Theor Appl Genet
DOI 10.1007/s00122-007-0626-x

ORIGINAL PAPER

Haplotype analysis of vernalization loci in European barley germplasm reveals novel *VRN-H1* alleles and a predominant winter *VRN-H1/VRN-H2* multi-locus haplotype

James Cockram · Elena Chiapparino · Scott A. Taylor · Konstantina Stamati · Paolo Donini · David A. Laurie · Donal M. O'Sullivan

Received: 8 May 2007 / Accepted: 31 July 2007

Abstract In barley, variation in the requirement for vernalization (an extended period of low temperature before flowering can occur) is determined by the *VRN-H1*, *-H2* and *-H3* loci. In European cultivated germplasm, most variation in vernalization requirement is accounted for by alleles at *VRN-H1* and *VRN-H2* only, but the range of allelic variation is largely unexplored. Here we characterise *VRN-H1* and *VRN-H2* haplotypes in 429 varieties representing a large portion of the acreage sown to barley in Western Europe over the last 60 years. Analysis of genotype, intron I sequencing data and growth habit tests identified three novel *VRN-H1* alleles and determined the most frequent *VRN-H1* intron I rearrangements. Combined analysis of *VRN-H1* and *VRN-H2* alleles resulted in the classification of seventeen *VRN-H1/VRN-H2* multi-locus haplotypes, three of which account for 79% of varieties. The molecular markers employed here represent powerful diagnostic tools for prediction of growth habit and assessment of varietal purity. These markers will also allow development of germplasm to test the behaviour of individual alleles with the aim of understanding the relationship between allelic variation and adaptation to specific agri-environments.

Keywords Vernalization requirement · Barley · *VRN-H1* · *VRN-H2* · Haplotype

Communicated by J. Dubcovsky.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-007-0626-x) contains supplementary material, which is available to authorized users.

J. Cockram · K. Stamati · P. Donini · D. M. O'Sullivan (✉)
National Institute of Agricultural Botany,
Huntington Road, Cambridge, CB3 0LE, UK
e-mail: donal.osullivan@niab.com

S. A. Taylor · D. A. Laurie
John Innes Centre, Norwich Research Park,
Colney, Norwich, NR4 7UH, UK

Present Address:
E. Chiapparino
CRA-Istituto Sperimentale per la Cerealicoltura,
Sant'Angelo Lodigiano (Lodi), Italy

Introduction

The molecular and genetic basis of the winter/spring growth habit (GH)—formally assessed as the requirement or not for an extended period of low temperature treatment (vernalization) to induce flowering—has been extensively studied in temperate cereals (recently reviewed by Cockram et al. 2007). Variation for vernalization requirement in cultivated barley (*Hordeum vulgare* ssp. *vulgare*) is controlled by the major loci *VRN-H1* and *VRN-H2* on chromosomes 5HL and 4HL, respectively (Takahashi and Yasuda 1971; Laurie et al. 1995). Comparative mapping shows that these loci are collinear with the orthologous *VRN-A^{m}1* and *VRN-A^{m}2* vernalization requirement loci in the diploid wheat, *Triticum monococcum* L. (Dubcovsky et al. 1998). *VRN-H1* is also colinear with equivalent loci on the group 5 chromosomes in hexaploid wheat (*T. aestivum* L.) and rye (*Secale cereale* L.) (Laurie 1997). Spring alleles at a third barley vernalization locus, *VRN-H3*, confer extremely early flowering and are found mostly in exotic barley genotypes (Takahashi and Yasuda 1971). As the majority of European varieties are thought to be fixed for winter alleles at *VRN-H3*, they can effectively be viewed as segregating for *VRN-H1* and *VRN-H2* only (Yasuda et al. 1993). Genetically,

the wheat and barley vernalization loci behave similarly, with recessive *vrn-2* alleles and dominant or semi-dominant *Vrn-1* and *Vrn-3* alleles conferring spring GH (Takahashi and Yasuda 1971; Yan et al. 2006).

Fine mapping of *VRN-A^{m}1* in *T. monococcum* identified the candidate gene as a MADS-box transcription factor related to the *Arabidopsis* meristem identity genes *APETALA1* (*AP1*) and *FRUITFUL* (*FUL*) (Yan et al. 2003). Soon after, *VRN-A^{m}2* was found to be encoded by the putative transcription factor *ZCCT1* (Yan et al. 2004). In vernalization requiring *T. monococcum* lines, expression of *VRN-A^{m}2* is high before and during the first week of vernalization under long day photoperiods. Subsequently, expression drops in proportion to the duration of cold treatment, and is stably repressed once the vernalization requirement has been satisfied. This is correlated with the concomitant upregulation of *VRN-A^{m}1*, whose expression remains elevated on subsequent return to non-vernalizing temperatures, resulting in competence to flower. These findings, combined with the epistatic genetic interaction, suggest a model in which *VRN-A^{m}2* acts directly or indirectly to suppress expression of *VRN-A^{m}1* in unvernalized winter lines (Yan et al. 2003; Yan et al. 2004). In this model spring growth habit can be derived by mutation at *VRN-A^{m}2*, so that no functional repressor is present, or by mutation within *VRN-A^{m}1* regulatory regions resulting in alleviation of *VRN-A^{m}2* mediated repression. In *T. monococcum*, spring *Vrn-A^{m}1* alleles carry small deletions spanning a putative *cis*-regulatory motif (the CArG-box) within the promoter (Yan et al. 2003) or insertions within its first intron (Dubcovsky et al. 2006), while spring *vrn-A^{m}2* alleles contain a point mutation resulting in an amino acid change at a conserved residue within the CCT domain or a complete deletion of *ZCCT1* (Yan et al. 2004).

The basis of allelic variation within orthologous genes at the collinear *VRN1* loci in *T. aestivum* and barley has recently been studied. Spring GH conferred by the homoeologous *VRN-A1-B1* and *-D1* genes of hexaploid wheat is associated with intron I deletions thought to span a vernalization critical regulatory region (Fu et al. 2005). An orthologous gene, *HvBM5A* (Schmitz et al. 2000), has been mapped to the collinear *VRN-H1* locus in barley, and displays similar patterns of cold-induced expression as well as containing intron I deletions associated with spring GH (Danyluk et al. 2003; Trevaskis et al. 2003, 2006; Fu et al. 2005; Karsai et al. 2005; Yan et al. 2005; von Zitzewitz et al. 2005; Dubcovsky et al. 2006; Szűcs et al. 2006, 2007). Comparison of wheat and barley sequence and deletion breakpoints has identified a 436 bp region of *VRN-H1* intron I as representing a putative vernalization critical region (Fu et al. 2005; von Zitzewitz et al. 2005). In barley, three *ZCCT*-like genes (*ZCCT-Ha, ZCCT-Hb* and *ZCCT-Hc*) have been mapped to the *VRN-H2* locus (Yan et al. 2004; Dubcovsky et al. 2005; von Zitzewitz et al. 2005; Karsai et al. 2005; Szűcs et al. 2006, 2007), with alleles associated with spring GH created by deletion of all three genes in almost all spring varieties examined to date (Yan et al. 2004; Dubcovsky et al. 2005; Karsai et al. 2005; von Zitzewitz et al. 2005).

The identification of *VRN-H1* and candidate genes for the *VRN-H2* locus in barley permits characterization of germplasm collections and identification of the alleles and allelic combinations utilized in modern cultivars, as well as providing the basis from which novel alleles and allelic combinations controlling vernalization requirement can be identified and assessed. In this paper, we describe the first large-scale survey of haplotype diversity of *VRN-H1* and *VRN-H2* in European germplasm, with the aim of identifying and classifying allelic variation at these loci, and to validate markers towards a diagnostic test for vernalization response. The haplotype data presented here will provide the basis from which breeders may accurately identify, utilize and trace allelic variation at the *VRN-H1* and *VRN-H2* in their breeding material. In addition, this data will permit identification and utilization of novel alleles or allele combinations not previously deployed.

Materials and methods

Germplasm, DNA extraction, phenotyping and nomenclature

A total of 429 spring, winter and facultative barley commercial varieties were sampled from 13 EU countries, selected to represent cultivated EU germplasm grown over the last 60 years. DNA extraction protocols and further details of plant varieties are described by Chiapparino et al. (2006). Growth habit scores were obtained from passport data, but where there was a need to verify GH, five individual plants from selected cultivars were phenotyped for heading date in 13-week glasshouse tests under long day photoperiods at 20°C. Under these conditions, winter varieties tend to flower very late. A late flowering phenotype was recorded if plants had not flowered after thirteen weeks, and were tentatively designated as displaying a winter GH. Flowering was recorded at developmental stage 49, according to the Zadock's scale (Tottman and Makepeace 1979). The Spanish varieties "Candela" and "Orria", listed as winter in their passport data, are better described as facultative (Dr Igartua, personal communication), a phenotypically spring GH class that can be regarded as "cold tolerant, vernalization unresponsive" (von Zitzewitz et al. 2005). It should be noted the barley vernalization loci and genes have been variously termed, and each has several pseudonyms: *VRN-H1* was initially termed *Sh2* or *Sgh2*, with the

candidate gene that underlies it variously termed *BM5A, HvBM5A* and *HvAP1a*. Similarly, *VRN-H2* was previously named *Sh* or *Sgh*. Here we use the vernalization locus nomenclature proposed by Dubcovsky et al. (1998) in combination with *HvBM5A* to denote the MADS-box gene that underlies *VRN-H1*. The three closely related members of the *ZCCT-H* gene family (*ZCCT-Ha*, *-Hb* and *-Hc*) follow the nomenclature established by Dubcovsky et al. (2005).

Surveying *VRN-H1* nucleotide sequence

A preliminary survey of genomic *VRN-H1* nucleotide variation was conducted by sequencing 4.2 kb from four gene fragments amplified from 11 varieties ("Arra", "Dandy", "Golden Promise", "Golf", "Igri", "Optic", "Panda", "Pearl", "Pohto" "Prisma" and "Triumph") chosen to represent all GH/row-number combinations released over a span of more than 60 years from a variety of western European countries, using the protocols described by Chiapparino et al. (2006). Complete sequence coverage was obtained for all exons, introns III, IV, V, VI and VII, with partial coverage of the promoter, 5′ UTR, 3′ UTR, and introns I and II.

Genotyping *VRN-H1* and *VRN-H2* polymorphisms

Ten *VRN-H1* markers, which discriminated between all the haplotypes identified were selected for genotyping across the complete varietal collection (Fig. S1). SNPs in the promoter (T-1,948/C, A-1,881/G, T-1,655/C,) and within intron VII (T + 14,567/C, G + 14,585/A) and the 3′ UTR (C + 14,828/G) were selected to identify any recombination that may have taken place within the relatively large *VRN-H1* gene, as exemplified by partial genomic sequence of "Tremois" (von Zitzewitz et al. 2005). SNPs were genotyped by dideoxy single-base primer-extension chemistry using the ABI PRISM® SNaPshot® Multiplex Kit (Applied Biosystems), as described in Chiapparino et al. (2006). Markers distributed within the intervening regions were as follows: a $(CGCT)_{2–5}$ SSR located in the 5′ UTR, previously published PCR/agarose-gel based assays (Fu et al. 2005) for the 5.2 kb intron I deletion that spans a vernalization critical region (von Zitzewitz et al. 2005; Fu et al. 2005) present in the spring cultivar "Morex", and for the absence of large intron I deletions, characteristic of the winter cultivar "Strider", as well as a 42 bp InDel towards the 3′ end of intron I, first identified by von Zitzewitz et al. (2005). Complete sequencing of *VRN-H1* intron I was conducted from one variety selected from each *VRN-H1* haplotype identified, and complete *VRN-H1* gene sequence (−2,074 to +14,952 bp) determined in the variety "Xenia". Ultimately, intron I deletion configuration for all lines were determined by PCR/agarose-gel based amplicon analysis, based on sequences presented here and previously published alleles (Fu et al. 2005; von Zitzewitz et al. 2005; Szűcs et al. 2007) using the primers listed in Table S1. All PCR assays based on presence/absence of amplicons were performed with a minimum of two replicates. A genotype assay permitting amplification of all three *ZCCT-H* genes at the *VRN-H2* locus was performed as described by Karsai et al. (2005). The *VRN-H3* allele present in "Xenia" was determined by direct sequencing of PCR amplicons, according to the winter and spring SNP associations reported by Yan et al. (2006). Primer sequences for *VRN-H1* amplification and sequencing, as well as all previously unpublished SSR, SNaPshot® and InDel assays are listed in Table S1. All *VRN-H1* SNP positions are annotated relative to the translation start site (+1 bp) in genomic sequence of the winter *vrn-H1* allele from the cultivar, "Strider" (AY750993). Details and accession numbers for all genomic *VRN-H1* regions sequenced are presented in Table S2.

Results

Pre-screening for *VRN-H1* polymorphism

To help establish whether published sequence adequately described allelic variation in European cultivated barley, partial genomic sequencing covering 4.2 kb was undertaken on a subset of 11 representative varieties (including winter and spring lines), and compared to previously published *VRN-H1* genomic sequences (Fu et al. 2005; von Zitzewitz et al. 2005; Szűcs et al. 2007). Forty-two polymorphic features were identified: 33 SNPs, 5 InDels and 4 SSRs (Table S3). The previously identified synonymous substitution at +13,745 bp within exon IV (CAA → CAG, Gln) was the only polymorphism identified within coding regions, supporting the theory that differences in GH are not due to changes in the predicted VRN-H1 protein (von Zitzewitz et al. 2005). One novel polymorphism (C+14,612/T) was identified in the spring cultivar "Dandy".

Alignment of the sequenced regions with previously published genomic *VRN-H1* sequence identified five *VRN-H1* haplotypes (Table S3). All three winter varieties sequenced here were grouped into the same haplotype, along with the previously published winter *VRN-H1* allele from the US variety "Strider". SNPs identical to those present in the winter *VRN-H1* allele were also found in the facultative lines "Dicktoo" and "88Ab536" (von Zitzewitz et al. 2005), as well as the spring variety "Arra". The two facultative lines have previously been shown to contain a deletion of the *ZCCT-H* genes (von Zitzewitz et al. 2005), which accounts for the observed spring GH. The spring phenotype of "Arra" is consistent with mutations in other regions of *VRN-H1*, explained subsequently (Table S2). The alignment in Table S3 shows the spring variety

"Tremois" to contain a promoter characteristic of spring varieties, while its partial 5′ UTR and intron I sequence are identical to those in the winter allele. Based on this preliminary sequence survey, a 6-SNP haplotype set was chosen that identified the five *VRN-H1* variants, including the putative recombination observed in "Tremois".

Investigation of germplasm with conflicting predicted and recorded growth habit scores

The first stage of genotyping across the complete varietal collection involved screening ten *VRN-H1* markers (Fig. S1) in conjunction with an assay testing for the presence of the *VRN-H2* candidate genes, *ZCCT-Ha*, *-Hb* and *-Hc*. *VRN-H1* groupings were made as follows: primary classification into haplogroups 1–5 was based on genotypes at SNP1 to SNP6, the $(CGCT)_n$ SSR in the 5′ UTR and the 42 bp intron I InDel, reflecting the overall pattern of sequence divergence across the gene. Sub-division of these groupings was made on the basis of the presence or absence of large intron I deletions, and initially resulted in the identification of eight *VRN-H1* haplotypes (1A, 1B, 2, 3, 4A, 4B, 5A and 5B), shown in Table 1. The predominant winter haplotype was designated 1A, with the order of subsequent haplogroups based on increasing marker differentiation. As GH depends on the allelic status at both *VRN-H1* and *VRN-H2* (Takahashi and Yasuda 1971), a haplotype nomenclature incorporating genotypes across both loci was established, by appending +Z or −Z (presence or absence of the three *ZCCT-H* genes) to the *VRN-H1* haplotype (Table 1). This allowed the diagnostic value of molecular marker combinations to be established by comparison of predicted GH (as inferred from multi-locus haplotypes) and recorded GH (as indicated in passport data), and permitted identification of putative errors in recorded GH.

Based on their similarity to previously known alleles (Tables 1, S3), haplotype 1A was predicted to correspond to the winter ("Strider") allele (von Zitzewitz et al. 2005), haplotypes 4B and 5A to correspond to spring alleles on the basis of the 5.2 kb intron I deletion (Fu et al. 2005; von Zitzewitz et al. 2005) and haplotypes 2, 3, 4A and 5B to the spring *VRN-H1* alleles of "Tremois" (von Zitzewitz et al. 2005), "Oregon Wolfe Barley Dominant" (Fu et al. 2005; von Zitzewitz et al. 2005), "Triumph" (von Zitzewitz et al. 2005) and "Calicachuma-sib" (Szűcs et al. 2007), respectively. Breakpoint analysis of haplotype 1B showed it did not contain the intact intron I amplicon characteristic of winter varieties, nor did it contain the 5.2 kb intron I deletion, and was thus assumed to carry a different intron I deletion affecting GH. All -Z haplotypes were also predicted to exhibit spring GH.

Table 1 *VRN-H1* and *ZCCT-H* multi-locus haplotypes

Haplotype	SNP1	SNP2	SNP3	SSR	Major IntronI Del (kb)	IntronI Ins 699bp	InDel 42bp	SNP4	SNP5	SNP6	*ZCCT-H* geno	Predicted *VRN-H1* allele	Predicted *VRN-H2* allele	Predicted plant GH	Recorded plant GH
1A+Z	T	A	C	2	0	0	1	T	G	C	+Z	W	W	W	W, U*
1A-Z	T	A	C	2	0	0	1	T	G	C	-Z	W	S	S	S
1B-Z	T	A	C	2	3.9	0	1	T	G	C	-Z	S	S	S	S
2+Z	C	G	C	2	0	1	1	T	G	C	+Z	S	W	S	S
2-Z	C	G	C	2	0	1	1	T	G	C	-Z	S	S	S	S
3+Z	T	G	T	2	6.3	0	0	C	A	C	+Z	S	W	S	S
3-Z	T	G	T	2	6.3	0	0	C	A	C	-Z	S	S	S	S
4A+Z	C	G	C	1	8.9	0	0	C	G	C	+Z	S	W	S	S
4A-Z	C	G	C	1	8.9	0	0	C	G	C	-Z	S	S	S	S
4B+Z	C	G	C	1	5.2	0	0	C	G	C	+Z	S	W	S	S
4B-Z	C	G	C	1	5.2	0	0	C	G	C	-Z	S	S	S	S
5A+Z	T	G	C	3	5.2	0	0	C	A	G	+Z	S	W	S	S, F
5A-Z	T	G	C	3	5.2	0	0	C	A	G	-Z	S	S	S	S
5B+Z	T	G	C	3	4.1	0	0	C	A	G	+Z	S	W	S	S
5B-Z	T	G	C	3	4.1	0	0	C	A	G	-Z	S	S	S	S
5C+Z	T	G	C	3	0.5	0	0	C	A	G	+Z	W	W	W	W
5C-Z	T	G	C	3	0.5	0	0	C	A	G	-Z	W	S	S	S

VRN-H1 markers | *VRN-H2* markers

Haplotype frequency, predicted *VRN-H1* and *VRN-H2* alleles and GH, as well as recorded plant GH, are indicated: *S* spring, *F* facultative, *W* winter, *U* unclassified. Ninety-seven percent of winter varieties belong to haplotype 1A + Z, however, four winter varieties ("Almunia", "Athene", "Birgit" and "Express") belong to haplotype 5C + Z. Two putative recombination events can be identified: (1) haplotype 2 combines a promoter SNP configuration characteristic of haplogroup 4 (boxed in solid black) into an otherwise winter 1A haplotype, associated with a spring GH. (2) The intron I deletion characteristic of the cultivar "Morex" (haplotype 5A) is also found in haplotype 4B (boxed in dashed black), suggesting a double recombination event within intron I. Genotype scores: *VRN-H1* 0.7 kb insertion: *0* insertion absent; *1* insertion present. *VRN-H1* 42 bp InDel: *0* smaller allele; *1* larger allele. SSR: *1* $(CGCT)_2$, *2* $(CGCT)_4$, *3* $(CGCT)_5$. *ZCCT-H* locus: *+Z* ZCCT-H gene cluster present; *−Z* *ZCCT-H* gene cluster absent. *No genotypic explanation was found for the intermediate flowering time of the variety "Xenia" (haplotype 1A + Z)

The recorded GH of 417 out of the 429 varieties agreed with the predicted GH on this basis. In the twelve cases where recorded GH was not explained by previously described allelic variation at *VRN-H1* or *VRN-H2*, further phenotypic evaluation and investigation of genotype was undertaken. In 7 out of these 12 varieties, the GH predicted by genotype was supported by our evaluation of flowering time under LD non-vernalizing conditions, and were not investigated further (corrected phenotypes are reported in Table S4). Four of the remaining five varieties ("Almunia", "Athene", "Birgit" and "Express") displayed apparent 5B + Z genotypes (predicted spring GH) whereas both passport GH designation and our phenotypic evaluation suggested they carried a winter growth habit. As close pedigree relationships exist between three of these four varieties ("Athene" and "Birgit" share 50% parentage, while "Athene" is a parent of "Express"), we postulated this group carries a putative novel winter *VRN-H1* allele, which we designated 5C (genotype details for which are reported below). Despite possessing the preliminary 1A + Z haplotype predominant among winter European barley varieties, "Xenia" flowered much later than all other spring varieties (80 days after germination), but 11 days before the arbitrary threshold of 13 weeks which was set to distinguish winter and spring GH. It is not established whether vernalization accelerates flowering in "Xenia", so a weak vernalization response cannot be discounted. Accordingly, GH for this variety remains unclassified, pending further investigation of flowering time in response to vernalization.

Characterisation of major *VRN-H1* intron I rearrangements

In barley, large deletions that span a vernalization critical region within *VRN-H1* intron I are thought to be responsible for the creation of vernalization insensitive spring alleles (Fu et al. 2005; von Zitzewitz et al. 2005; Szűcs et al. 2007). With the aim of verifying the intron I configurations of all the remaining spring *VRN-H1* haplotypes defined here, and to characterise the putative *VRN-H1* winter 5C allele, we sequenced intron I from one variety from each of the nine *VRN-H1* haplotypes. A total of seven major intron I rearrangements were identified (Fig. 1), with four of these having previously been reported: the 6.3 kb deletion present in haplotype 3 (Fu et al. 2005; von Zitzewitz et al. 2005; Szűcs et al. 2007), the 8.9 kb deletion present in haplotype 4A (von Zitzewitz et al. 2005), the 5.2 kb deletion present in haplotypes 4B and 5A (Fu et al. 2005; von Zitzewitz et al. 2005) and the 4.1 kb deletion present in haplotype 5B (von Zitzewitz et al. 2005; Szűcs et al. 2007). In addition, we found that the 10.8 kb intron I sequence of the European winter variety "Panda" was identical to the previously sequenced US winter variety "Strider" apart from the deletion of a region including a MITE element, as previously

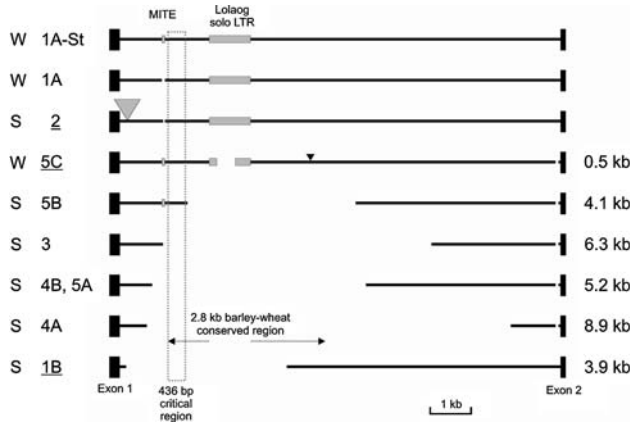

Fig. 1 Schematic representation of intron I from an exemplar variety from each *VRN-H1* haplotype: "Panda" (*1A*), "Etu" (*1B*), "Varunda" (*2*), "Dandy" (*3*), "Optic" (*4A*), "Prisma" (*4B*), "Golden Promise" (*5A*), "Oriol" (*5B*), "Express" (*5C*) and the reference winter *vrn-H1* allele from "Strider" (AY750993, *1A-St*). Deletion sizes are indicated to the right of the figure. The intron I sequence of the spring variety "Xenia" is identical to "Strider", apart from an additional base at +4,391 bp. Haplotypes with novel intron I deletions are underlined. The predicted *VRN-H1* allelic state, spring (*S*) or winter (*W*), is indicated. *Black boxes* indicate exon 1 and exon 2. The positions of the MITE and Lolaog solo LTR are indicated by *grey boxes*. *Gaps* represent deletions relative to the winter *vrn-H1* "Strider" allele. The insertion position of a putative transposable element and a 24 bp insertion are indicated by *grey* and *black triangles*, respectively. Positions of the 2.8 kb conserved region between barley and wheat, and the vernalization critical region defined by von Zitzewitz et al. (2005), are indicated

observed in a sequenced intron I fragment from the winter variety "Igri" (von Zitzewitz et al. 2005), a T/C SNP at +4,391 bp and an additional base at +4,246 bp.

Three novel *VRN-H1* intron I rearrangements were identified (Fig. 1). (1) The spring variety "Etu" (haplotype 1B) contained a novel 3.9 kb intron I deletion removing the putative vernalization critical region (Fu et al. 2005; von Zitzewitz et al. 2005). This deletion represents the furthest extent of any deletion breakpoint towards the 5′ end of intron I identified to date, while the 3′ breakpoint is located close to the end of the 2.8 kb conserved region between barley and wheat. (2) Sequencing of haplotype 2 ("Varunda") showed a 0.7 kb insertion was identified 170 bp downstream of the start of intron I. Sequence analysis indicates this insertion is likely to be a transposable element, as it shows characteristic features including short host sequence duplication and inverted repeats at its ends flanking imperfect direct repeats. Apart from this insertion, "Varunda" intron I sequence was identical to the winter variety "Panda" and partial sequence from "Igri" (von Zitzewitz et al. 2005), suggesting this novel allele represents a recent insertion event associated with spring GH in this variety. PCR analysis indicates this insertion is present in the spring variety "Tremois", providing further evidence that it displays *VRN-H1* haplotype 2 (data not shown). (3)

Sequencing of haplotype 5C exemplar variety "Express" showed that it contained a largely intact intron I with a previously undescribed 486 bp deletion. This deletion is not within the putative vernalization critical region (Fu et al. 2005; von Zitzewitz et al. 2005), and is unlikely to affect gene regulation or function as it removes part of a retrotransposon solo Long Terminal Repeat (LTR) found in intron I of the winter cultivar "Strider".

Further investigation of the basis for intermediate GH in "Xenia"

VRN-H1 intron I was also sequenced in the intermediate flowering time variety "Xenia", as preliminary multi-locus haplotyping predicts winter alleles at *VRN-H1* and *VRN-H2* (1A + Z). We found "Xenia" to contain the MITE and flanking regions present in the winter *VRN-H1* allele of the US variety "Strider", combined with the additional base at +4,431 bp identified in the European winter allele sequenced here. Thus, intron I sequencing in this variety failed to account for its intermediate flowering date. Accordingly, we sequenced the "Xenia" *VRN-H1* allele from −2,074 to +14,953 bp. No further sequence differences were identified between "Xenia" and "Strider" *VRN-H1* alleles (although the inability to completely sequence across a long "TA" repeat motif in the promoter prevents discounting this region as polymorphic to the "Strider" allele). To determine whether the intermediate phenotype of "Xenia" could be explained by spring alleles at *VRN-H3*, the SNP markers at positions 270 and 384 bp from the start of intron I, reported to be associated with spring alleles by Yan et al. (2006) were genotyped. However, as the "TC" SNP configuration at these positions shows "Xenia" has a winter *vrn-H3* haplotype, the intermediate flowering phenotype remains unexplained.

Frequency of multi-locus haplotypes and correlation with growth habit

Identification of the seven major *VRN-H1* intron I rearrangements listed above was followed by a final phase of genotyping, in which intron I configuration in each of the 429 varieties was determined by a new set of PCR/agarose-gel based assays based on the seven fully determined intron I sequences (primer sequences and combinations listed in Table S1; genotype data presented in Table S4). Breakpoint analysis showed that all members of spring haplotypes 1B, 2, 3, 4A, 4B, 5A and 5B as well as winter haplotype 5C were, without exception, positive for the deletion exhibited by the exemplar variety for the group, confirming the robustness of these groupings.

In this survey, all winter varieties contained the *ZCCT-H* locus (preliminary haplotypes 1A + Z, 115 varieties; 5C + Z, 4 varieties), consistent with a requirement for the presence of these genes in winter germplasm. The majority of spring lines carry a deletion of the *ZCCT-H* locus. However, 13% of spring varieties were found to contain all three *ZCCT-H* genes (Fig. 2; Table S4). Thus, a total of nine *VRN-H1* haplotypes were identified that, when combined with the *ZCCT-H* genotypes, resulted in seventeen multi-locus haplotypes (Table 1). Full genotypes and recorded GH for all varieties analysed are presented in Table S4 and online at http://www.niab.com/research/researchpublications/.

Three common multi-locus haplotypes (1A + Z, 4A-Z and 5A-Z) account for 27, 26 and 26% of the collection, respectively (Fig. 2). The remaining 14 haplotypes account for just 21% of the collection, and are each found at a frequency of <4%. Haplotype groupings show a clear distinction between spring and winter GH, with much greater genetic diversity identified in spring relative to winter germplasm, as illustrated by the number of spring multilocus haplotypes identified (Fig. 2). Haplotype 1A + Z, predicted to be the only possible winter haplotype based on previously described *VRN-H1* alleles, accounts for 98% of winter varieties. However, the discovery of a novel haplotype (5C + Z) in winter varieties "Athene", "Almunia", "Birgit" and "Express" (haplotype 5C + Z) shows a second class of winter *VRN-H1* allele is present at low frequency in European germplasm. Predicted GH scores, as defined by multi-locus haplotypes of *VRN-H1* and *VRN-H2*, were therefore consistent with observed phenotype, with the sole exception of "Xenia", for which acceleration of flowering by vernalization has not been determined.

The markers utilised in this study to survey and classify *VRN-H1* alleles can be reduced to a combination of intron I

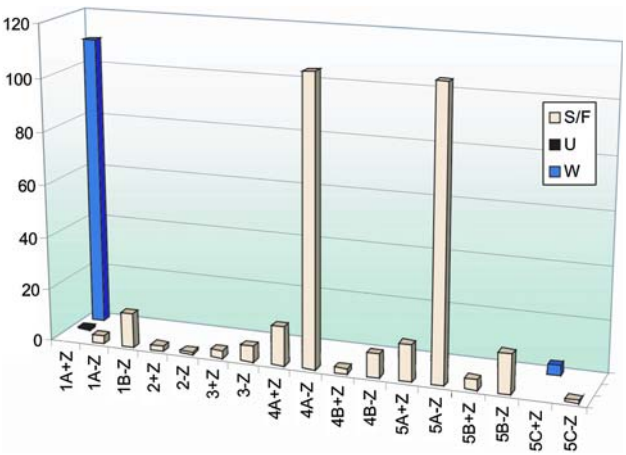

Fig. 2 Histogram of *VRN-H1* and *VRN-H2* multi-locus haplotype frequencies in spring (S), facultative (F) and winter (W) barley cultivars. All but four winter varieties ("Almunia", "Athene", "Birgit" and "Express", haplotype 5C + Z) are classified as haplotype 1A + Z. The variety "Xenia"of unclassified growth habit displays a 1A + Z haplotype

deletion type and SNP1, providing a minimal marker set that define each of the haplotypes identified. Analysis of the large germplasm set screened here shows the previously published PCR assay, reported to distinguish between winter and spring *VRN-H1* alleles (Fu et al. 2005), would incorrectly classify *VRN-H1* haplotype 5C alleles as spring, as the 486 bp intron I deletion within the solo LTR removes the forward primer binding site. Modification of this assay enables detection, and discrimination between the haplotype 1A and 5C winter *VRN-H1* alleles, as illustrated in Fig. S2.

Discussion

No frame-shifts, non-synonymous substitutions or premature stop codons within *VRN-H1* coding regions were identified. This suggests a highly conserved VRN-H1 peptide is likely to be essential for plant growth and reproduction, irrespective of GH, and that the differentiation of spring and winter alleles is due to polymorphism in putative *cis*-regulatory regions outside the exons. In contrast to the coding regions, plentiful variation was identified in the promoter, introns and UTRs, permitting identification of a set of polymorphisms that allowed the survey of germplasm for intragenic recombination as well as novel *VRN-H1* haplotypes. Both scenarios were encountered, and this information was critical in differentiating the predominant winter haplotype from spring-associated haplotypes that would otherwise have confounded the correlation between *VRN-H1* genotypes and vernalization requirement.

This study identified two novel intron I rearrangements likely to result in spring alleles at *VRN-H1* (haplotypes 1B and 2). These, in combination with the four previously published deletions (Fu et al. 2005; von Zitzewitz et al. 2005; Szűcs et al. 2007) result in a total of six known spring *VRN-H1* intron I rearrangements (Table 1). Analysis of the relative effects these alternative spring alleles have on flowering time may help further refine the vernalization critical region thought to be involved in the cold-activated transcriptional control of expression, as well as allowing the relative strengths of these spring alleles to be determined. Interestingly, the presence of three types of intron I deletion within *VRN-H1* haplogroup 5 and their association with winter (486 bp deletion) and spring (4.1 and 5.2 kb deletions) varieties, suggests the former rearrangement does not affect GH. The small deletion in the novel winter allele identified in this study is located within the retrotransposon solo LTR present in all winter barley accessions sequenced to date. Although the solo LTR is located in a region of intron I that is highly conserved between barley and wheat, the absence of a similar element in winter wheat means it is not evolutionarily conserved and therefore less likely to be essential for cold-responsive regulation of *VRN-H1*. The location and size of the deletion found in the four winter haplotype 5C + Z varieties, suggests it does not result in significant disruption of a vernalization critical region, since winter GH is retained (however, it should be noted that "Express" intron I contains six additional SNPs relative to the winter "Strider" allele within the 2.8 kb conserved region of intron I). As vernalization requirement is not a bimodal character, but shows continuous phenotypic distribution between spring and winter GH (Takahashi and Yasuda 1971; Szűcs et al. 2007), further phenotypic evaluation is required to determine the relative strength of this novel variant. This study found just one variety ("Xenia") whose relatively early flowering time in non-vernalizing LD conditions compared to the winter varieties screened could not be predicted by its multi-locus haplotype (1A + Z) or genomic sequence presented here. Despite determining the full sequence of its *VRN-H1* allele and a portion of *VRN-H3*, we found only typically winter haplotypes at both loci. Therefore, variation conferring earliness in "Xenia" may reside either in as yet unexplored regions of vernalization genes (e.g. mutations within the *ZCCT-H* genes) or in additional loci controlling flowering time. Further phenotypic evaluation is required to determine whether "Xenia" possesses a vernalization requirement and should therefore be considered as be an "early winter" rather than a "late spring" type.

The presence of just two allelic states at the *ZCCT-H* locus, suggests that only spring-associated variants in which all three *ZCCT-H* genes are deleted have been widely utilized in European spring barley germplasm. Detailed analysis of deletion breakpoints at the *ZCCT-H* locus over multiple varieties would be required to verify whether all lines carry the same ancestral deletion, or whether multiple mutation events with the same effect have occurred. Interestingly, previous studies have identified two examples of alternative deletion variants: one single and one double *ZCCT-H* gene deletion, both found in non-European varieties (Yan et al. 2004; Dubcovsky et al. 2005; Szűcs et al. 2007), demonstrating additional *ZCCT-H* locus deletions are present in cultivated barley. Here we show that a significant minority of European spring varieties still carry the *ZCCT-H* gene cluster.

In contrast to the relatively frequent occurrence of winter alleles at the *ZCCT-H* locus in spring varieties, the invariant incidence of spring *VRN-H1* alleles in confirmed spring European germplasm is curious, as the genetic model of interaction between *VRN-H1* and *VRN-H2* predicts that spring alleles at either locus result in spring GH. This suggests that breeders have knowingly, or unknowingly, selected for spring alleles at *VRN-H1* in preference to those at *VRN-H2*, which might be explicable if the former confers a stronger spring phenotype. However, this is not supported

by recent analysis of flowering time in a "Morex" × *H. vulgare* ssp. *spontaneum* F_2 population, where the presence of spring *Vrn-H1* alleles in combination with winter *Vrn-H2* alleles, and vice versa, result in identical flowering times (Dubcovsky et al. 2005). The unequal distribution of spring alleles at *VRN-H1* and *VRN-H2* in spring germplasm may reflect relative mutation frequencies combined with the comparative ease with which dominant alleles (i.e. spring *Vrn-H1* alleles) could be selected over recessive ones (spring *vrn-H2* alleles), with the resulting skew in landrace allele frequencies still evident in the modern cultivars which succeeded them.

The presence of just one major winter multi-locus haplotype (1A + Z) that is (with the exception of the phenotypically undetermined variety "Xenia") composed uniquely of winter varieties and that accounts for 97% of all winter varieties screened, is curious given the winter form is believed to be ancestral (Takahashi and Yasuda 1971) and therefore potentially more diverse. However, European winter barley is widely recognised as having a much narrower germplasm base than spring cultivars. Winter types were not widely grown in Europe until the 1950s (Fischbeck 2003; Melchinger et al. 1994), with the majority of 6-row winter varieties developed from just five progenitors ("Dea", "Firlbecks" "4zlg", "Hauters" and "Vogelsanger Gold"). This, along with the requirement to maintain winter alleles at all three VRN loci is likely to explain the lack of *VRN-H1* haplotype variation in European winter barley. The comparatively large number of distinct spring *VRN-H1* haplotypes may reflect strong selection for spring *VRN-H1* alleles during and post-domestication.

VRN-H1 allele evolution

Drawing evolutionary relationships between *VRN-H1* haplotypes is problematic due to the complex interplay between mutation and recombination. However, some relationships can be deduced from the genotyping and sequencing undertaken here. (1) The subdivision of haplogroups 4 and 5 into haplotypes 4A/4B and 5A/5B/5C results solely from the presence of different intron I deletions, indicating shared ancestry between members of respective haplogroups. *VRN-H1* intron I sequencing suggests a series of deletions of increasing size could account for the conversion of winter to spring alleles within haplogroup 5 (Fig. 1). More likely, haplogroup 5 deletion forms could have arisen independently from a single ancestral allele. (2) The observation that haplotypes 5A and 4B share the same deletion breakpoints suggest a double recombination event involving intron I or gene conversion may have taken place. Although formally possible, it seems unlikely that an identical deletion has taken place independently on more than one occasion. (3) The spring *VRN-H1* haplotype 1B is identical to winter haplotype 1A except for the presence of a novel 3.9 kb deletion, suggesting the former is recently derived from the latter, and may represent a recent derivation of a spring allele from a winter *VRN-H1* allele. (4) Evidence of recombination within the *VRN-H1* promoter was identified: varieties belonging to *VRN-H1* haplotype 2 ("Dram", "Mazurka" and "Varunda") are unusual in matching winter haplotype 1A genotypes at all *VRN-H1* markers, except at the first two promoter SNPs (SNP configuration "CG"). This SNP combination is characteristic of spring *VRN-H1* haplotypes 4A and 4B, indicating the promoter region has likely recombined upstream of the $(CAGT)_n$ SSR in the 5′ UTR, and may contribute to the creation of the predicted spring allele (Table 1). Evidence of similar recombination can be seen by comparison with previously published genomic fragments from the spring variety "Tremois" (von Zitzewitz et al. 2005), which suggest it may belong to the same haplotype as "Varunda". The recent finding that the US spring variety "Oregon Wolfe Barley Dominant" contains a SNP within the putative promoter CArG box regulatory motif (whose disruption is thought to be responsible for spring alleles at the collinear locus of the wheat A genome) supports the existence of promoter polymorphisms that could affect vernalization requirement in barley (Szűcs et al. 2007), although this variety also carries a 6.3 kb intron I deletion characteristic of haplotype 3. The *VRN-H1* sequence of "Varunda" also differs from winter alleles by the insertion of a transposable element (TE) upstream of the putative vernalization critical region proposed by von Zitzewitz et al. (2005) and Fu et al. (2005). This insertion may regulate *VRN-H1* expression due to an effect on transcript processing and regulation inherent to the presence of the TE, as shown in the case of a TE insertion within the first intron of the *Arabidopsis* MADS-box gene *FLC* (Liu et al. 2004). Alternatively the barley TE insertion could result in the physical disruption of a previously unrecognised motif critical for cold-mediated regulation. PCR analysis indicates "Tremois" contains an intron I TE insertion of similar size to that found in haplotype 2, and may explain its dominant spring *Vrn-H1* allele. Interestingly, insertion of a transposable element within intron I is reported to be associated with spring alleles at an orthologous *VRN* locus in wheat (Dubcovsky et al. 2006) and in the promoter of *VRN-B3* (Yan et al. 2006), indicating the activity of such elements may play an important role in the creation of novel alleles at agronomically important loci in cereals.

Potential use of diagnostic markers for vernalization requirement in breeding and varietal identification

The germplasm survey undertaken in this study has resulted in the characterization of three novel *VRN-H1* alleles present at low frequencies within EU barley. The combination

ZCCT-H markers (Karsai et al. 2005) with PCR assays which take account of these novel *VRN-H1* alleles amount to a diagnostic test capable of predicting the GH of all but one ("Xenia") of the 429 European barley lines sampled here. This assay should find application in breeding programmes moving alleles between winter and spring backgrounds, purity testing, and the assessment of alleles or allelic combinations on flowering-time and agronomic traits. Furthermore, the definition of a series of putative *cis*-regulatory motif disruptions allows the performance of these alleles and allelic combinations in different agri-environments to be assessed and utilized accordingly.

Acknowledgments This work was supported in part by Defra project VS0137 and the GEDIFLUX project (contract No. QLK5-CT-2001-00934), supported by the European Commission. S.T. was funded by grant 208/D14003 from the UK Biotechnology and Biological Sciences Research Council, UK.

References

Chiapparino E, Donini P, Reeves J, Tuberosa R, O'Sullivan DM (2006) Distribution of β-amylase I haplotypes among European cultivated barleys. Mol Breed 18:341–354

Cockram J, Jones H, Leigh FJ, O'Sullivan D, Powell W, Laurie DA, Greenland AJ (2007) Control of flowering time in temperate cereals: genes, domestication and sustainable productivity. J Exp Bot 58:1231–1244

Danyluk J, Kane NA, Breton G, Limin AE, Fowler DB, Sarhan F (2003) TaVRT-1, a putative transcription factor associated with vegetative to reproductive transition in cereals. Plant Physiol 132:1849–1860

Dubcovsky J, Lijavetzky D, Appendino L, Tranquilli G (1998) Comparative RFLP mapping of *Triticum monococcum* genes controlling vernalization requirement. Theor Appl Genet 97:968–975

Dubcovsky J, Chen C, Yan L (2005) Molecular characterization of allelic variation at the *VRN-H2* vernalization locus in barley. Mol Breed 15:395–407

Dubcovsky J, Loukoianov A, Fu D, Valarik M, Sanchez A, Yan L (2006) Effect of photoperiod on the regulation of wheat vernalization genes *VRN1* and *VRN2*. Plant Mol Biol 60:469–480

Fischbeck G (2003) Diversification through breeding. In: von Bothmer R, van Hintum T, Khuepffer H, Sato K (eds) Diversity in barley (*Hordeum vulgare*). Elsevier, Amsterdam, pp 29–52

Fu D, Szűcs P, Yan L, Helguera M, Skinner JS, von Zitzewitz J, Hayes PM, Dubcovsky J (2005) Large deletions within the first intron in *VRN-1* are associated with spring growth habit in barley and wheat. Mol Genet Genomics 273:54–65

Karsai I, Szűcs P, Mészáros K, Filichkina T, Hayes PM, Skinner JS, Láng L, Bedő Z (2005) The *Vrn-H2* locus is a major determinant of flowering time in a facultative × winter growth habit barley (*Hordeum vulgare* L.) mapping population. Theor Appl Genet 110:1458–1466

Laurie DA, Pratchett N, Bezant JH, Snape JW (1995) RFLP mapping of five major genes and eight quantitative trait loci controlling flowering time in a winter × spring barley (*Hordeum vulgare* L.) cross. Genome 38:575–585

Laurie DA (1997) Comparative genetics of flowering time. Plant Mol Biol 35:167–177

Liu J, He Y, Amasino R, Chen X (2004) siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in Arabidopsis. Genes Dev 18:2873–2878

Melchinger AE, Graner A, Singh M, Messmer M (1994) Relationships among European barley germplasm: I. genetic diversity among winter and spring cultivars revealed by RFLPs. Crop Sci 34:1191–1199

Schmitz J, Franzen R, Ngyuen TH, Garcia-Maroto F, Pozzi C, Salamini F, Rohde W (2000) Cloning, mapping and expression analysis of barley MADS-box genes. Plant Mol Biol 42:899–913

Szűcs P, Karsai I, von Zitzewitz J, Mészáros K, Cooper LLD, Gu YQ, Chen THH, Hayes PM, Skinner JS (2006) Positional relationships between photoperiod response QTL and photoreceptor and vernalization genes in barley. Theor Appl Genet 112:1277–1285

Szűcs P, Skinner JS, Karsai I, Cuesta-Marcos A, Haggard KG, Corey AE, Chen THH, Hayes PM (2007) Validation of the *VRN-H2/VRN-H1* epistatic model in barley reveals that intron length variation in *VRN-H1* may account for a continuum of vernalization sensitivity. Mol Genet Genomics 277:249–261

Takahashi R, Yasuda S (1971) Genetics of earliness and growth habit in barley. In: Nilan RA (ed) Barley genetics II (proceedings of the second international barley genetics symposium). Washington State University Press, Pullman, pp 388–408

Tottman DR, Makepeace RJ (1979) An explanation of the decimal code for the growth stages of cereals, with illustrations. Annals Appl Biol 93:221–234

Trevaskis B, Bagnall DJ, Ellis MH, Peacock J, Dennis EJ (2003) MADS box genes control vernalization-induced flowering in cereals. Proc Natl Acad Sci USA 100:13099–13104

Trevaskis B, Hemming MN, Peacock WJ, Dennis ES (2006) *HvVRN2* responds to daylength, whereas *HvVRN1* is regulated by vernalization and developmental status. Plant Physiol 140:1397–1405

Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J (2003) Positional cloning of the wheat vernalization gene *VRN1*. Proc Natl Acad Sci USA 100:6263–6268

Yan LL, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, SanMiguel P, Bennetzen JL, Echenique V, Dubcovsky J (2004) The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. Science 303:1640–1644

Yan L, von Zitzewitz J, Skinner JS, Hayes PM, Dubcovsky J (2005) Molecular characterization of the duplicated meristem identity genes *HvAP1a* and *HvAP1b* in barley. Genome 48:905–912

Yasuda S, Hayashi J, Moriya I (1993) Genetic constitution for spring growth habit and some other characters in barley cultivars in the Mediterranean coastal regions. Euphytica 70:77–83

Yan L, Fu D, Li C, Blechl A, Tranquilli G, Bopnafede M, Sanchez A, Valarik M, Yasuda S, Dubcovsky J (2006) The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*. Proc Natl Acad Sci USA 103:19581–19586

von Zitzewitz J, Szűcs P, Dubcovsky J, Yan L, Francia E, Pecchioni N, Casas A, Chen THH, Hayes P, Skinner J (2005) Molecular and structural characterization of barley vernalization genes. Plant Mol Biol 59:449–467