*The Plant Journal* (2006) **47**, 811–826 doi: 10.1111/j.1365-313X.2006.02831.x

TECHNICAL ADVANCE

# Mapped *Ds*/T-DNA launch pads for functional genomics in barley

**Tiehan Zhao[1], Margaret Palotta[2], Peter Langridge[2], Manoj Prasad[3], Andreas Graner[3], Paul Schulze-Lefert[1,*] and Thomas Koprek[1,*]**

[1]*Department of Plant–Microbe Interactions, Max Planck Institute for Plant Breeding Research, Carl-von-Linné Weg 10, 50829 Cologne, Germany,*

[2]*Australian Centre for Plant Functional Genomics, Waite Campus, University of Adelaide, South Australia 5064, Australia, and*

[3]*Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstrasse 3, 06466 Gatersleben, Germany*

Received 11 January 2006; revised 25 April 2006; accepted 22 May 2006.

*For correspondence (fax +49 221 5062 353; e-mail schlef@mpiz-koeln.mpg.de and koprek@mpiz-koeln.mpg.de).

**Summary**

**A system for targeted gene tagging and local saturation mutagenesis based on maize transposable elements (*Ac*/*Ds*) was developed in barley (*Hordeum vulgare* L.). We generated large numbers of transgenic barley lines carrying a single copy of the non-autonomous maize *Ds* element at defined positions in the genome. Independent *Ds* lines were either generated by activating *Ds* elements in existing single-copy lines after crossing with AcTPase-expressing plants or by *Agrobacterium*-mediated transformation. Genomic DNA flanking *Ds* and T-DNA insertion sites from over 200 independent lines was isolated and sequenced, and was used for a sequence based mapping strategy in a barley reference population. More than 100 independent *Ds* insertion sites were mapped and can be used as launch pads for future targeted tagging of genes in the vicinity of the insertion sites. Sequence analysis of *Ds* and T-DNA flanking regions revealed a sevenfold preference of both mutagens for insertion into non-redundant, gene-containing regions of the barley genome. However, whilst transposed *Ds* elements preferentially inserted adjacent to regions with a high number of predicted and experimentally validated matrix attachment regions (nuclear MARs), this was not the case for T-DNA integration sites. These findings and an observed high transposition frequency from mapped launch pads demonstrate the future potential of gene tagging for functional genomics and gene discovery in barley.**

**Keywords: mapped *Ds*/T-DNA insertions, barley, functional genomics, transposon tagging.**

## Introduction

Barley is an agronomically important crop and represents a reference species for the Triticeae. Numerous advantages make barley an ideal candidate for genetic studies. Barley is a true diploid with seven cytologically distinct chromosomes. As such, recessive mutations can be more easily identified in barley than in its polyploid relatives, wheat and oats (Scholtz *et al.*, 2001). Barley chromosomes are homeologous with those of wheat (Moore *et al.*, 1995), which allows barley to serve as a reference for the polyploid members of the Triticeae. Although class II transposable elements of the *hAT* transposon superfamily have recently been discovered in barley (Muehlbauer *et al.*, 2006), and there are large numbers of retrotransposons (Kalendar *et al.*, 2000; Waugh *et al.*, 1997) comprising up to 85% of the barley genome (Bennetzen, 1998; Ramakrishna *et al.*, 2002), barley still appears ideal for implementing a gene-tagging system in the Triticeae to accelerate gene identification ('forward genetics').

The barley research community has amassed a large array of genomics resources (for review see Hayes *et al.*, 2003), including genetic stocks, linkage maps, quantitative trait loci (QTL) databases, expressed sequence tags (ESTs), single nucleotide polymorphisms (Kota *et al.*, 2003), the Affymetrix barley microarray (Wanamaker and Close, 2003: HarvEST: Triticeae. Release 0.99, http://138.23.191.152:blast/blast.html) and a large-insert bacterial artificial chromosome (BAC) library (Yu *et al.*, 2000). While barley genome sequencing has been precluded to date because of its

large genome size (5300 Mbp; Arumuganathan and Earle, 1991; Bennett and Smith, 1976), the development of methods such as methylation filtration (Rabinowicz *et al.*, 1999) led to a 20-fold gene enrichment with a gene space of about 250 Mbp (R. Martienssen, pers. comm.), permitting sequencing of the gene-containing regions. Further development of the existing gene-tagging system in barley (Cooper *et al.*, 2004) would complement the current genomics tools in barley and greatly facilitate functional gene analysis. Maize *Activator* (*Ac*) and *Dissociation* (*Ds*) transposable elements have been utilized as mutagens in many heterologous systems. Two-element (*Ac/Ds*) systems have been developed in order to better control the transposition of the transposon. In a two-element system, the non-autonomous *Ds* element may be activated by crossing with a stabilized autonomous element containing the transposase gene (*Ac*TPase). After its transposition, the *Ds* element may be isolated and stabilized by segregation of the elements in the following generation (Koprek *et al.*, 2000). The ability of *Ds* to transpose to linked sites (Parinov *et al.*, 1999; Ito *et al.*, 1999) allows for targeted mutagenesis of genetically linked loci if the position of a *Ds* element is known. *Ds* transpositions that land in genes of interest can be utilized directly, facilitating the functional analysis and cloning of those genes. Mapped *Ds* elements are expected to be a particularly useful resource in barley, as previous mutagenesis and mapping experiments revealed map positions of several hundred loci controlling a diverse range of traits including agronomically useful characteristics (http://www.untamo.net/cgi-bin/ace/searches/basic).

Initial steps towards gene tagging in barley were the establishment of efficient transformation systems in different barley cultivars (Jähne *et al.*, 1994; Wan and Lemaux, 1994), the demonstration of *Ds* excision in barley tissue in transient experiments (McElroy *et al.*, 1997), and the stable introduction of *Ac* and *Ds* into barley (Koprek *et al.*, 2000; Scholtz *et al.*, 2001). Koprek *et al.* (2000, 2001) have provided evidence of high transposition frequencies of *Ds* elements and characterized the transpositional behavior of *Ds* in barley. Similarly, Scholtz *et al.* (2001) also demonstrated successful transposition of *Ac* in barley, aiming at the development of a one-element system.

The establishment of *Agrobacterium*-mediated barley transformation systems (Tingay *et al.*, 1997; Wang *et al.*, 1998) resulted in very high transformation frequencies with single-copy transgenes. Advanced transformation methods are critical for the generation of large numbers of independent T-DNA lines serving as launch sites for *Ds* elements in targeted tagging approaches. The frequently observed tendency of T-DNA to insert into gene-containing regions (Barakat *et al.*, 2000; Sha *et al.*, 2004; An *et al.*, 2003; Pan *et al.*, 2005; Szabados *et al.*, 2002) adds the extra bonus of having the launch sites in positions that promise successful gene tagging.

Our long-term objective is to develop and make available to the research community a comprehensive set of *Ds* insertion lines in barley containing a single *Ds* element assigned to a general barley linkage map for targeted gene tagging. A total of 101 insertion sites have been mapped and characterized. In order to assess the prospects of our tagging system for functional genomics in barley and to reveal possible shortcomings, we studied more than 200 *Ds* and T-DNA insertion sites. Whilst both *Ds* and T-DNA mutagens showed a remarkable preference for integration in genic regions of the barley genome, DNA pattern analysis of flanking sequences revealed seemingly differential integration site selection. This was experimentally supported by *in vitro* binding assays of DNA fragments from *Ds* integration sites to the nuclear matrix. Our data demonstrate that the *Ds* launch pads provide a solid foundation for tagging approaches in the reference species barley.

## Results

### *Generation of a collection of barley lines containing* Ds *elements and T-DNA insertions*

Barley lines containing transposed *Ds* elements were generated from originally four independent single-copy *Ds* plants. Two of these (*Ds* lines A and B) carried the *Ds* element in its original integration site after transformation, while two more lines (*Ds* lines C and D) carried a single copy of *Ds* in a new integration site after one initial transposition. In all lines, transposition of *Ds* was activated by crossing the starter lines *Ds*A-*Ds*D with AcTPase-expressing plants (Koprek *et al.*, 2000). Resulting $F_2$ progeny were analyzed for transposition events in which the new integration site had segregated away from the old *Ds* insertion site and the source of *Ac*TPase. In total, 104 of these lines were identified and used for further analysis.

Barley T-DNA insertion lines were generated using a highly efficient *Agrobacterium*-mediated transformation protocol (Wang *et al.*, 2001). We constructed two T-DNA vectors based on plasmid pWBVec8 (Tingay *et al.*, 1997) that contained two different *Ds* insertion cassettes (designated pWBV-Ds-Ubi-bar-Ds and pWBV-Ubi-Ds-bar; Figure S1). In total, 171 independent lines were generated with plasmid pWBV-Ds-Ubi-bar-Ds and 149 lines with plasmid pWBV-Ubi-Ds-bar. Of these, 81 and 50 were found to contain the respective plasmids as a single-copy insert based on genomic DNA blot analysis (data not shown).

Activation of *Ds* elements from T-DNA launch pads by crossing with plants expressing the *Ac* transposase gene resulted in high *Ds* excision frequencies of 49.2% in lines carrying the plasmid pWBV-Ds-Ubi-Bar-Ds (based on the analysis of 18 independent T-DNA integration sites; the activity of each *Ds* was tested in at least 17 $F_2$ progeny). In $F_2$ lines carrying plasmid pWBV-Ubi-Ds-Bar, excision frequen-

 

cies were even higher (67.7%; five independent integration sites examined) and exceeded on average the excision frequencies observed in the different *Ds* starter lines (20.4%; Koprek *et al.*, 2000). In all lines, at least 86% of the excised *Ds* elements re-inserted in the barley genome as confirmed by Southern blot analysis. Additional re-insertions could be obscured if the *Ds* element had transposed to an unlinked site that had segregated away from the excision site in the $F_2$ generation, making it impossible to detect excision and re-insertion in the same plant.

### *Sequence analysis of* Ds *and T-DNA flanking regions*

We adopted an effective inverse PCR (ICPR) method to retrieve *Ds* flanking regions of 101 single-copy *Ds*-containing plants for subsequent DNA sequence analysis. In total, we obtained from 97 plants genomic sequences between 300 and 850 bp in length (average length 550 bp), which were used for subsequent computational analysis. Of these sequences, 78 were isolated from the *Ds* 3′ end and 19 from the *Ds* 5′ end. Isolated fragments of four additional *Ds*-containing plants were less than 100 bp in length and were therefore omitted from further analysis. Interestingly, BLASTN search results showed that 33% of the *Ds* insertions had a significant sequence homology to monocot ESTs, with an E value threshold of $<10^{-5}$ (Table S1).

We used a modified adaptor-ligation PCR (ALPCR) method to isolate T-DNA flanking sequences from 132 T-DNA lines. This enabled us to isolate longer genomic fragments with an average size of 830 bp (ranging between 200 bp and 2.0 kb). We recovered T-DNA flanking sequences from both the left (LB) and the right border (RB) of the T-DNA from 21 insertion lines, whilst from the other lines, only LB (66 lines) or RB flanking sequences (35 lines) were isolated. BLASTN searches using the same *e* value threshold of $<10^{-5}$ as for the *Ds* elements showed that 40% of the T-DNA insertion sites had significant homology to monocot EST database entries. Lines from which LB and RB sequence information was available showed consistent BLASTN search results for both flanking regions in six cases, while the other lines had no homology to cereal ESTs for one or both flanking regions. Inconsistencies between LB and RB were not discovered in cases where BLASTN hits were found on both sites of the insert. All sequences have been deposited in Genbank and the accession numbers are given in Table S1.

### *Both* Ds *and T-DNA preferentially integrate into genomic regions of low redundancy*

The results obtained by sequence analyses of integration sites and by DNA hybridization experiments using isolated fragments from these sites demonstrated a clear preference of the *Ds* element and T-DNA for integration in genomic regions of low redundancy (Figure 1a). Of all *Ds* elements analyzed, 41% were found to be re-inserted in single-copy genomic regions after transposition, whereas 31% were re-inserted in regions of low redundancy of less than five copies (estimated value based on hybridization results). Only 28% of the transposed *Ds* elements had re-inserted in regions of moderate to high repetitiveness based on BLASTN search results, revealing that these elements had either re-inserted in a gene of a large gene family (e.g. *Ds*A5 in a receptor-like kinase gene) or in retrotransposons (*Ds*A24 in a BARE-1 retrotransposon). Although the latter mobile elements comprise more than 70% of the barley genome (Vicient *et al.*, 2001), these served as integration sites in only 23% of the *Ds* lines and 29% of the T-DNA lines (Figure 1b). Insertions in other repetitive elements were rare (2% and 4%, respectively; Figure 1b). Similar overall results were found for the primary T-DNA insertion sites (36% of the insertion sites were found in single-copy regions, 27% in areas of low redundancy, and 36% in repetitive regions; Figure 1a). We conclude that both *Ds* and T-DNA preferentially insert into barley genomic regions of low redundancy.

In-depth sequence analysis of genomic DNA directly adjacent to the *Ds* and T-DNA inserts showed that 33% of the transposed *Ds* elements and 41% of the T-DNA had inserted in coding regions (Figure 1b). Such unambiguous assignments to coding regions failed for a relatively high percentage of sequenced integration sites (23% of the *Ds* insertions and 20% of the T-DNA insertions, excluding insertions in retrotransposons). This could be a consequence of the limited genomic sequence information that is available for *H. vulgare*, which complicates the identification of insertions in non-coding 5′ and 3′ gene regulatory regions as well as in introns. In 24 lines, genomic sequences immediately adjacent to the insertion site did not show recognizable relatedness to expressed genes but significant homology to ESTs was detected further upstream or downstream. In those cases, we inferred that the insertion sites were within an intron or presumptive 5′ or 3′ regulatory region of a gene. The low number of such events restricts conclusive statements about preferred insertions in those regulatory regions.

### *Analysis of the DNA structure around insertion sites*

In order to detect DNA motifs or other structural characteristics that might facilitate *Ds* and T-DNA insertion, we conducted a computational analysis of the flanking genomic regions. Genomic regions adjacent to 83 *Ds* integration sites have been further studied. A large proportion (60%) of the flanking sequences contained an unusually high AT content, ranging between 65% and 70% compared to an overall AT content in barley of about 54–57% (Melzer and Kleinhofs, 1987). The disproportionate share of AT content in *Ds* flanking regions appeared to be mainly because of distinct AT-rich motifs such as A-boxes, T-boxes and DNA unwinding motifs. These were found in 33 *Ds* flanking

 

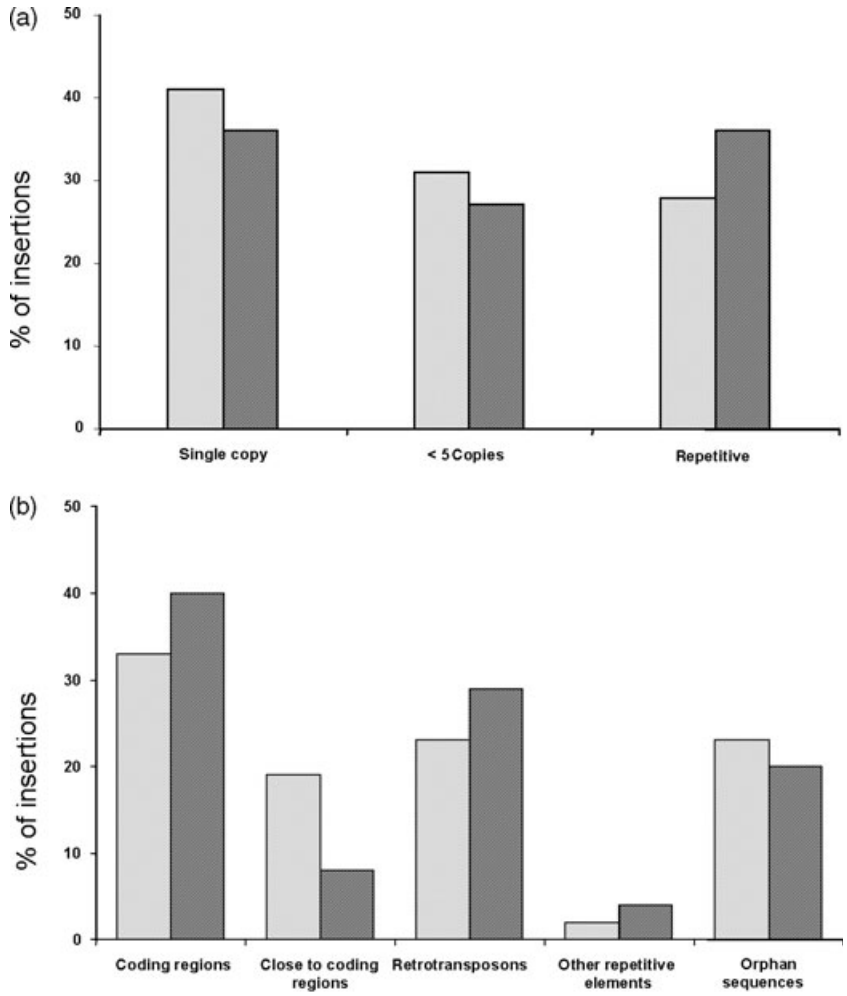


**Figure 1.** Classification of *Ds* and T-DNA insertion sites.
(a) Classification by copy number. Copy number was estimated by hybridization of flanking regions to genomic DNA, sequence analysis and mapping results.
(b) Distribution of *Ds* and T-DNA insertions within genome regions based on BLASTN searches of EST and contiguous barley BAC sequences (2.23 Mb). Light-gray columns represent *Ds* insertions and dark-gray columns represent T-DNA insertions. 'Close to coding region' means that the coding region is not immediately adjacent to the insert.

sequences. Such motifs are frequently found in nuclear matrix attachment regions (MARs), and prompted us to analyze the isolated sequences for other typical matrix binding motifs using the software program MAR-WIZ (Singh *et al.*, 1997). Interestingly, 25 out of 33 sequences contained multiple copies of up to six different typical MAR motifs. In addition to the above mentioned motifs, we found Ori-patterns, topoisomerase II binding sites, and kinked and curved DNA structures. The frequency of some of these motifs (kinked DNA, curved DNA and AT-richness patterns) was significantly higher in the *Ds* flanking sequences than in 2.23 Mb available contiguous barley genomic DNA derived from BAC clones (Figure 2). DNA motifs of curved and kinked DNA structures in particular were over-represented in the *Ds* flanking sequences. Because of the high number of these motifs, the MAR-WIZ program predicted high matrix binding strengths between 0.70 and 0.96 for 25 *Ds* flanking regions (scale between 0.00 and 1.00; default threshold level of 0.60).

Similar analyses were performed with T-DNA flanking sequences derived from 122 integration sites. Unlike *Ds* flanking regions, T-DNA flanking sequences contained in general neither unusually high numbers of MAR or other motifs, and the predicted MAR binding strength was only above the threshold level of 0.60 for 12 sequences. This is not significantly different from the frequency of MAR motifs in available contiguous genomic barley DNA sequences. However, the number of predicted kinked DNA structures was higher in T-DNA flanking sequences than in barley BAC DNA (Figure 2). In order to detect less evident differences between the two mutagens and to obtain a reference value, the number of motifs has been calculated per 500 base pairs (Figure 2b). Apart from these DNA motif differences at T-DNA and *Ds* integration sites, the average AT content of T-DNA flanking regions (58%) was lower than the AT content in *Ds* flanking regions and more similar to the average AT content in the barley genome.

### In vitro *binding assay to corroborate preferential* Ds *insertion in regions flanking MAR motif clusters*

*In vitro* matrix binding studies were carried out to test the significance of computationally predicted MARs using iso-

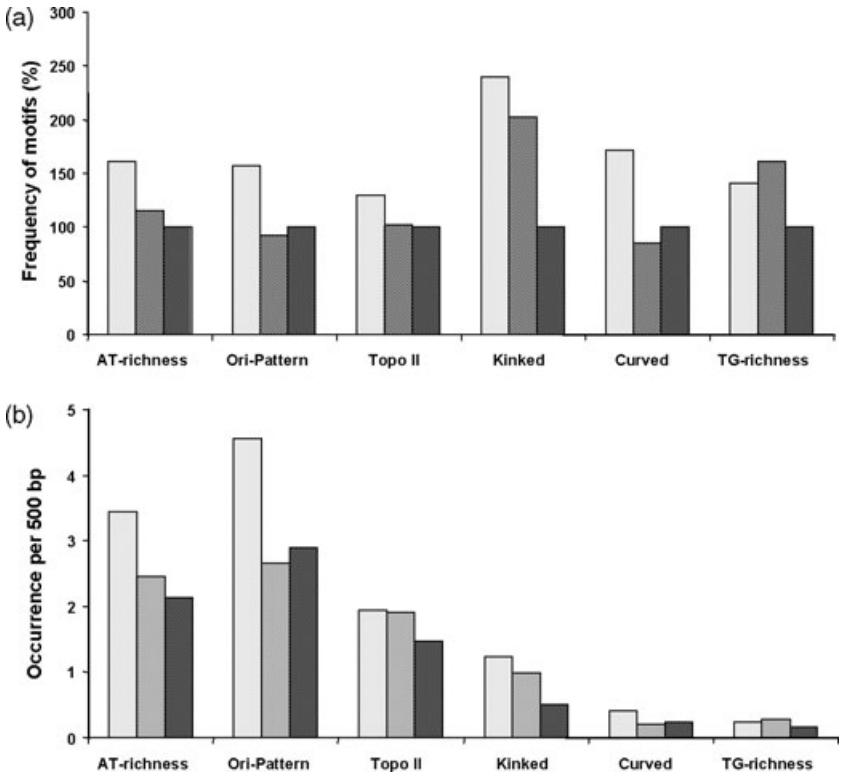


**Figure 2.** Frequency of DNA motifs in *Ds* and T-DNA flanking regions.
(a) Frequency of motifs in *Ds* and T-DNA flanking regions compared with their incidence in genomic barley DNA (the latter has been set as 100%). Genomic sequence information is derived from 2.23 Mb BAC sequences.
(b) Incidence of sequence motifs within 500 bp of flanking DNA. Light-gray columns represent *Ds* insertions, dark-gray columns represent T-DNA insertions, and genomic sequences are shown as black columns.

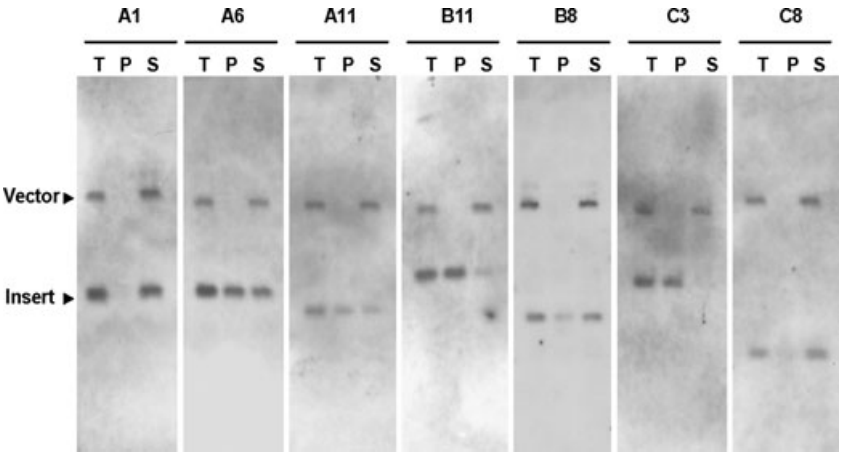


**Figure 3.** Nuclear matrix binding assay for predicted matrix attachment regions flanking *Ds* insertions.
Lanes marked 'T' represent aliquots of the total radiolabeled probe used in the assay. Lanes marked 'P' represent aliquots of the probe found in the pellet, i.e. bound to the nuclear matrix. Lanes marked 'S' represent aliquots of the probe found in the supernatant, i.e. not bound to the nuclear matrix. A1, A6, A11, B11, B8, C3 and C8 designate DNA fragments flanking *Ds* insertions in the respective plant lines.

lated nuclear matrixes derived from barley leaf tissue (Nomura *et al.*, 1997). In total, 36 *Ds* flanking sequences were used for the *in vitro* binding studies, including 25 sequences that contained five or more different MAR motifs. Three sequences had a nuclear matrix binding affinity between 0.10 and 0.60 and served as negative controls. The other sequences contained MAR motifs of only one or two out of eight different motifs, which were evenly distributed across the isolated fragments and had a predicted MAR binding strength of just above 0.60. Typical results for the nuclear matrix binding assays are shown in Figure 3. For 22 *Ds* flanking sequences, significant binding to the nuclear matrix preparations was detected (Table 1). Fragments without detectable binding affinity had a predicted binding strength below or just above the 0.60 threshold. Only three sequences with relatively high predicted matrix binding affinity (0.74, 0.76 and 0.83) did not bind to the matrix. These fragments contain typical MAR motifs that are evenly distributed over the whole sequence of the fragment, whereas validated matrix binding fragments often contain multiple MAR motifs clustered within a short stretch of DNA. Collectively, the experimentally determined 'binding activity' of

 

**Table 1** Predicted and measured matrix attachment regions binding strength of *Ds* flanking genomic regions

| Line | Predicted binding strength[a] | Measured binding strength[b] (%) |
|---|---|---|
| A1 | – | – |
| A2 | 0.86 | 69.5 |
| A4 | 0.87 | 80.3 |
| A5 | 0.89 | 84.7 |
| A6 | 0.96 | 87.5 |
| A7 | 0.84 | 69.9 |
| A9 | 0.74 | – |
| A10 | 0.78 | 33.0 |
| A11 | 0.81 | 45.1 |
| A12 | 0.84 | 66.3 |
| A13 | 0.91 | 88.5 |
| A15 | 0.82 | 91.8 |
| A17 | 0.83 | – |
| A20 | 0.66 | – |
| A21 | 0.70 | 55.9 |
| B1 | 0.72 | 18.1 |
| B2 | 0.84 | 70.9 |
| B3 | 0.82 | 60.0 |
| B4 | 0.84 | 83.3 |
| B5 | 0.86 | 54.5 |
| B8 | 0.61 | 16.9 |
| B9 | 0.78 | 59.2 |
| B11 | 0.94 | 95.1 |
| B12 | 0.64 | 32.2 |
| B15 | 0.76 | 45.3 |
| C3 | 0.94 | 89.0 |
| C4 | 0.86 | 86.5 |
| C6 | 0.66 | – |
| C8 | 0.73 | 8.8 |
| D3 | 0.78 | 31.5 |
| D5 | 0.84 | 84.2 |
| D8 | 0.62 | – |
| F2 | 0.68 | – |
| F6 | 0.76 | – |

[a]Binding strength predicted by MAR-WIZ program.
[b]*In vitro* binding strength as a percentage of total labeled input DNA measured using IMAGEQUANT software.

*Ds* flanking regions to barley nuclear matrix preparations corresponds well with computationally predicted motifs for MARs.

### *Nucleotide composition at the insertion sites*

To reveal additional characteristic features at the insertion sites based on the nucleotide composition, we plotted the nucleotide frequency versus the position of the nucleotide relative to the insertion site (Figure 4). The increased AT content and specific MAR motifs near *Ds* insertions sites did not appear to be prominent at specific positions next to the insertion sites, and the overall AT content was approximately the same in all positions. Likewise, the nucleotide composition next to the T-DNA LB and RB seems to lack specific characteristics. However, the two nucleotide positions immediately adjacent to the T-DNA RB are clearly biased. In the first position next to the T-DNA RB sequence, 47% of the sequences contain an adenine residue; the average A content over 100 bp upstream and 100 bp downstream is 29%. Likewise, in 40% of all sequences, a guanidine residue was found in the second position flanking the RB (average G content is 21%). The high number of A and G residues in these positions may result from the mechanism of T-DNA insertion or may indicate a preferred target site (Brunaud *et al.*, 2002).

### *Mapping of* Ds *and T-DNA insertion sites*

Using two different mapping procedures, 225 *Ds* and T-DNA flanking sequences (98 from *Ds* lines, 122 from T-DNA lines and five transposed *Ds* elements derived from T-DNA lines) have been analyzed for polymorphisms in different mapping populations. Most of the *Ds* elements were genetically mapped by RFLP analysis using isolated flanking regions as hybridization probes for the detection of polymorphisms in seven different barley mapping populations. Flanking regions were isolated from either the 3′ or 5′ end of each *Ds* element. Sixty-five *Ds* insertions sites were analyzed, of which 27 were derived from starter line *Ds*A, 29 from starter line *Ds*B, 21 from *Ds*C and 17 from *Ds*D. In addition, the four initial single-copy start positions were mapped. Thirty of these lines produced single hybridization bands, and four fragments produced two or three analyzable bands (*Ds*A, *Ds*D, *Ds*B9, *Ds*B10) that mapped to the same locus. Three flanking regions produced two bands that mapped to two different positions.

Because the linkage map positions were determined in different mapping populations, we integrated these based on the presence of common markers in the general barley bin map (Kleinhofs and Graner, 2001) (Figure 5a,b). The insertion site of starter line *Ds*A was found on chromosome 6H, and five transposed *Ds* elements derived from this line mapped to the same position or were very close within the same bin. However, 11 other transposition events derived from the same line mapped to other chromosomes. Transposed *Ds* elements derived from the starter line *Ds*B had also re-inserted to linked and unlinked sites (4 and 10, respectively; Figure 5a). All but one re-insertion site of transposed *Ds* elements derived from starter lines *Ds*C and *Ds*D were unlinked to the original insertion sites. This might indicate starter line-specific quantitative differences in transposing to linked and unlinked sites.

For genetic mapping of T-DNA insertions and some of the transposed *Ds* elements, we adopted a restriction digest-based assay of PCR products (CAPS assay). PCR amplicons were separated on high-resolution SSCP gels to detect even the smallest differences in fragment lengths. Polymor-

 

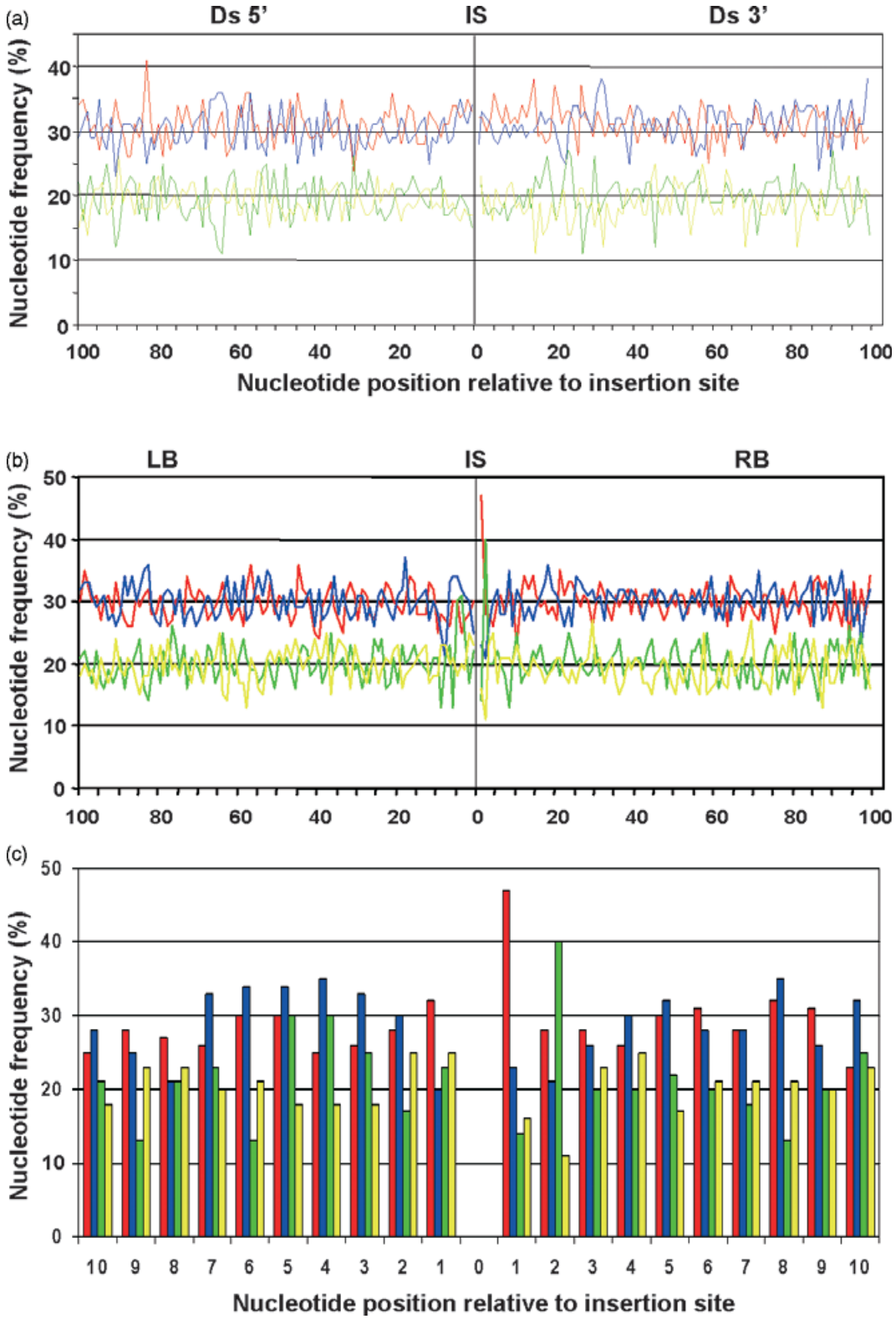


**Figure 4.** Relative base composition of the *H. vulgare* DNA at *Ds* and T-DNA insertion sites.
(a) Nucleotide composition of the 5′ and 3′ sides of *Ds* insertion sites.
(b) Nucleotide composition of left border (LB) and the right border (RB) T-DNA flanking sequences.
(c) High-resolution nucleotide composition 10 bp upstream and downstream of the T-DNA insertion site.

 

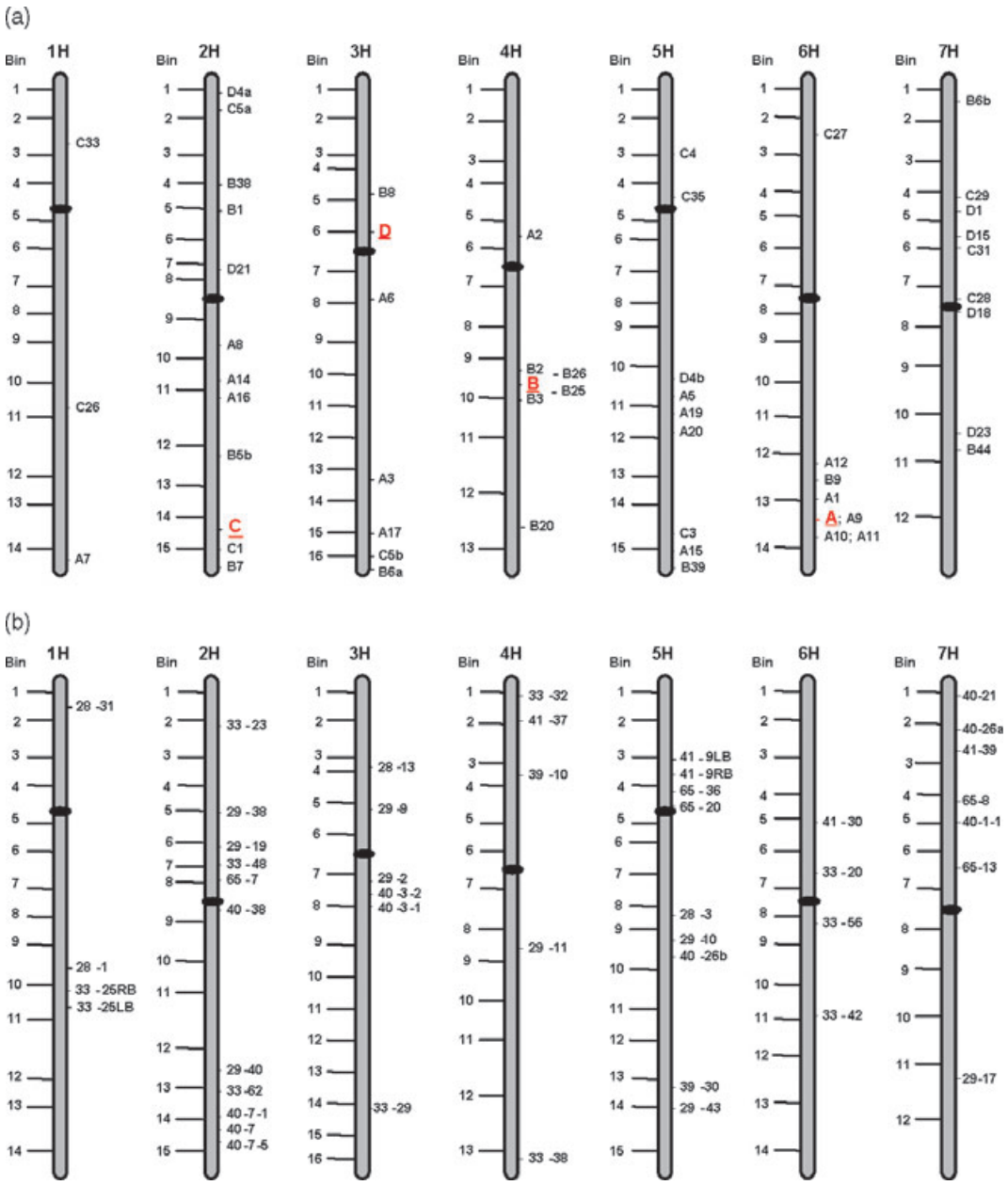


**Figure 5.** Positions of *Ds* and T-DNA insertions in the barley bin map.
Because the linkage map positions were determined in different mapping populations, map positions were assigned based on the presence of markers common to the bin map. The populations used for mapping of different insertions sites were as described in Experimental procedures. Centromeres of chromosomes are shown as black ovals.
(a) Bin map positions of *Ds* insertions. Positions of *Ds* in four starter lines are underlined and in gray.
(b) Bin map positions of T-DNA insertions. For validation of the mapping procedure, LB and RB flanking sequences were mapped from lines 33-25 and 41-9. These lines appear twice on the bin map.

phisms were detected in 54 T-DNA and 29 *Ds* flanking regions, of which in total 46 T-DNA and 17 *Ds* insertion sites could be mapped unambiguously in the Oregon Wolfe barley (OWB) mapping population and were integrated into the general barley bin map (Figure 5a,b). Other fragments were either present on multiple barley chromosomes or were homomorphic.

## Discussion

### *Generation of single-copy* Ds *lines*

We have described here two complementary approaches to generate single-copy *Ds*-containing barley plants. In the first approach, *Ds* elements in existing single-copy plants were

 

reactivated by crossing with AcTPase-expressing plants. Resulting $F_2$ plants carrying a transposed *Ds* element were then used for further analysis. This approach allows facile generation of numerous independent *Ds* lines that can serve as starter lines. Our mapping results show that the *Ds* re-insertion sites are well distributed on all seven barley chromosomes and are mostly unlinked to the original integration site. This seeming discrepancy from earlier findings in other species (Chen *et al.*, 1987; Chomet *et al.*, 1987; Parinov *et al.*, 1999; Ito *et al.*, 1999) is probably because of the relatively small number of transposition events analyzed from only four starter lines. In addition, in two out of the four *Ds* lines examined (*DsA* and *DsB*), the transposition frequencies to linked sites were significantly higher than expected from a random distribution (31% and 29%, respectively). In addition, our selection process of transposition events by Southern blot analysis may have eliminated very close transposition events (within a few kb).

The introduction of T-DNA carrying *Ds* elements via *Agrobacterium*-mediated transformation has the advantage of mostly random integration in all chromosomes. Even though there are preferences for integration in specific chromosomal areas, it appears to be more likely that an even distribution of T-DNA insertions as launch pads for further transposition will be obtained than with *Ds* elements generated from just a few original starter lines. In this study, we have shown that both approaches have the inherent advantage that *Ds* elements as well as T-DNA insert preferentially in gene-containing regions (An *et al.*, 2003; Barakat *et al.*, 2000; Chen *et al.*, 1987; Chomet *et al.*, 1987; Enoki *et al.*, 1999; Greco *et al.*, 2001; Jeon *et al.*, 2000; Kolesnik *et al.*, 2004; Sha *et al.*, 2004). Considering the expected bias of *Ds* for transpositions to linked sites, these locations promise to be efficient starting points for gene tagging after activation of *Ds*. Furthermore, gene-containing areas are in general not permanently methylated, and reactivation from these positions is therefore more efficient than from launch pads generated by direct gene transfer methods, which frequently result in integrations in methylated regions (Kohli *et al.*, 1999; reviewed in Greco *et al.*, 2003; Martienssen and Colot, 2001). These considerations are corroborated by results obtained from our activation experiments for *Ds* elements from T-DNA launch pads, showing high transposition frequencies (average of 51.6% from T-DNA launch pads versus 20.4% from *Ds* starter lines generated by direct DNA transfer). The observed, and in some cases significant, differences in *Ds* excision frequencies in different *Ds* starter lines may be a consequence of different chromatin structures as well as differences in the methylation status at the integration site (reviewed in Okamoto and Hirochika, 2001). Interestingly, *Ds* transposition frequencies reported from T-DNA launch pads in rice and in this study in barley are similar (50–74% and 51.6%, respectively; Chin *et al.*, 1999; Greco *et al.*, 2001; Kolesnik *et al.*, 2004; Nakagawa *et al.*, 2000), indicating that genome complexity does not affect *Ds* activity as long as the start sites are located in genic regions. The observed high transposition frequencies in *Ds* lines and especially of T-DNA lines are a prerequisite for a targeted gene-tagging strategy. Therefore, the insertion lines presented here are a potentially powerful tool for functional genomics in barley.

### *Preferential insertion of* Ds *and T-DNA into single-copy regions of the barley genome*

For both *Ds* and T-DNA, we observed a striking preference for integration in single-copy genomic regions or regions of low redundancy. Considering the abundance of repetitive DNA, which accounts for about 90% of the barley genome (Schulman and Kalendar, 2005) and consists mostly of transposable elements of various classes (Schulman and Kalendar, 2005), the observed insertions of either mutagen in non-repetitive regions are approximately sevenfold over-represented (Figure 1a). This coincides with a preferential integration into gene-containing regions (accounting for about 10% of the barley genome; Sandhu and Gill, 2002). The successful application of the methylation filtration method (Rabinowicz *et al.*, 1999) in maize (Palmer *et al.*, 2003) and sorghum (Bedell *et al.*, 2005) enabled the enrichment of gene-containing sequences with a gene space of only 425 Mbp in maize and 245 Mbp in sorghum. In barley, methylation filtration resulted in a 20-fold gene enrichment, and suggests a gene space of only about 250 Mbp (R. Martienssen, pers. comm.) compared with the whole genome of about 5300 Mbp (i.e. 4.7% of the genome). Taking this into account, the observed percentage of *Ds* and T-DNA integrations into gene-containing regions is even more remarkable, thus indicating the feasibility of saturation mutagenesis in plants with large genomes such as barley.

Why do both mutagens preferentially insert in gene-rich regions? We cannot exclude the possibility that the selection process after T-DNA transformation might favor integration events into open chromatin structures that are non-methylated and therefore allow for stable expression of the selectable marker gene. Consequently, transgenic plants carrying the T-DNA insert in or near genic regions will be predominantly recovered. This would explain the high percentage of T-DNA insertions in coding regions. However, this cannot explain the observed biased insertion of transposed *Ds* elements into genic regions, because these were mostly re-insertion events in progeny of starter lines without marker selection. This rather suggests that integration site selectivity is an inherent *Ds* feature or a consequence of the chromatin structure of barley gene islands (e.g. open and non-methylated chromatin structure). A bias of insertion of *Ds* and T-DNA into hypomethylated and gene-rich regions has indeed been observed in various plants species including rice, barley and Arabidopsis, and suggests accessibility

 

of DNA as one determinant of the integration process (Barakat *et al.*, 2000; Greco *et al.*, 2001; Jeon *et al.*, 2000; Koprek *et al.*, 2001; Nakagawa *et al.*, 2000; Parinov *et al.*, 1999; Yamazaki *et al.*, 2001; Sha *et al.*, 2004; An *et al.*, 2003; Greco *et al.*, 2003; Kuromori *et al.*, 2004; Pan *et al.*, 2005; Raina *et al.*, 2002; Szabados *et al.*, 2002).

Despite the fact that both *Ds* and T-DNA mainly integrate into gene-rich regions, our analysis of flanking genomic sequences suggests mutagen-specific preferences to insert into specific regions of genes (Figures 1b and 2). The observed higher frequency of *Ds* element re-insertion adjacent to a high density of predicted MAR motifs might be linked to the finding that this mutagen disrupted coding regions with lower frequency in comparison to T-DNA (33% and 40%, respectively; 225 integration sites analyzed), whilst integrations close to coding regions were found twice as often (19% and 8%, respectively). Thus, the effectiveness of *Ds* elements in producing null mutations of single genes in barley might be somewhat lower than using T-DNA integrations. Considering the high efficiency of *Agrobacterium*-mediated transformation of barley and the high frequency of T-DNA insertion in coding regions, the establishment of a system for T-DNA-based saturation mutagenesis in barley appears to be feasible and may serve as an alternative to targeted gene tagging. Our findings contrast with those reported for Arabidopsis. While in this unusually small plant genome *Ds* elements insert more often into coding regions (Pan *et al.*, 2005), T-DNA was preferentially found to insert in the 5′- and 3′-regulatory regions of genes (Pan *et al.*, 2005; Szabados *et al.*, 2002). Similar differences in insertion site preference have also been observed in rice for T-DNA and the retrotransposon *Tos*17 (Sallaud *et al.*, 2004). In fact, T-DNA insertions in regulatory regions are two to three times more frequent than those in coding regions (Szabados *et al.*, 2002). In contrast, analysis of the insertion positions within genes demonstrated that *Ds* transposed equally into exons and introns, and there is no preference toward the 5′ end of the gene as described in Arabidopsis (Parinov *et al.*, 1999). A similar *Ds* distribution has been reported in rice (Kuromori *et al.*, 2004). Moreover, in maize, the majority of *Ac* insertions occurred at the middle of the gene (75% in exons), and no bias was observed near the ATG start codon (Cowperthwaite *et al.*, 2002). Whether this seemingly unusual behavior of *Ds* in barley is a consequence of its genome composition (e.g. frequency and density of MARs, relative abundance and classes of retrotransposons) remains to be validated once more contiguous barley genomic sequences can be directly assigned to *Ds* integration sites.

### *Insertion site composition*

In-depth computational analysis of *Ds* and T-DNA flanking regions failed to unveil an insertion site preference conforming to specific consensus sequences as has been described for some transposable elements (reviewed by Craig, 1997). Nonetheless, we observed in many of the *Ds* flanking sequences an unusually high AT content as a consequence of repeatedly occurring AT-rich motifs such as A-boxes, T-boxes and AT-DNA unwinding motifs. AT-rich *Ds* flanking regions had on average an AT content of 67.3%, while the AT content in T-DNA flanking regions (58.1%) was similar to the overall AT content of the barley genome (54–57%, Melzer and Kleinhofs, 1987). Because these AT-rich motifs are frequently found in MARs, and certain transposable elements show a pronounced tendency for insertion near MARs (Bruni *et al.*, 1995; Romig *et al.*, 1994; Stephanova *et al.*, 1993; Tikhonov *et al.*, 2001), we analyzed the AT-rich *Ds* flanking regions for other typical MAR motifs using the MAR-WIZ algorithm that has been developed for the prediction of MAR binding activity of DNA sequences *in silico* (Singh *et al.*, 1997). We identified other MAR motifs in *Ds* flanking sequences, such as kinked and curved DNA structures, topoisomerase II binding sites and Ori patterns, and these occurred in the AT-rich *Ds* flanking regions more often than in T-DNA flanking regions and in contiguous barley genomic DNA of sequenced BACs.

Our data using *in vitro* binding assays confirmed for most of the sequences the calculated binding capacity to isolated nuclear matrices, suggesting that the *Ds* flanking regions are functional MARs or contain functional elements of MARs. The integration of *Ds* elements in the vicinity of MARs is not unparalleled. Various reports have shown that various classes of foreign DNA have a tendency to be integrated next to MARs. Several Long Terminal Repeats (LTR) and non-LTR retrotransposons in plants and animals appear to have a strong tendency to insert into or near MARs (Bruni *et al.*, 1995; Romig *et al.*, 1994; Stephanova *et al.*, 1993; Tikhonov *et al.*, 2001). The close vicinity of transposed *Ds* elements to MARs strongly suggests that structural DNA features of the target site are significant determinants for integration rather than primary nucleotide sequences adjacent to the integration site. Despite the fact that MARs are frequently associated with gene-containing regions, we did not observe a significant correlation between MAR binding properties and hits in BLASTN search results. *Ds* flanking regions with MAR binding capacity had significant BLASTN results as often as *Ds* flanking regions without MAR binding properties. This may be a result of the relatively low number of lines analyzed, and may be different after analysis of additional insertion sites. Although 12 T-DNA lines exhibited some MAR binding property, those lines were not further analyzed in binding assays because the predicted binding strength was just above the threshold level of 0.6 and typical motifs were not more common than in genomic barley DNA.

At T-DNA insertions, we found immediately adjacent to the T-DNA RB an abrupt change in the nucleotide composition that differs significantly from the regional composition (Figure 4). Adenine and guanine residues were clearly over-represented in the first and second positions adjacent to the

 

RB sequence. These differences might reflect the previously proposed role of micro-homologies between the host DNA and T-DNA sequences in the integration process (Tinland, 1996). Large-scale analysis of T-DNA integration sites in Arabidopsis showed that sequences flanking the T-DNA borders share similarity with the complement at the end of the T-DNA (Brunaud *et al.*, 2002). Our data are consistent with these findings in that complementary nucleotides are predominantly found in the first and second position flanking the T-DNA right border, and are therefore in accordance with the suggested models of T-DNA integration by Brunaud *et al.* (2002). However, unlike in Arabidopsis (Brunaud *et al.*, 2002), we did not observe an elevated number of short AT-rich sequences that are known to exert strong bending properties of the DNA (Bolshoy *et al.*, 1991) and are thought to play a role in the integration of foreign DNA in eukaryote genomes (Liao *et al.*, 2000; Müller and Varmus, 1994).

The observed differences in the insertion site preferences both at the nucleotide and the structural level between the transposed *Ds* elements and the T-DNA insertions may be a consequence of the different size and structure of both mutagens. Despite the very similar structure of both mutagens, the T-DNA constructs contain additional right and left border sequences that are required for *Agrobacterium*-mediated transformation. This, and the different integration processes (transposition versus T-DNA integration), could have an influence on target site preferences.

### *Mapping and distribution of insertion sites*

Using an RFLP and an SSCP analysis-based mapping strategy, we determined the location of 55 independent single-copy *Ds* insertions and 46 T-DNA integration sites (Figure 5a,b). The integration of the mapped *Ds* and T-DNA integration sites into the barley bin map (Kleinhofs and Graner, 2001) appears to be useful as it will facilitate the integration of functional and structural genomics resources in barley. By assigning the *Ds* and T-DNA loci to bins on the OWB map, researchers can identify the insertions closest to mapped morphological loci, genes of known function or QTL regions identified in other cultivars. These *Ds* insertions can then serve as targets for reactivation in order to tag linked genes. In addition, use of the bin map allows integration of mapped insertion sites obtained from other mapping populations (Cooper *et al.*, 2004).

Our data suggest a relatively even distribution of the insertions in the barley genome, without significant gaps, although this remains to be re-evaluated when the density of starter lines increases further. Similarly, the relatively small number of starter lines does not allow us to draw conclusions about differences in the transpositional behaviour in different starter lines. Although the origin of the four *Ds* starter lines differed (original integration versus primary transposition of *Ds*), we did not observe significant differences concerning the transposition to linked sites or preferences for integration in genic regions.

Considering the observed high reactivation frequency of *Ds* elements after crossing with *AcTPase*-expressing plants, the system offers great potential as a tool for targeted tagging of a plethora of mapped loci with known mutant phenotypes. As existing barley TILLING populations and comparable methods cannot be utilized in such 'forward genetics' experiments (Caldwell *et al.*, 2004), and map-based cloning in barley remains both costly and time-consuming (Buschges *et al.*, 1997; Collins *et al.*, 2003; Shirasu *et al.*, 1999; Zhou *et al.*, 2001), the mapped *Ds* launch pads described here help to close a significant gap in the barley functional genomics toolbox.

## Experimental procedures

### *Plant material*

Two different populations containing the maize transposable elements *Ac* and *Ds* were generated. In one population, the transposons were introduced into barley (cv. Golden Promise) as described previously (Koprek *et al.*, 2000). Barley plants containing single-copy transposed elements were generated from originally two independent single-copy *Ds* lines (*Ds* starter lines A and B). In these two lines, which had been produced by particle bombardment, the *Ds* elements resided in their original integration site after transformation. Transposition of *Ds* was initiated by crossing the starter lines with previously described AcTPase-expressing plants (Koprek *et al.*, 2000). $F_2$ progeny resulting from these crosses were analyzed for transposition events. Two lines carrying a single-copy transposed *Ds* element were used as additional starter lines (*Ds* starter lines C and D), in which the *Ds* element was reactivated for a secondary transposition as explained above. In order to simplify the subsequent isolation of *Ds* flanking genomic regions and to make interpretation of sequence analysis unambiguous, we selected plants carrying only a new single-copy insertion site. In total, 104 lines were identified in which the *Ds* insertion site had segregated away from the original *Ds* integration site and the *Ac*TPase gene. In all of these plants, the structure of the transposed *Ds* element is identical (corresponding to the *Ds* cassette as shown in Figure S1). From these plants, as well as from the original lines *Ds*A, *Ds*B, *Ds*C and *Ds*D, *Ds* flanking DNA was isolated, sequenced, and used for mapping.

The second population consists of barley T-DNA insertion lines that were generated by *Agrobacterium*-mediated transformation with two different plasmids carrying the *Ds* element (Figure S1). The transformation was carried out was according to the method described by Wang *et al.* (2001). Integration of complete T-DNA cassettes was validated by genomic Southern analysis using two vector-derived hybridization probes. *Ds* elements were activated by crossing the T-DNA containing plants with *AcTPase*-expressing plants. In most cases, the *AcTPase*-expressing lines were the same as for the activation in the *Ds* starter lines.

All Ds-, T-DNA- and *AcTPase*-containing barley lines are available upon request from the corresponding authors.

### *Growth of donor plant material for barley transformation*

Barley (*Hordeum vulgare* L. cv. Golden Promise) was grown in growth chambers at 15–16°C/12°C day per night temperatures with a

 

16 h photoperiod (cool white 215W, Silwana, Danvers, MA, USA) and an 8 h dark period at an average relative humidity of about 50%. Plants were fertilized once per week with NPK (20–20–20).

### *Vector construction*

*pWBV-Ds-Ubi-bar-Ds.* The *Hind*III and *Cla*I fragment (blunted) of pSP-Ds-Ubi-bar-Ds (Koprek *et al.*, 2000), which contains the *Ds* element carrying a *bar* gene under the control of the maize ubiquitin 1 promoter, was ligated to the *Not*I-digested (blunted) pWBVec8+A (Wang *et al.*, 1998), resulting in construct pWBV-Ds-Ubi-bar-Ds.

*pWBV-Ubi-Ds-Bar.* The Ubi-bar cassette from pSP-Ds-ubibar-Ds was removed by digestion with *Not*I. Into the *Not*I site of the resulting pSP-Ds and into the *Not*I-digested vector backbone pWBVec8+N, we cloned additional I-*Sce*I sites. The ubiquitin promoter and first intron and the *bar* gene were released from pSP-Ds-ubi-bar-Ds, blunt-ended and ligated into the *Pst*I-digested and blunt-ended fragments of pSP-Ds, to result in pSP-Ubi-Ds-bar, which was cloned into the multiple cloning site of pWBVec8+N resulting in pWBV-Ubi-Ds-Bar.

*pWBV-AcAcPase.* The *Pst*I fragment of pUC-codA-Act-Ac-Ac, which contains *Ac*TPase under the control of a putative *Ac* promoter, was ligated to the *Pst*I site of pBlue+AN to result in pBlue-Ac-Ac. The *Asc*I- and *Not*I-digested cassette from pBlue-Ac-Ac was inserted into the *Asc*I and *Not*I restriction site of pWBV+8A to generate pWBV-Ac-Ac.

### *Isolation of* Ds *flanking genomic regions*

Genomic DNA was isolated according to the method described by Cone *et al.* (1989). *Ds* flanking regions were isolated using IPCR or adaptor-ligation PCR (ALPCR). For IPCR, 5 μg of genomic DNA were digested with *Bst*YI (10 U, Invitrogen, Karlsruhe, Germany) in a 100 μl reaction mixture at 60°C overnight, purified with nucleospin columns (Machery & Nagel, Dveren, Germany), and ligated in a total volume of 100 μl with T4 ligase (10 U, Invitrogen) at 16°C for 12–16 h. Ligation products were again column-purified and used directly for PCR in the following reaction mixture: 5 μl ligated DNA, 5 μl PCR-buffer (10 ×), 4 μl dNTPs (2.5 mM), 3 μl $MgCl_2$ (25 mM), 1 μl (10 pmol/$\mu l^{-1}$) of each primer, and 2 U Taq polymerase (Invitrogen) in a total volume of 50 μl. PCR conditions for the first PCR were: 94°C for 2 min, then 30 cycles of 94°C for 1 min, 55°C for 1 min and 72°C for 3 min, with a final extension at 72°C for 10 min. First-cycle PCR products were diluted 1:20 in sterile $ddH_2O$ and used for a second round of PCR using nested primers and the same PCR conditions. Primers for the first PCR were as follows: for the 5′ end of *Ds*: Ds5′fwd-1, 5′-ACCTCGGGTTCGAAATCGATCG and Ds5′Rev-1, 5′-CGGAAACGGGATATACAAAACGG; for the 3′ end of *Ds*: Ds3′Fwd-1, 5′-CATCCAGAAATTGCGTGGCGGA and Ds3′ Rev-1, 5′-CGACCGGATCGTATCGGTTTTCG. For the secondary PCR reaction, the following nested primers were used: for the 5′ end: Ds5′fwd-2, 5′-ATCGGTTATACGATAACGGTCGG and Ds5′rev-2, 5′-CGGAAACGGTAGAGCTAGTTTCC and for the 3′ end: Ds3′fwd-2, 5′-ATTCCTTTCCCACCGCTCCTTCGC and Ds3′ rev-2, 5′-CGATTACCGTATTTATCCCGTTCG.

For ALPCR, 5 μg genomic DNA were digested with *Dra*I, *Eco*RI or *Ssp*I overnight in a 100 μl volume at the appropriate temperature. DNA fragments were ligated with adaptor 1 and adaptor 2 using T4 ligase at 16°C overnight in a volume of 100 μl. The PCR conditions for first-round PCR were 10 μl ligated DNA (approximately 500 ng), 5 μl buffer, 6 μl dNTP (2.5 μM), 2 μl (10 μM) primer AP1, 2 μl (10 μM) Ds3R1 or Ds5F1, and 0.5 U Taq DNA polymerase in a 50 μl volume reaction. PCR cycling was performed as follows: 94°C for 2 min, then 31 cycles of 94°C for 1 min, 55°C for 1 min and 72°C for 3 min. The secondary PCR was conducted in the same way but with 1 or 2 μl of a 1:10 dilution of the first-round PCR product and with nested primers AP2 and Ds3R2 or Ds5F2 and 35 cycles. The sequences of the oligonucleotides are: adaptor 1: 5′-GTAATACGACTCACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGT-3′; adaptor 2: 3′-H2N-CCCGACCA-$PO_4$-5′; adaptor primer 1 (AP1): 5′-GTAATACGACTCACTTATAGGGC-3′; adaptor primer 2 (AP2): 5′-ACTATAGGGCACGCGTGGT-3′; Ds3 reverse primer 1 (Ds3R1): GATTTCGACTTACCCCGACCGGATCG; Ds3 reverse primer 2 (Ds3R2): GTTACCGGTATATCCCGTTTTCGTTTCCG; Ds5 forward primer 1 (Ds5F1): GTGAAACGGTCGGGAAACTAGCTCTACCG; Ds5 forward primer 2 (Ds5F2): ACGATAACGGTCGGTACGGGATTTTCC.

PCR products were electrophoretically separated on ethidium bromide-stained 1% agarose gels, fragments of more than 400 bp were excised, purified (nucleospin kit; Machery & Nagel) and directly used for sequencing. For *in vitro* binding studies, fragments were subcloned into the pGEM-T-Easy cloning vector (Promega, Mannheim, Germany).

### *Sequencing and sequence analysis*

DNA sequences were determined by the Max-Planck-Institut fuer Zuechtungs-forschung (MPIZ) DNA core facility on Applied Biosystems (Weiterstadt, Germany) Abi Prism 377, 3100 and 3730 sequencers using BigDye-terminator version 3.1 chemistry. Premixed reagents were obtained from Applied Biosystems. Oligonucleotides were purchased from Invitrogen. BLASTN and TBLASTX (Altschul *et al.*, 1990, 1997) searches were performed using genomic DNA sequences from wild-type Golden Promise corresponding to the DNA flanking each *Ds* insertion event. The following databases were utilized: NCBI Blast (http://www.ncb.nlm.nil.gov/BLAST), Gramene (http://www.gramene.org/multi/blastview), IPK barley ESTs database (http://pgrc.ipk-gatersleben.de) and TIGR gene indices (http://tigrblast.tigr.org/tgi/). All sequences have been deposited in Genbank and the accession numbers are given in Table S1.

To detect typical motifs of MAR within the sequenced *Ds* flanking regions and for prediction of matrix binding strength, the program MAR-WIZ version 1.5 (Singh *et al.*, 1997; http://www.future-soft.org.MAR-Wiz/) was used. A more detailed search for MAR consensus sequences was performed using the computer program BIOEDIT 7.0.1 (Hall, 1999). These sequences included A-box motifs (AATAAAYAAA) (Gasser and Laemmli, 1986; Hall *et al.*, 1991), T-boxes (TTWTWTTWTT) (Mirkovitch *et al.*, 1988), topoisomerase II binding sites (Cockerill and Garrard, 1986; Käs and Laemmli, 1992), and DNA strand unwinding sequences (AATATT and AATATATTT) (Bode *et al.*, 1992; Kohwi-Shigematsu *et al.*, 1997; Mielke *et al.*, 1990), as well as kinkable dinucleotides with potential for forming bends (Homberger, 1989). Motif pattern search was conducted on both sense and antisense DNA strands.

### *Isolation of barley nuclei*

Five grams of young leaves of plants carrying a single copy of a transposed Ds element were ground in liquid $N_2$; the powder was resuspended in 25 ml of Hamilton buffer (10 mM Tris–HCl, pH 7.6, 1.14 M sucrose, 5 mM $MgCl_2$ and 5 mM β-mercaptoethanol; Hamilton *et al.*, 1972). The slurry was gently stirred and filtered through three layers of Miracloth (Calbiochem, La Jolla, CA, USA) in 10 ml of

 

the same buffer containing 0.3% of Triton X-100. The suspension was kept in a centrifuge tube on ice for 15 min with occasional gentle inversion of the tube. After centrifugation for 10 min at 1000 ***g***, the pellet was washed twice more in same buffer to remove chloroplasts and mitochondria. The final pellet was resuspended and washed in 5 ml of extraction buffer without Triton X-100. The quality and number of intact nuclei were determined after staining with 4,6-diamino-phenylindole (DAPI) and inspection under a fluorescence microscope, and were diluted and stored at a density of $10^7$ nuclei $ml^{-1}$.

### *Isolation of nuclear matrices*

Nuclear matrices for the binding assay were essentially prepared according to the method described by Nomura *et al.* (1997) using a high salt concentration extraction treatment. Nuclei of $10^7$ cells were washed once in 10 ml MES buffer (50 mM MES, pH 7.0, 50 mM NaCl, 5 mM $MgCl_2$ and 1 mM dithiothreitol), centrifuged at 2000 ***g*** for 10 min at 4°C, and digested with *Eco*RI (500 U for $10^7$ nuclei) at 37°C for 2h. After centrifugation at 2000 ***g*** for 10 min, the pellet was washed with 50 mM MES, pH 7.0, 5 mM $MgCl_2$ and 1 mM PMSF. For complete removal of nucleic acids, the suspension was treated with 50 μg $ml^{-1}$ DNase I and 20 μg $ml^{-1}$ RNase A at 37°C for 1 h. An equal volume of ice-cold 4 M NaCl in 50 mM MES, pH 7.0, 5 mM $MgCl_2$ and 1 mM PMSF was added, and the samples were incubated at room temperature for 10–20 min with several inversions of the tube. Insoluble nuclear proteins containing the nuclear matrix fraction were pelleted by centrifugation (10 000 ***g***) for 10 min at 4°C. The supernatant was carefully removed and the pellet was washed twice with 10 ml of 2 M NaCl in 50 mM MES, pH 7.0, 5 mM $MgCl_2$ and 1 mM PMSF to remove all residual nucleases. The nuclear matrices were washed three times as described above in the matrix binding buffer (20 mM Tris–HCl, pH 7.4, 70 mM NaCl, 20 mM KCl, 10 mM $MgCl_2$, 0.125 mM spermidine and 0.05 mM spermine).

### *Binding assays*

Determination of the binding of the *Ds* flanking genomic fragments to isolated nuclear matrices was performed by the exogenous binding assay as previously described (Allen *et al.*, 1996). The cloned fragments were released from the plasmid by *Eco*RI digestion. Both the released insert and the plasmid backbone were [α]-$^{32}$P end-labeled with the Klenow fragment of DNA polymerase I (Promega) and included in the binding assay. The radioactively labeled vector in an equimolar ratio served as an internal control (non-MAR-binding DNA fragment) and allowed better quantification of the ratio between bound and unbound fragments. Nuclear matrixes isolated from 500 000 nuclei and approximately 20 ng of labeled probe (insert and vector) were incubated with sheared *Escherichia coli* competitor DNA (final concentration 100 μg $ml^{-1}$) in a total volume of 500 μl at 37°C overnight with gentle agitation. After separation into pellet (bound DNA fragments) and supernatant (unbound DNA fragments) fractions by centrifugation (10 000 ***g*** in a benchtop microcentrifuge for 30 min at 4°C), DNA was purified by incubation at 37°C for 60 min with 0.1% SDS and 50 μg $ml^{-1}$ proteinase K, followed by phenol–chloroform extraction. The DNA precipitate was dissolved in 30 μl TE. For all experiments we have used the 'equal fractions' method of presenting data as described by Kay and Bode (1995), i.e. an equal fraction of the total volume of each of the supernatant and resuspended pellet fractions is subjected to electrophoresis. With this approach, the amount of radioactivity in a band from the pellet or supernatant is directly proportional to the amount of labeled probe DNA in each fraction. Therefore, the ratio of matrix-bound probe to unbound probe can be directly compared from lane to lane. In the experiments presented here, we loaded 10 μl of pellet, 10 μl of supernatant and 20 μl total fraction (consisting of 10 μl of each, pellet and supernatant) on 1.0% agarose gels. After electrophoresis, the gels were dried onto nylon membranes, followed by overnight autoradiography.

### *Image analysis*

Exposed X-ray films were analyzed by IMAGEQUANT version 5.0 (Molecular Dynamics, Sunnyvale, CA, USA). The mean gray value for each fragment (vector backbone or insert) was computed, and the values for the local background were subtracted to obtain the relative band intensity of the fragment. Mean gray values for the background were determined by measuring the gray value of a same-sized area adjacent to the band. Mean gray values of the bands from the 'total' fraction representing equimolar amounts of the labeled empty vector and labeled insert DNA were used as a reference and were set as 100%. MAR binding strength was calculated as a percentage value for the matrix-bound fraction relative to the value in the 'total' fraction. Comparison between the band intensities of the vector in the total and supernatant fraction and between the vector backbone and the matrix-bound fraction in the total fraction was used to control the recovery of labeled fragments from the binding assay.

### *Genetic mapping of T-DNA and* Ds *insertion sites by PCR-SSCP*

The distribution of T-DNA insertions on barley chromosomes was determined by PCR analysis of wheat-barley addition lines (Islam *et al.*, 1981). The sequences of isolated T-DNA and *Ds* flanking regions were used to design primers for the amplification and identification of polymorphisms of the corresponding alleles in different mapping populations. PCR amplifications were carried out using conditions that are appropriate for the amplification of the expected PCR product. All PCR reagents were from Bioline (Luckenwalde, Germany). Plant lines carrying the isolated flanking region on one chromosome only were further used for mapping. PCR products were amplified from OWB parents Dom and Rec. Resulting PCR fragments were digested with various restriction enzymes (*Alu*I, *Bstu*I, *Hae*III, *Hha*I, *Hin*fI, *Mse*I, *Msp*I, *Rsa*I, *Sau*3A, *Sau*96I, *Scr*FI, *Taq*I and *Tsp*509) and analyzed for polymorphisms on high-resolution Mutation Detection Enhancement (MDE) gels (Sentinelli *et al.*, 2000). Once a polymorphism was detected, the corresponding sequence was amplified from each member of the Oregon-Wolfe-Barley Doubled Haploid (OWB DH) mapping population and digested with the appropriate restriction enzyme. Depending upon the size difference of the generated DNA fragments, individual members of the mapping population were assayed and scored for the polymorphisms on MDE gels or on 2–3% agarose gels stained with ethidium bromide.

### *Mapping of* Ds *insertion sites by RFLP*

Genomic DNA samples from 14 different genotypes that are the parents of seven mapping populations were each digested with six restriction enzymes (*Bam*HI, *Dra*I, *Eco*RI, *Eco*RV, *Hind*III, *Xba*I) and a Southern blot was performed with these samples as described by Kleinhofs *et al.* (1993). The following mapping populations were used to detect RFLP polymorphisms and for subsequent mapping: OWB-Dom × OWB-Rec, Steptoe × Morex (Kleinhofs *et al.*, 1993),

 

Igri × Franka (Graner *et al.*, 1994, Galleon × Haruna Nijo (Karakousis *et al.*, 2003a), Chebec × Harrington (Barr *et al.*, 2003), Clipper × Sahara (Karakousis *et al.*, 2003b), Sloop × Halcyon (Read *et al.*, 2003). Membranes were hybridized with $^{32}$P-labeled probes generated using genomic DNA flanking single-copy *Ds* insertion sites. The procedure used for RFLP analysis has been described by Kleinhofs *et al.* (1993). Linkage map construction was performed using the program MAP MANAGER (Manly *et al.*, 2001).

## Acknowledgements

This work was supported by the Deutsche Forschungs Gemeinschaft (DFG grant number 1791/2-1(2-2) and by the Plant Genome Programme (GABI) of the Federal Ministry of Education and Research (BMBF grant number PTJ-BIO/0312858). Unpublished data have been provided by Dr Robert Martienssen, Cold Spring Harbor Laboratory (NY, USA).

## Supplementary Material

The following supplementary material is available for this article online:

**Figure S1.** Schematic representation of plasmids used for Agrobacterium-mediated transformation. Only restriction sites which were relevant for Southern analysis are shown. The black bar indicates the size standard, red bars indicate the probes used for Southern hybridization. RB, right border; LB, left border; Nos, Nos terminator; 35ST, 35S terminator; Hpt, hygromycin phospho transferase; Ds, Ds element; UbiProm, ubiquitin promoter; Bar, phosphinothricin acetyl transferase gene.

**Table S1** Summary of *Ds* and T-DNA insertion site mapping in barley based on chromosome and bin location, cereal EST matches and E values

This material is available as part of the online article from http://www.blackwell-synergy.com

## References

**Allen, G.C., Hall, G., Michalowski, S., Newman, W., Weissinger, A.K., Spiker, S. and Thompson W.F.** (1996) High-level transgene expression in plant cells: effects of a strong scaffold attachment region from tobacco. *Plant Cell*, **8**, 899–913.

**Altschul, S., Gish, W., Miller, W., Myers, E. and Lipman, D.** (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.

**Altschul, S., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman D.J.** (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402.

**An, S., Park, S., Jeong, D.H. *et al.*** (2003) Generation and analysis of end sequence database for T-DNA tagging lines in rice. *Plant Physiol.* **133**, 2040–2047.

**Arumuganathan, K. and Earle, E.D.** (1991) Nuclear DNA content of some important plant species. *Plant Mol. Biol. Rep.* **9**, 208–218.

**Barakat, A., Gallois, P., Raynal, M., Mestre-Ortega, D., Sallaud, C., Guiderdoni, E., Delseny, M. and Bernardi, G.** (2000) The distribution of T-DNA in the genomes of transgenic *Arabidopsis* and rice. *FEBS Lett.* **471**, 161–164.

**Barr, A.R., Karakousis, A., Lance, R.C.M. *et al.*** (2003) Mapping and QTL analysis of the barley population Chebec × Harrington. *Aust. J. Agric. Res.* **54**, 1125–1130.

**Bedell, J.A., Budiman, M.A., Nunberg, A. *et al.*** (2005) Sorghum genome sequencing by methylation filtration. *PLoS Biol.* **3**, e13.

**Bennett, M.D. and Smith, J.B.** (1976) Nuclear DNA amount in angiosperms. *Philos. Trans. Roy. Soc. Lond. B Biol. Sci.* **B274**, 227–274.

**Bennetzen, J.I.** (1998) The structure and evolution of angiosperm nuclear genomes. *Curr. Opin. Plant Biol.* **2**, 103–108.

**Bode, J., Kohwi, Y., Dickinson, L., Joh, T., Klehr, D., Mielke, C. and Kohwi-Shigematsu, T.** (1992) Biological significance of unwinding capability of nuclear matrix-associating DNA. *Science*, **255**, 195–197.

**Bolshoy, A., McNamara, P., Harrington, R.E. and Trifonov, E.N.** (1991) Curved DNA without A-A: experimental estimation of all 16 DNA wedge angles. *Proc. Natl Acad. Sci. USA*, **88**, 2312–2316.

**Brunaud, V., Balzergue, S., Dubreucq, B. *et al.*** (2002) T-DNA integration into the *Arabidopsis* genome depends on sequences of pre-insertion sites. *EMBO Rep.* **3**, 1152–1157.

**Bruni, R., Argentini, C., D'Ugo, E., Ciccaglione A. and Rapicetta, M.** (1995) Recurrence of WHV integration in the b3n locus in woodchuck hepatocellular carcinoma. *Virology*, **214**, 229–234.

**Buschges, R., Hollricher, K., Panstruga, R. *et al.*** (1997) The barley Mlo gene: a novel control element of plant pathogen resistance. *Cell*, **88**, 695–705.

**Caldwell, D.G., McCallum, N., Shaw, P., Muehlbauer, G.J., Marshall, D.F. and Waugh, R.** (2004) A structured mutant population for forward and reverse genetics in barley (*Hordeum vulgare* L.). *Plant J.* **40**, 143–150.

**Chen, J., Greenblatt, I.M. and Dellaporta, S.L.** (1987) Transposition of *Ac* from the *P* locus of maize into unreplicated chromosomal sites. *Genetics*, **117**, 109–116.

**Chin, H.G., Choe, M.S., Lee, S.H. *et al.*** (1999) Molecular analysis of rice plants harboring an *Ac/Ds* transposable element-mediated gene trapping system. *Plant J.* **19**, 615–623.

**Chomet, P.S., Wessler, S. and Dellaporta, S.L.** (1987) Inactivation of the maize transposable element Activator (*Ac*) is associated with its DNA modification. *EMBO J.* **6**, 295–302.

**Cockerill, P.N. and Garrard, W.T.** (1986) Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. *Cell*, **44**, 273–282.

**Collins, N.C., Thordal-Christensen, H., Lipka, V. *et al.*** (2003) SNARE protein-mediated disease resistance at the plant cell wall. *Nature*, **425**, 973–977.

**Cone, K.** (1989) Yet another rapid plant DNA prep. *Maize Gen. Coop. Newslett.* **63**, 68.

**Cooper, L.D., Marquez-Cedillo, L., Singh, J. *et al.*** (2004) Mapping *Ds* insertions in barley using a sequence-based approach. *Mol. Gen. Genomics*, **272**, 181–193.

**Cowperthwaite, M., Park, W., Xu, Z., Yan, X., Maurais, S.C. and Dooner, H.K.** (2002) Use of the transposon *Ac* as a gene-searching engine in the maize genome. *Plant Cell*, **14**, 713–726.

**Craig, N.L.** (1997) Target site selection in transposition. *Annu. Rev. Biochem.* **66**, 437–474.

**Enoki, H., Izawa, T., Kawahara, M., Komatsu, M., Koh, S., Kyozuka, J. and Shimamoto, K.** (1999) *Ac* as a tool for the functional genomics of rice. *Plant J.* **19**, 605–613.

**Gasser, S.M. and Laemmli, U.K.** (1986) Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of *D. melanogaster*. *Cell*, **46**, 521–530.

**Graner, A., Bauer, E., Kellermann, A., Kirchner, S., Muraya, J.K., Jahoor, A. and Wenzel, G.** (1994) Progress of RFLP-map construction in winter barley. *Barley Genet. Newsl.* **23**, 53–59.

**Greco, R., Ouwerkerk, P.B., Taal, A.J., Favalli, C., Beguiristain, T., Puigdomenech, P., Colombo, L., Hoge, J.H. and Pereira, A.** (2001)

Early and multiple *Ac* transpositions in rice suitable for efficient insertional mutagenesis. *Plant Mol. Biol.* **46**, 215–227.

**Greco, R., Ouwerkerk, P.B., De Kam, R.J., Sallaud, C., Favalli, C., Colombo, L., Guiderdoni, E., Meijer, A.H., Hoge, J.H. and Pereira, A.** (2003) Transpositional behaviour of an Ac/Ds system for reverse genetics in rice. *Theor. Appl. Genet.* **108**, 10–24.

**Hall, T.A.** (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**, 95–98.

**Hall, G., Allen, G.C., Loer, D.S., Thompson, W.F. and Spiker, S.** (1991) Nuclear scaffolds and scaffold-attachment regions in higher plants. *Proc. Natl Acad. Sci. USA*, **88**, 9320–9324.

**Hamilton, R.H., Kunsch, U. and Temperli, A.** (1972) Simple rapid procedures for the isolation of tobacco leaf nuclei. *Anal. Biochem.* **49**, 48–57.

**Hayes, P.M., Castro, A., Marquez-Cedillo, L. *et al.*** (2003) Genetic diversity for quantitatively inherited agronomic and malting quality traits. In *Diversity in Barley (Hordeum vulgare*) (von Bothmer, R., Knüpffer, H., van Hintum, T. and Sato, K., eds). Amsterdam: Elsevier, pp. 201–226.

**Homberger, H.P.** (1989) Bent DNA is a structural feature of scaffold-attached regions in *Drosophila melanogaster* interphase nuclei. *Chromosoma*, **98**, 99–104.

**Islam, A.K.M.R., Shepherd, K. and Sparrow, D.H.B.** (1981) Isolation and characterisation of euplasmic wheat–barley chromosome addition lines. *Heredity*, **46**, 161–174.

**Ito, T., Seki, M., Hayashida, N., Shibata, D. and Shinozaki, K.** (1999) Regional insertional mutagenesis of genes on *Arabidopsis thaliana* chromosome V using the *Ac*/*Ds* transposon in combination with a cDNA scanning method. *Plant J.* **17**, 433–444.

**Jähne, A., Becker, D., Brettschneider, R. and Lörz, H.** (1994) Regeneration of transgenic, microspore-derived, fertile barley. *Theor. Appl. Genet.* **89**, 525–533.

**Jeon, J.S., Lee, S., Jung, K.H. *et al.*** (2000) T-DNA insertional mutagenesis for functional genomics in rice. *Plant J.* **22**, 561–570.

**Kalendar, R., Tanskanen, J., Immonen, S., Nevo, E. and Schulman, A.H.** (2000) Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence. *Proc. Natl Acad. Sci. USA*, **97**, 6603–6607.

**Karakousis, A., Barr, A.R., Kretschmer, J.M. *et al.*** (2003a) Mapping and QTL analysis of the barley population Galleon × Haruna Nijo. *Aust. J. Agric. Res.* **54**, 1131–1135.

**Karakousis, A., Barr, A.R., Kretschmer, J.M., Manning, S., Jefferies, S.P., Chalmers, K.J., Islam, A.K.M. and Langridge, P.** (2003b) Mapping and QTL analysis of the barley population Clipper × Sahara. *Aust. J. Agric. Res.* **54**, 1137–1140.

**Käs, E. and Laemmli, U.K.** (1992) In-vivo topoisomerase II cleavage of the Drosophila histone and satellite III repeats DNA sequence and structural characteristics. *EMBO J.* **11**, 705–716.

**Kay, V. and Bode, J.** (1995) Detection of scaffold-attached regions (SARs) by in vitro techniques: activities of these elements *in vivo*. *Methods Mol. Cell. Biol.* **5**, 186–194.

**Kleinhofs, A. and Graner, A.** (2001) An integrated map of the barley genome. In *DNA-Based Markers in Plants* (Phillips, R.L. and Vasil, I.K., eds). Dordrecht, The Netherlands: Kluwer Academic Publishers, pp. 187–199.

**Kleinhofs, A., Kilian, A., Saghai-Maroof, M.A., Biyashev, R.M. and Hayes, P.** (1993) A molecular, isozyme and morphological map of the barley (*Hordeum vulgare*) genome. *Theor. Appl. Genet.* **86**, 705–712.

**Kohli, A., Gahakwa, D., Vain, P., Laurie, D.A. and Christou, P.** (1999) Transgene expression in rice engineered through particle bombardment: molecular factors controlling stable expression and transgene silencing. *Planta*, **208**, 88–97.

**Kohwi-Shigematsu, T., Maass, K. and Bode, J.** (1997) A thymocyte factor SATB1 suppresses transcription of stably integrated matrix-attachment region-linked reporter genes. *Biochemistry*, **36**, 12005–12010.

**Kolesnik, T., Szeverenyi, I., Bachmann, D., Kumar, C., Jiang, S., Ramamoorthy, R., Cai, M., Ma, Z., Sundaresan, V. and Ramachandran, S.** (2004) Establishing an efficient *Ac/Ds* tagging system in rice: large-scale analysis of *Ds* flanking sequences. *Plant J.* **37**, 301–314.

**Koprek, T., McElroy, D., Louwerse, J., Williams-Carrier, R. and Lemaux, P.G.** (2000) An efficient method for dispersing *Ds* elements in the barley genome as a tool for determining gene function. *Plant J.* **24**, 253–263.

**Koprek, T., Rangel, S., McElroy, D., Louwerse, J.D., Williams-Carrier, R.E. and Lemaux, P.G.** (2001) Transposon-mediated single-copy gene delivery leads to increased transgene expression stability in barley. *Plant Physiol.* **125**, 1354–1362.

**Kota, R., Rudd, S., Facius, A., Kolesov, G., Thiel, T., Zhang, H., Stein, N., Mayer, K. and Graner, A.** (2003) Snipping polymorphisms from large EST collections in barley (*Hordeum vulgare* L.). *Mol. Genet. Genomics*, **270**, 24–32.

**Kuromori, T., Hirayama, T., Kiyosue, Y., Takabe, H., Mizukado, S., Sakurai, T., Akiyama, K., Kamiya, A., Ito, T. and Shinozaki, K.** (2004) A collection of 11 800 singlecopy *Ds* transposon insertion lines in *Arabidopsis*. *Plant J.* **37**, 897–905.

**Liao, G.C., Rhem, E.J. and Rubin, G.M.** (2000) Insertion site preferences of the P transposable element in *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA*, **97**, 3347–3351.

**Manly, K.F., Robert, H.C. and Meer, J.M.** (2001) Map Manager QTX, cross-platform software for genetic mapping. *Mamm. Genome*, **12**, 930–932.

**Martienssen, R. and Colot, V.** (2001) DNA methylation and epigenetic inheritance in plants and filamentous fungi. *Science*, **293**, 1070–1074.

**McElroy, D., Louwerse, J., McElroy, S. and Lemaux, P.G.** (1997) Development of a simple transient assay for *Ac/Dc* activity in cells of intact barley tissue. *Plant J.* **11**, 157–165.

**Melzer, J.M. and Kleinhofs, A.** (1987) Molecular genetics of barley. In *Proceedings of the Fifth International Barley Genetics Symposium* (Yanuda, S. and Konishi, T., eds). Okayama: Sanyo Press, pp. 481–491.

**Mielke, C., Kohwi, Y., Kohwi-Shigematsu, T. and Bode, J.** (1990) Hierarchical binding of DNA fragments derived from scaffold-attached regions: correlation of properties in vitro and function in vivo. *Biochemistry*, **29**, 7475–7485.

**Mirkovitch, J., Gasser, S.M. and Laemmli, U.K.** (1988) Scaffold attachment of DNA loops in metaphase chromosomes. *J. Mol. Biol.* **200**, 101–109.

**Moore, G.** (1995) Cereal genome evolution: pastoral pursuits with 'Lego' genomes. *Curr. Opin. Genet. Dev.* **5**, 717–724.

**Muehlbauer, G.J., Bhau, S.B., Syed, N.H., Heinen, S., Cho., S., Marshall, D., Pateyron, S., Buisine, N., Chalhoub, B. and Flavell, A.J.** (2006) A hAT superfamily transposase recruited by the cereal grass genome. *Mol. Genet. Genomics*, e-publication ahead of print, February 2006.

**Müller, H. and Varmus, H.** (1994) DNA bending creates favored sites for retroviral integration: an explanation for preferred insertion sites in nucleosomes. *EMBO J.* **13**, 4704–4714.

**Nakagawa, Y., Machida, C., Machida, Y. and Toriyama, K.** (2000) Frequency and pattern of transposition of the maize transposable element *Ds* in transgenic rice plants. *Plant Cell Physiol.* **41**, 733–742.

 

**Nomura, K., Saito, W., Ono, K., Moriyama, H., Takahashi, S., Inoue, M. and Masuda, K.** (1997) Isolation and characterization of matrix associated region DNA fragments in rice (*Oryza sativa* L.). *Plant Cell Physiol.* **38**, 1060–1068.

**Okamoto, H. and Hirochika, H.** (2001) Silencing of transposable elements in plants. *Trends Plant Sci.* **6**, 527–534.

**Palmer, L.E., Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V.S., Nascimento, L.U., Dike, S., De la Bastide, M., Martienssen, R.A. and McCombie, W.R.** (2003) Maize genome sequencing by methylation filtration. *Science*, **302**, 2115–2117.

**Pan, X., Li, Y. and Stein, L.** (2005) Site preferences of insertional mutagenesis agents in Arabidopsis. *Plant Physiol.* **137**, 168–175.

**Parinov, S., Sevugan, M., Ye, D., Yang, W.C., Kumaran, M. and Sundaresan, V.** (1999) Analysis of flanking sequences from dissociation insertion lines: a database for reverse genetics in *Arabidopsis*. *Plant Cell*, **11**, 2263–2270.

**Rabinowicz, P.D., Schutz, K., Dedhia, N., Yordan, C., Parnell, L.D., Stein, L., McCombie, W.R. and Martienssen, R.A.** (1999) Differential methylation of genes and retrotransposons facilitates shotgun sequencing of the maize genome. *Nature Genet.* **23**, 305–308.

**Raina, S., Mahalingam, R., Chen, F. and Fedoroff, N.** (2002) A collection of sequenced and mapped *Ds* transposon insertion sites in *Arabidopsis thaliana*. *Plant Mol. Biol.* **50**, 93–110.

**Ramakrishna, W., Dubcovsky, J., Park, Y.J., Busso, C., Emberton, J., SanMiguel, P. and Bennetzen, J.L.** (2002) Different types and rates of genome evolution detected by comparative sequence analysis of orthologous segments from four cereal genomes. *Genetics*, **162**, 1389–1400.

**Read, B.J., Raman, H., McMichael, G. *et al.*** (2003) Mapping and QTL analysis of the barley population Sloop × Halcyon. *Aust. J. Agric. Res.* **54**, 1145–1153.

**Romig, H., Ruff, J., Fackelmeyer, F., Patil, M. and Richter, A.** (1994) Characterization of two intronic nuclear matrix attachment regions in the human DNA topoisomerase I gene. *Eur. J. Biochem.* **221**, 411–419.

**Sallaud, C., Gay, C., Larmande, P. *et al.*** (2004) High throughput T-DNA insertion mutagenesis in rice: a first step towards in silico reverse genetics. *Plant J.* **39**:450–465.

**Sandhu, D. and Gill, K.S.** (2002) Gene-containing regions of wheat and other grass genomes. *Plant Physiol.* **128**, 803–811.

**Scholtz, S., Lörz, H. and Lutticke, S.** (2001) Transposition of maize transposable element Ac in barley (*Hordeum vulgare* L.). *Mol. Gen. Genet.* **264**, 653–661.

**Schulman, A.H. and Kalendar R.** (2005) A movable feast: diverse retrotransposons and their contribution to barley genome dynamics. *Cytogenet. Genome Res.* **110**, 598–605.

**Sentinelli, F., Lovari, S., Vitale, M., Giorgi, G., Mario, U. and Baroni, M.** (2000) A simple method for non-radioactive PCR-SSCP using MDETM gel solution and midi gel format: application for the detection of variants in the GLUT1 and CTLA-4 genes. *J. Biotechnol.* **78**, 201–204.

**Sha, Y., Li, S., Pei, Z., Luo, L., Tian, Y. and He, C.** (2004) Generation and flanking sequence analysis of a rice T-DNA tagged population. *Theor. Appl. Genet.* **108**, 306–314.

**Shirasu, K., Lahaye, T., Tan, M.-W., Zhou, F., Azevedo, C. and Schulze-Lefert, P.** (1999) A novel class of eukaryotic zinc-binding proteins is required for disease resistance signaling in barley and development in *C. elegans*. *Cell*, **99**, 355–366.

**Singh, G.B., Kramer, J.A. and Krawetz, S.A.** (1997) Mathematical model to predict regions of chromatin attachment to the nuclear matrix. *Nucleic Acids Res.* **25**, 1419–1425.

**Stephanova, E., Stancheva, R. and Avramova, Z.** (1993) Binding of sequences from the 5′ and 3′ nontranscribed spacers of the rat rDNA locus to the nucleolar matrix. *Chromosoma*, **102**, 287–295.

**Szabados, L., Kovacs, I., Oberschall, A. *et al.*** (2002) Distribution of 1000 sequenced T-DNA tags in the Arabidopsis genome. *Plant J.* **32**, 233–242.

**Tikhonov, A.P., Lavie, L., Tatout, C., Bennetzen, J.L., Avramova, Z. and Deragon, J.M.** (2001) Target sites for SINE integration in Brassica genomes display nuclear matrix binding activity. *Chromosome Res.* **9**, 325–337.

**Tingay, S., McElroy, D., Kalla, R., Fieg, S., Wang, M., Thornton, S. and Brettell, R.** (1997) *Agrobacterium tumefaciens*-mediated barley transformation. *Plant J.* **11**, 1369–1376.

**Tinland, B.** (1996) The integration of T-DNA into plant genomes. *Trends Plant Sci.* **1**, 178–184.

**Vicient, C.M., Jaaskelainen, M.J., Kalendar, R. and Schulman, A.H.** (2001) Active retrotransposons are a common feature of grass genomes. *Plant Physiol.* **125**, 1283–1292.

**Wan, Y. and Lemaux, P.G.** (1994) Generation of large number of independently transformed fertile barley plants. *Plant Physiol.* **104**, 37–48.

**Wang, M.B., Li, Z., Matthews, P.R., Upadhyaya, N.M. and Waterhouse, P.M.** (1998) Improved vectors for *Agrobacterium tumefaciens*-mediated transformation of monocot plants. *Acta Hortic.* **461**, 401–407.

**Wang, M.B., Abbott, D.C., Upadhyaya, N.M., Jacobsen, J.V. and Waterhouse, P.M.** (2001) *Agrobacterium tumefaciens*-mediated transformation of an elite Australian barley cultivar with virus resistance and reporter genes. *Aust. J. Plant Physiol.* **28**, 149–156.

**Waugh, R., McLean, K., Flavell, A., Pearce, S., Kumar, A., Thomas, B.B.T. and Powell, W.** (1997) Genetic distribution of BARE-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). *Mol. Gen. Genet.* **253**, 687–694.

**Yamazaki, M., Tsugawa, H., Miyao, A., Yano, M., Wu, J., Yamamoto, S., Matsumoto, T., Sasaki, T. and Hirochika, H.** (2001) The rice retrotransposon *Tos17* prefers low copy number sequences as integration targets. *Mol. Gen. Genet.* **265**, 336–344.

**Yu, Y., Tomkins, J.P., Waugh, R., Frisch, D.A., Kudrna, D., Kleinhofs, A., Brueggeman, R.S., Muehlbauer, G.J., Wise, R.P. and Wing, R.A.** (2000) A bacterial artificial chromosome library for barley (*Hordeum vulgare* L.) and the identification of clones containing putative resistance genes. *Theor. Appl. Genet.* **101**, 1093–1099.

**Zhou, F.S., Kurth, J.C., Wei, F., Elliott, C., Vale, G., Yahiaoui, N., Keller, B., Somerville, S., Wise, R. and Schulze-Lefert, P.** (2001) Cell-autonomous expression of barley Mla1 confers race-specific resistance to the powdery mildew fungus via a Rar1-independent signaling pathway. *Plant Cell*, **13**, 337–350.

 