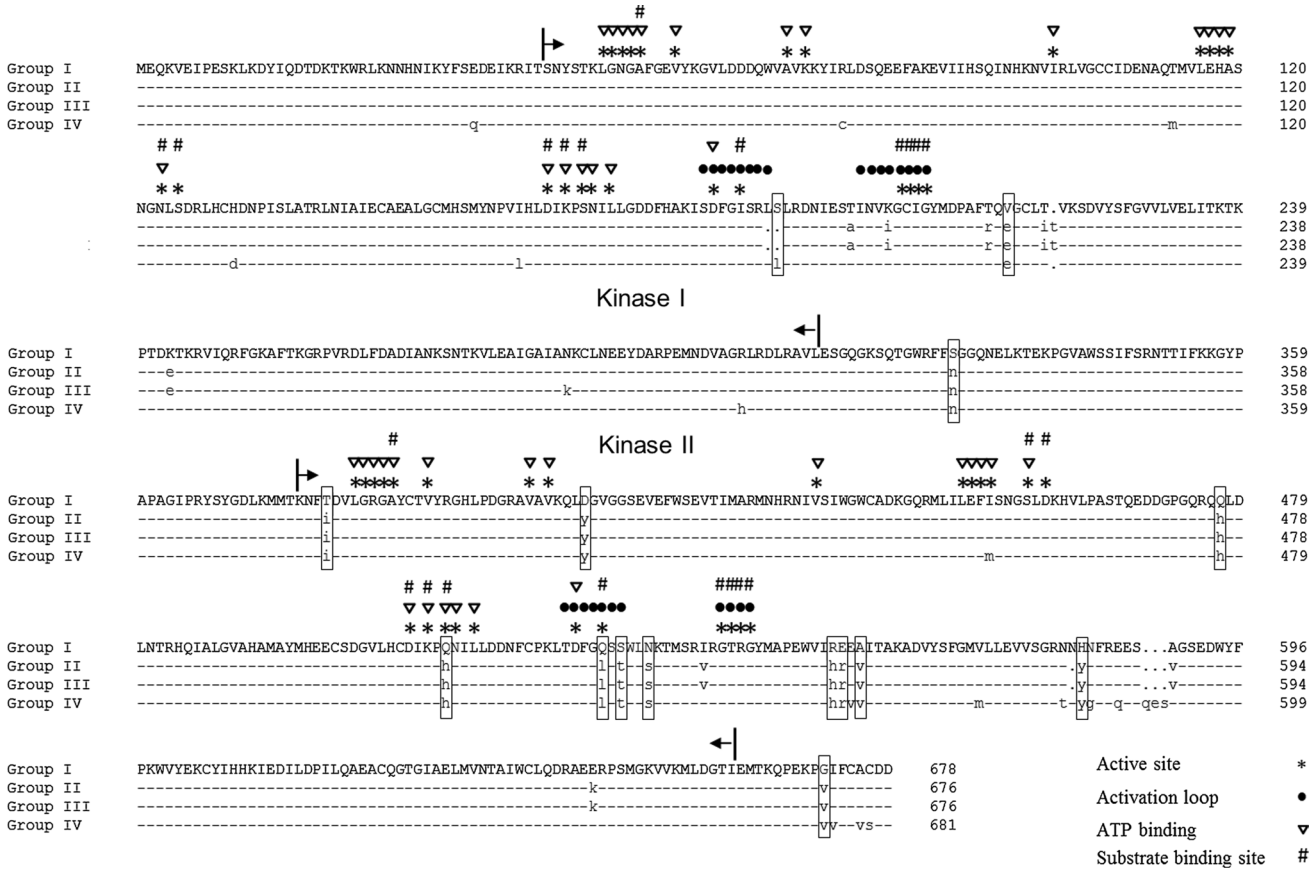

Fig. 3 Amino acid sequence alignment of the *Un8* candidate gene alleles derived from 21 resistant and susceptible barley lines (Table 1). The dominant features of the deduced protein are the two tandem kinase domains (each of which is delimited by *forward* and *reverse-pointing arrows*). Fifteen amino acid polymorphisms (indicated by *boxes*) differentiate the deduced protein sequence in resistant lines from the deduced protein sequences in susceptible lines. Location of predicted active sites, activation loops, ATP binding sites and substrate binding sites are indicated above the relevant amino acids. Group I: sequence from the five resistance lines (TR09398, TR11698, TR12135, HB11316, 'AC Metcalfe'); Group II: sequence from seven susceptible lines (TR07728, 'Bowman', 'Morex', 'OAC 21', TR12737, 'Calcule', 'Streif'); Group III: sequence from six susceptible lines ('Barke', 'Baudin','Quilmes Carisma', 'Champlain', 'Jet', 'Montcalm'); Group IV: sequence from three susceptible lines ('Quilmes Ayelén', 'CDC Austenson', 'Optic')

personal communication), this study has created a high resolution map of the *Un8* region consisting of sixteen markers and spanning a distance of 2.853 cM near the distal end of chromosome arm 1HL.

Positional cloning of genes in barley is hampered by the large genome (5.1 Gb) and high percentage of repetitive sequences (84 %) (The International Barley Genome Sequencing Consortium 2012), however, the *Un8* gene is located near the distal end of chromosome arm 1HL where three recombination 'hot spots' were identified, with an estimated physical/genetic ratio of 0.6 Mb/cM (Künzel et al. 2000) which assisted fine mapping of the *Un8* gene in this study. The location of the *Un8* gene supports prior observations that many barley resistance genes are found distally in regions of high recombination (The International Barley Genome Sequencing Consortium 2012). In the present study, the interval spanning markers *Un8* SNP4 and 0498L15 F3/R3 is less than 265 kb and 16 recombinants were identified from the 4625 F_4 lines screened (or 0.346 cM). This produced a physical to genetic distance ratio of 0.76 Mb/cM, similar to the estimate of Künzel et al. (2000).

Comparative studies in the *Un8* region with rice and Brachypodium

Rice diverged from barley approximately 50 million years ago (Dubcovsky et al. 2001; Paterson et al. 2004) which predated the divergence of barley and Brachypodium (Bossolini et al. 2007; The International Brachypodium Initiative 2010). As such, Brachypodium shows a closer relationship with Triticeae than rice or other species like sorghum (The International Brachypodium Initiative 2010) and may be a better model for comparative study (Huo

et al. 2008; Mayer et al. 2011). For example, the analogue of the barley *Rpg1* stem rust-resistance gene can be found within the syntenic region in Brachypodium, but not in rice (Brueggeman et al. 2002; Drader and Kleinhofs 2010). Similarly, resistance gene analogues to the *Yr26* wheat stripe rust-resistance gene were located by syntenic mapping in Brachypodium, but no such genes were identified in the syntenic region of rice (Zhang et al. 2013). The syntenic relationship was conserved slightly better in Brachypodium than in rice for the *Un8* region. Only barley markers 13742 and 8487 localized to different locations on Brachypodium chromosome 2. These same two markers also localized to alternate chromosomes in rice, as did barley marker 1406 (Fig. 1). However, no resistance genes (or analogues) were identified in the *Un8* syntenic region of either Brachypodium or rice.

Un8 candidate gene

Only two of the 17 genes predicted to exist within the BAC clones spanning the *Un8* locus appeared to play a role related to disease resistance and both were located within the interval delimited by *Un8* SNP4 and 0498L15 F8/R8. One of these two genes was predicted to be a cell wall invertase (β-fructofuranosidase). These genes are up-regulated in response to pathogen infection and, via the import of hexose sugars to the site of infection, help increase plant metabolism to mount an effective defence (Proels and Hückelhoven 2014). However, a recombination event was identified in the 'Harrington' × TR306 mapping population between *Un8* and the predicted cell wall invertase gene located in BAC HVVMRXALLmA0180J17. The second disease-related gene was a predicted protein kinase resistance gene analogue that was located close to the 0751D06 F6/R6 marker co-segregating with *Un8*. Protein kinases, such as receptor-like protein kinases and mitogen-activated protein kinases, are representatives of one of the main protein classes associated with plant disease resistance. In barley, most of the kinase-containing resistance genes located on both arms of chromosome 1H tended to cluster distally (The International Barley Genome Sequencing Consortium 2012).

Domain annotation of the deduced protein sequence of the *Un8* candidate gene showed that it contained two tandem protein kinase domains. Both of the catalytic domains were classified into the tyrosine kinase subfamily using the SMART database. However, the CDD database placed the first catalytic domain into the tyrosine-specific kinase subfamily (smart00219) while the second domain contained a pfam00069 protein kinase domain similar to that found in *Rpg1* (Brueggeman et al. 2002). If the *Un8* candidate gene is ultimately proven to be the *Un8* resistance gene, then it and *Rpg1* would be the only barley resistance proteins reported to contain two protein kinase domains, although they do not share a high degree of similarity with only 26 % (186 of 703 amino acid residues) overall identity at the protein level (26 % within the protein kinase I domain and 30 % within the protein kinase II domain). Moreover, *Rpg1* is classified as a receptor-like protein kinase and the *Un8* candidate gene as a wall-associated protein kinase. However, because both the genes mediate durable resistance to barley biotrophic pathogens (lasting over 50 years in both cases) it would be interesting to determine if the longevity of their resistances is based on a similar mechanism.

Among the 21 barley lines which were studied, 15 amino acids substitutions differentiated the deduced protein sequence present in resistant lines from all three deduced protein sequences present in susceptible lines. Three of these amino acid substitutions, at positions 513, 530 and 532, were of particular interest because they occurred within predicted functional regions of the kinase II domain and therefore could affect function of the deduced protein (and thus resistance). The amino acid at position 513 was predicted to form part of the active site, ATP binding site and substrate binding site, while the amino acids at positions 530 and 532 composed part of the activation loop. Small changes in amino acid sequence have been shown to affect the function of resistance genes. For example, the difference in function of the resistant and susceptible alleles of the rice blast resistance gene *Pi-ta* could be explained by only one amino acid difference (Bryan et al. 2000). In contrast, the susceptible alleles of *Rpg1* appeared to be non-functional because of the complete absence of the gene or due to the presence of various stop codons (Brueggeman et al. 2002). However, subsequent work indicated that substitution of catalytically active amino acids, specifically lysine 461 and 462 in the second kinase domain, could render the gene non-functional (Nirmala et al. 2006).

The predicted wall-associated protein kinase gene identified in this study is a strong candidate to be the *Un8* gene due to its prediction from multiple sources, the presence of a corresponding RNA sequence aligned with the candidate gene position within the cv. 'Morex' whole genome assembly and the presence of amino acid substitutions, three of which were in predicted functional regions of the protein, that were consistently associated with resistance to true loose smut in a panel of barley lines from diverse origins. However, because the BAC clones used in this study are derived from 'Morex', a susceptible variety, it is also possible that 'Morex' does not contain an *Un8* allele and the gene thus would not be present within the clones. Additionally, there were two predicted genes within the interval delimited by *Un8* SNP4 and 0498L15 F8/R8 which could also be the *Un8* gene. However, these two genes were considered weak candidates since no annotated function was associated with either and there was no RNA sequence identified for one of them.

In conclusion, a physical map surrounding the *Un8* true loose smut resistance gene was constructed and one putative disease resistance gene analogue sequence was found in the region which was considered as a strong candidate for the *Un8* gene. Until such time as a perfect marker is created for *Un8*, the 0751D06 F6/R6 marker, which not only cosegregates perfectly with *Un8* in the TR09398 × TR07728 mapping population, but also in a broader spectrum of barley populations (P. Eckstein, personal communication), will be very useful for MMAS efforts as it will alleviate some of the prior issues related to recombination between *Un8* and previous markers which caused incorrect phenotypic predictions in barley breeding lines at the CDC. Next steps will focus on expression analysis of the candidate gene and transformation of the candidate gene into a susceptible barley line to definitively prove that it is the *Un8* gene.

Author contribution statement W. Z., P. E. E., A. D. B., and G. J. S. designed the research; W. Z., P. E. E., M. C. and D. V. performed research; W. Z., P. E. E., A. D. B., and G. J. S. analyzed the experimental data; W. Z. wrote the manuscript; N. S., A. D. B., and G. J. S. involved in critical revision of the manuscript; A. H., S. B., and N. S. sequenced the BACs and established the minimal tiling path of the FPC for the *Un8* region.

Acknowledgments The authors would like to thank the Saskatchewan Agriculture Development Fund for their financial support that allowed this research to be conducted, and the China Scholarship Council for their financial support of W. Zang's study. Physical mapping and sequence analysis of the barley genome was supported by grants of the German Ministry of Education and Research (BARLEX: FKZ 0314000; TRITEX: FKZ 0315954) to N. S. and U. S. We would also like to thank Dr. Timothy J. Close for providing the HarvEST Barley (v.1.98) and Dr. Stefan Taudien, Martin Mascher, Dr. Matthias Platzer, and Dr. Uwe Scholz for providing access to unpublished barley genome sequence data.

Conflict of interest The authors have no conflict of interest to declare.

Ethical standard The experiments comply with the current laws of the countries in which they were performed.

References

Acevedo-Garcia J, Collins NC, Ahmadinejad N, Ma L, Houben A, Bednarek P, Benjdia M, Freialdenhoven A, Altmüller J, Nürnberg P, Reinhardt R, Schulze-Lefert P, Panstruga R (2013) Fine mapping and chromosome walking towards the *Ror1* locus in barley (*Hordeum vulgare* L.). Theor Appl Genet 126:2969–2982

Andrews JE (1956) Inheritance of reaction to loose smut, *Ustilago nuda*, and to stem rust, *Puccinia graminis tritici*, in barley. Can J Agr Sci 36:356–370

Ariyadasa R, Mascher M, Nussbaumer T, Schulte D, Frenkel Z, Poursarebani N, Zhou R, Steuernagel B, Gundlach H, Taudien S, Felder M, Platzer M, Himmelbach A, Schmutzer T, Hedley PE, Muehlbauer GJ, Scholz U, Korol A, Mayer KF, Waugh R, Langridge P, Graner A, Stein N (2014) A sequence-ready physical map of barley anchored genetically by two million single-nucleotide polymorphisms. Plant Physiol 164:412–423

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

Brueggeman R, Rostoks N, Kudrna D, Kilian A, Han F, Chen J, Druka A, Steffenson B, Kleinhofs A (2002) The barley stem rust-resistance gene *Rpg1* is a novel disease-resistance gene with homology to receptor kinases. Proc Natl Acad Sci USA 99:9328–9333

Bryan GT, Wu KS, Farrall L, Jia Y, Hershey HP, McAdams SA, Faulk KN, Donaldson GK, Tarchini R, Valent B (2000) A single amino acid difference distinguishes resistant and susceptible alleles of the rice blast resistance gene *Pi-ta*. Plant Cell 12:2033–2046

Close TJ, Bhat PR, Lonardi S, Wu Y, Rostoks N, Ramsay L, Druka A, Stein N, Svensson JT, Wanamaker S, Bozdag S, Roose ML, Moscou MJ, Chao S, Varshney RK, Szucs P, Sato K, Hayes PM, Matthews DE, Kleinhofs A, Muehlbauer GJ, DeYoung J, Marshall DF, Madishetty K, Fenton RD, Condamine P, Graner A, Waugh R (2009) Development and implementation of high-throughput SNP genotyping in barley. BMC Genom 10:582

Drader T, Kleinhofs A (2010) A synteny map and disease resistance gene comparison between barley and the model monocot *Brachypodium distachyon*. Genome 53:406–417

Dubcovsky J, Ramakrishna W, SanMiguel PJ, Busso CS, Yan L, Shiloff BA, Bennetzen JL (2001) Comparative sequence analysis of colinear barley and rice bacterial artificial chromosomes. Plant Physiol 25:1342–1353

Eckstein PE, Duncan SJL, Scoles GJ, Rossnagel BG (1993) An RFLP for true loose-smut resistance and for scald resistance. In: Proceedings of the 15th North American Barley Research Workshop, Guelph, Ontario, Canada

Eckstein PE, Krasichynska N, Voth D, Duncan S, Rossnagel BG, Scoles GJ (2002) Development of PCR-based markers for a gene (*Un8*) conferring true loose smut resistance in barley. Can J Plant Pathol 24:46–53

Huo N, Lazo GR, Vogel JP, You FM, Ma Y, Hayden DM, Coleman-Derr D, Hill TA, Dvorak J, Anderson OD, Luo MC, Gu YQ (2008) The nuclear genome of *Brachypodium distachyon*: analysis of BAC end sequences. Funct Integr Genomics 8:135–147

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Künzel G, Korzun L, Meister A (2000) Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. Genetics 154:397–412

Li CD, Eckstein PE, Lu M, Rossnagel BG, Scoles GJ (2001) Targeted development of a microsatellite marker associated with a true loose smut resistance gene in barley (*Hordeum vulgare* L.). Mol Breed 8:235–242

Liu J, Huang SM, Sun MY, Liu SY, Liu YM, Wang WX, Zhang XR, Wang HZ, Hua W (2012) An improved allele-specific PCR primer design method for SNP marker analysis and its application. Plant Methods 8:34

Livingston JE (1942) The inheritance of resistance to *Ustilago nuda*. Phytopathology 32:451–466

Lüpken T, Stein N, Perovic D, Habekuss A, Krämer I, Hähnel U, Steuernagel B, Scholz U, Zhou R, Ariyadasa R, Taudien S, Platzer M, Martis M, Mayer K, Friedt W, Ordon F (2013) Genomics-based high-resolution mapping of the BaMMV/BaYMV resistance gene *rym11* in barley (*Hordeum vulgare* L.). Theor Appl Genet 126:1201–1212

Lüpken T, Stein N, Perovic D, Habekuss A, Serfling A, Krämer I, Hähnel U, Steuernagel B, Scholz U, Ariyadasa R, Martis M,

Mayer K, Niks RE, Collins NC, Friedt W, Ordon F (2014) High-resolution mapping of the barley *Ryd3* locus controlling tolerance to BYDV. Mol Breed 33:477–488

Mathre DE (1997) Compendium of barley diseases, 2nd edn. APS, St. Paul, pp 46–47

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

Menzies JG, Steffenson BJ, Kleinhofs A (2010) A resistance gene to *Ustilago nuda* in barley is located on chromosome 3H. Can J Plant Pathol 32:247–251

Metcalfe DR (1966) Inheritance of loose smut resistance. III. Relationships between the 'Russian' and 'Jet' genes for resistance and genes in 10 barley varieties of diverse origin. Can J Plant Sci 46:487–495

Metcalfe DR, Johnston WH (1963) Inheritance of loose smut resistance. II. Inheritance of resistance in three barley varieties to races 1, 2 and 3 of *Ustilago nuda* (Jens.) Rostr. Can J Plant Sci 42:390–396

Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb Protoc 2010: pdb-prot5448

Mueller KJ (2006) Susceptibility of German spring barley cultivars to loose smut populations from different European origins. Eur J Plant Pathol 116:145–153

Nirmala J, Brueggeman R, Maier C, Clay C, Rostoks N, Kannangara CG, von Wettstein D, Steffenson BJ, Kleinhofs A (2006) Subcellular localization and functions of the barley stem rust resistance receptor-like serine/threonine-specific protein kinase *Rpg1*. Proc Natl Acad Sci USA 103:7518–7523

Paterson AH, Bowers JE, Chapman BA (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc Natl Acad Sci USA 101:9903–9908

Popovic Z, Menzies JG, Matheson F, Recksiedler B, Knox R, Orr D, Ne Rauhala (1998) Cereal smut survey, 1997. Can Plant Dis Surv 78:82

Procunier JD, Xu J, Kasha KJ (1991) A rapid and reliable DNA extraction method for higher plants. Barley Genet Newsl 20:74–75

Proels RK, Hückelhoven R (2014) Cell-wall invertases, key enzymes in the modulation of plant metabolism during defence responses. Mol Plant Pathol 15:858–864

Rivandi J, Miyazaki J, Hrmova M, Pallotta M, Tester M, Collins NC (2011) A SOS3 homologue maps to *HvNax4*, a barley locus controlling an environmentally sensitive Na^+ exclusion trait. J Exp Bot 62:1201–1216

Robertson DW, Wiebe GA, Shands RG (1947) A summary of linkage studies in barley, Supplement I, 1940–1946. J Am Soc Agron 39:464–473

Schaller CW (1949) Inheritance of resistance to loose smut, *Ustilago nuda*, in barley. Phytopathology 39:959–979

Shahinnia F, Druka A, Franckowiak J, Morgante M, Waugh R, Stein N (2012) High resolution mapping of dense spike-ar (*dsp.ar*) to the genetic centromere of barley chromosome 7H. Theor Appl Genet 124:373–384

Shavrukov Y, Bovill J, Afzal I, Hayes JE, Roy SJ, Tester M, Collins NC (2013) HVP10 encoding V-PPase is a prime candidate for the barley *HvNax3* sodium exclusion gene: evidence from fine mapping and expression analysis. Planta 237:1111–1122

Silvar C, Perovic D, Scholz U, Casas AM, Igartua E, Ordon F (2012) Fine mapping and comparative genomics integration of two quantitative trait loci controlling resistance to powdery mildew in a Spanish barley landrace. Theor Appl Genet 124:49–62

Skoropad WP, Johnson LPV (1952) Inheritance of resistance to *Ustilago nuda* in barley. Can J Bot 30:525–536

Tekauz A (2003) Diseases of barley. In: Bailey KL, Gossen BD, Gugel RK, Morrall RAA (eds) Diseases of field crops in Canada, 3rd edn. Canadian Phytopathological Society and University Extension Press, University of Saskatchewan, Saskatoon, pp 30–53

The International Barley Genome Sequencing Consortium (2012) A physical, genetic and functional sequence assembly of the barley genome. Nature 491:711–716

The International Brachypodium Initiative (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature 463:763–768

Thomas PL, Menzies JG (1997) Cereal smuts in Manitoba and Saskatchewan, 1989–1995. Can J Plant Pathol 19:161–165

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

Zhang X, Han D, Zeng Q, Duan Y, Yuan F, Shi J, Wang Q, Wu J, Huang L, Kang Z (2013) Fine mapping of wheat stripe rust resistance gene *Yr26* based on collinearity of wheat with *Brachypodium distachyon* and rice. PLoS One 8:e57885