Theor Appl Genet (2001) 102:773–781

R.-Y. Chang · L.S. O'Donoughue · T.E. Bureau

Inter-MITE polymorphisms (IMP): a high throughput transposon-based genome mapping and fingerprinting approach

Received: 24 March 2000 / Accepted: 28 July 2000

Abstract Miniature inverted-repeat transposable elements or MITEs represent a large superfamily of transposons that are moderately to highly repetitive and frequently associated with plant genes. These attributes were exploited in the development of a powerful marker technology called Inter-MITE polymorphism, or IMP, which involves the amplification between two adjacent MITEs. In this report, we describe the utility of the IMP approach in the mapping and fingerprinting of the barley genome. MITEs were systematically mined from barley genomic gene sequences by computer-assisted database searches and structural analysis. Barley MITEs include members of the *Stowaway* family and a new family we have named *Barfly*. Using these barley MITEs, a total of 88 IMP markers were mapped onto an existing barley RFLP map that was based on a doubled-haploid segregating population between *Hordeum vulgare* and *Hordeum spontaneum*. We demonstrate that the IMP approach can be effectively applied in the fingerprinting of barley cultivars and for genetic similarity analysis. We also provide evidence that barley MITE-based primers can be effectively used in the mapping and fingerprinting of other cereals, suggesting that the IMP approach has broad applicability.

Keywords IMP · MITE · Mapping · Fingerprinting · Barley

Communicated by G. Wenzel

R.-Y. Chang · T.E. Bureau (✉)
McGill University, Department of Biology,
1205 Docteur Penfield Avenue, Montreal,
Quebec H3A 1B1, Canada
e-mail: thomas_bureau@maclan.mcgill.ca

L.S. O'Donoughue
DNA LandMarks, Inc., P.O. Box 6, St-Jean-sur-Richelieu,
Quebec J3B 6Z1, Canada

Introduction

The development of the restriction fragment length polymorphism (RFLP; Bostein et al. 1980), random amplified polymorphic DNA (RAPD; Welsh and McClelland 1990;Williams et al. 1990) and amplified fragment length polymorphism (AFLP; Vos et al. 1995) techniques as molecular mapping tools has facilitated the rapid evolution of genome mapping and fingerprinting technologies. Modern genome mapping and fingerprinting techniques have been made even more powerful by exploiting the use of repetitive genomic sequences usually derived from retroelements (Nelson et al. 1989) and simple sequence repeats (SSRs; Litt and Lutty 1989; Tautz 1989; Weber and May 1989; Zietkiewicz et al. 1994).

Retroelements are only one class of transposable elements found in eukaryotic genomes. In fact, many transposable elements have been identified by genetic and molecular approaches. Transposable elements have also been mined by computer-based sequence-similarity searches. For example, Bureau et al. (1996) identified numerous members of a novel superfamily of transposable elements called miniature inverted-repeat transposable elements or MITEs. Although MITEs have terminal inverted-repeats (TIRs), they differ from other well-characterized plant transposable elements in that they are small in size (<500 bp) and have a target-site preference for either TAA or TA. Furthermore, MITEs appear to prefer insertion near or within genes and are present in moderate to high copy number in many, if not all, plant genomes. We suggest that these characteristics lend themselves to the development of a PCR-based approach to identify markers linked to characters of agronomic importance in plants.

We identified potentially informative transposable elements for barley genome mapping and fingerprinting purposes by computer-based sequence similarity searches. These elements were exploited in a novel PCR-based technique that we have termed inter-MITE polymorphisms or IMPs. In this approach, polymorphisms were revealed with primers designed from the TIRs of MITEs.

Table 1 MITEs associated with barley gene sequences. e, end of available DNA sequence; t, truncated sequence; ND, not determined

Name of gene	Accession #	Location	Size (bp)
Stowaway			
Protein kinase	X65606	Intron 2, 2261–2341	81
Rubisco activase RcaA	M55449	Intron 3, 6201–6359	159
1,3–1,4-ß-Glucanase	M62740	Intron 1, 1202–1524	323
Cold-regulated protein	M60732	3´UTR, 604–652e	NA
Pathogenesis-related Hv-1	X58564	3´UTR, 669–809e	NA
Rubisco activase RcaB	M55449	Intron 1, 1975–2272	298
Serine carboxypeptidase II	Y09602	5´, 1302t-1363	ND
Serine carboxypeptidase II	Y09602	5´, 1458–1499 t	ND
Pathogenesis-related protein 4	Y10814	3´UTR, 640–714e	ND
Caffeic acid 0-methyltransferase	U54767	Intron 2, 3671–3752	82
Caffeic acid 0-methyltransferase	U54767	Intron 3, 3967–4126	160
ß-Glucosidase BGQ60	L41869	3´, 3016–3128	113
Lipoxygenase isoenzyme I	U83904	Intron 2, 4986–5131	146
Chitinase CHI26	L34210	3´UTR, 2721–2804	84
Homeobox	X83518	Intron 3, 7203–7295	92
Homeobox	X83518	Intron 3, 7538–7617	80
Homeobox	X83518	Intron 3, 8130–8206	77
Plastocyanin NK 1558	Z28347	5´, 80–295	216
Plastocyanin NK 1558	Z28347	3´, 1801–1884e	ND
Xylose isomerase	X95256	Intron 5, 1554–1622	69
Lipid transfer protein Ltp4.3	Z66528	3´UTR, 1703–1783	81
Blz-1	X80068	Intron 1, 945–1032	88
Blz-1	X80068	Intron 3, 2808–2901	94
Heat shock protein 17 kDa	X64560	5´, 132–223	92
Tourist			
Acyl-carrier protein 1	M58753	3´, 3369–3496	128
Chalcone synthase	X58339	Intron 1, 2898–3065	168
Barfly			
Homeobox	X83518	Intron 4, 8748–8987	240
Seed protein	X64254	5´, 1e-213	ND

We assessed the usefulness of these MITE-based primers by examining segregation patterns in a *Hordeum* spp. doubled-haploid (DH) mapping population and in fingerprinting 26 cultivars of *Hordeum vulgare* and one accession of cultivar of *Hordeum spontaneum*.

Materials and methods

Plant materials

The mapping population consisted of 88 DH individuals from a cross between *H. vulgare* cultivar Lina and *H. spontaneum* Canada Park. This population is being used to construct a linkage map based mostly on RFLP markers (O'Donoughue, unpublished results). A total of 27 lines were used in the fingerprinting experiments including 26 *H. vulgare* entries and one *H. spontaneum* entry, Canada Park. The collection included two-row and six-row types. Among the two-row types, both spring and winter cultivars were included. All 27 lines were previously used in an RFLP genotyping study (O'Donoughue and Ménard, unpublished results) and therefore the RFLP-based genetic relationships among these cultivars were known. A subset of a wheat (*Triticum aestivum* L.) W-7984×Opata85 (van Deynze et al. 1995) and of an oat (*Avena sativa* L.) Kanota×Ogle (O'Donoughue et al. 1995) recombinant inbred line (RIL) population, as well as the maize (*Zea mays* spp. *mays*) lines B73, Mo17, W22, Argentina and Uruguay, were used to assess the general applicability of the IMP approach.

Computer-based sequence similarity searches

DNA sequence identification and analysis were accomplished using the programs available from the University of Wisconsin Genetics Computer Group (UWGCG, version 9.0) suite and the BLAST search algorithm accessed via the website at the National Center for Biotechnology Information (NCBI, January 1998). Barley gene sequences were compiled using the program STRINGS and the keywords hordeum, vulgare, spontaneum, and barley. Each genomic gene sequence was used as a query in a search of the Genbank and EMBL DNA sequence databases. Non-coding sequences sharing significant (BLAST score >80) sequence similarity were compiled. Each group of sequences were analyzed for features indicative of transposable elements (Bureau and Wessler 1994; Bureau et al. 1996).

Polymerase chain reaction

Two degenerate MITE-based primers (TEM-1 and TEM-10) were designed and evaluated in this study. TEM-1, 5´-(AG)TAT TT(TA)GGAACGGAGGGAG-3´, is a barley *Stowaway* element (Bureau and Wessler 1994) primer based on 44 TIR sequences. TEM-10, 5´-TCCCCA(CT)T(AG)TGACCA(CGT)CC-3´, was designed from the consensus sequence of seven *Barfly* TIRs. The parentheses indicate degeneracy at these positions. Both TIRs were used to design the primers. Primers were labeled with IRD700 fluorescent dye (LI-COR, Inc., Lincoln, Nebraska). PCR products were visualized on a 6% polyacrylamide denaturing gel cast with 41-cm-long glass plates. Gel electrophoreses were carried out using a LI-COR 4200 automated detection system.

PCR-amplifications were performed in a 20 µl vol containing 2.5 mM of $MgCl_2$, 0.4 mM of dNTPs, 1.0 µM (non-labeled) or 0.1 µM (labeled) of primer and 0.5 units of Ampli*Taq* DNA polymerase (Perkin-Elmer). The following profile was used for both

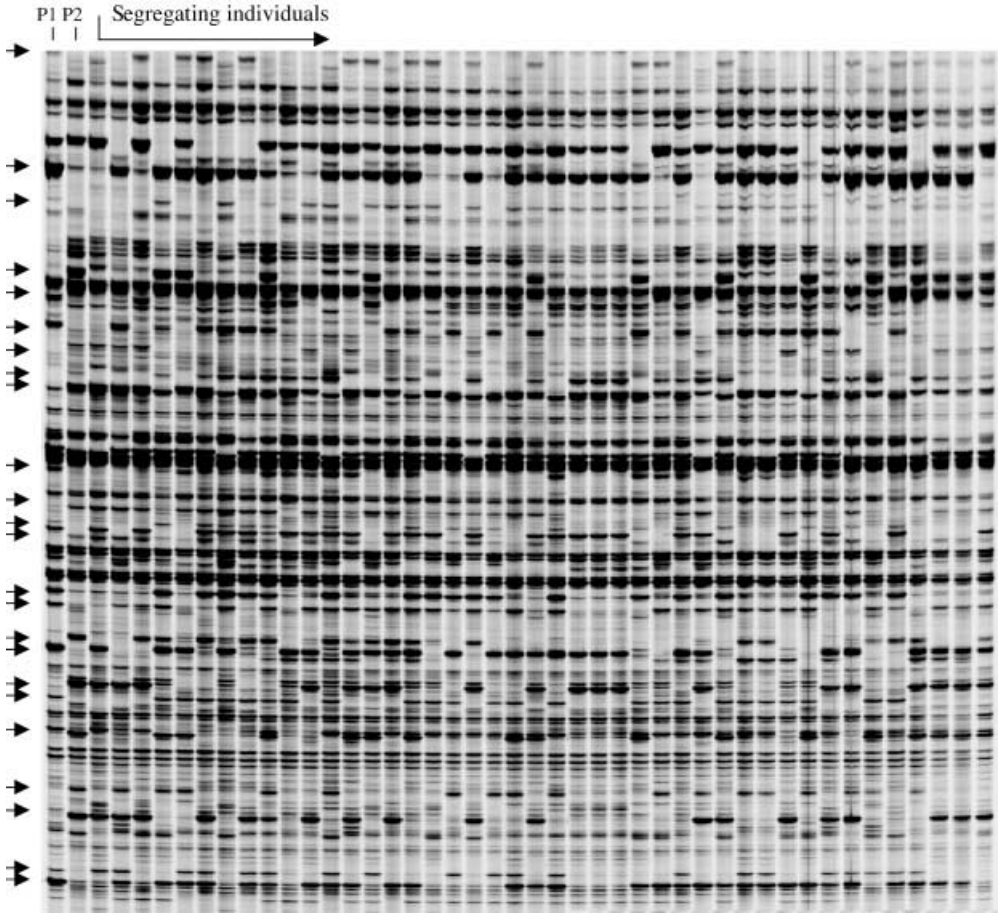

Fig. 1 A subset of the PCR results of IRD700 fluorescence dye-labeled TEM-1 primer, visualized on a 6% polyacrylamide gel with the LI-COR 4200 automated system. Parent *H. vulgare*, Lina (*P1*); parent *H. spontaneum*, Canada Park (*P2*), and the segregating individuals from the Lina×Canada Park DH population are shown

the pre-amplification and amplification steps: an initial denaturation of 1 min 30 s at 94°C; followed by 35 cycles of 30 s at 94°C, 45 s at 58°C (TEM-1) or 56°C (TEM-10), and 1 min at 72°C; with a final extension of 5 min at 72°C. PCR-amplifications were done in two steps. The first step was a pre-amplification with non-labeled primers for 35 cycles with 100 ng of DNA template. An aliquot of 2 µl of the pre-amplification product was used for the second step of amplification.

Data collection and statistical analyses

For the fingerprinting and genetic similarity analyses, polymorphic bands were scored as present (1), absent (0), or as missing data (9) for each individual. The resulting raw-data matrices were used to generate relative genetic-similarity (GS) matrices using Nei and Li's measurement (Nei and Li 1979), $2n_{xy}/(n_x+n_y)$, where n_x and n_y are the numbers of bands in lines x and y, respectively, and n_{xy} is the number of bands shared by both lines. Only polymorphic bands were used to calculate the GS values. Dendrograms were generated based on the GS matrices using the unweighted pair-group method arithmetic average (UPGMA). A combined dendrogram resulting from analyses with the two MITE primers (TEM-1 and TEM-10) was generated. The normalized Mantel statistic (Z; Mantel 1967) was used to compare the GS matrix based on the IMP markers with that of the same lines based on 313 polymorphic RFLP marker fragments (O'Donoughue and Menard, unpublished results). The test of significance was performed by comparing the observed Z-value with the distribution of 1000 random permutations of the matrices. All these statistical analyses were performed with the NTSYS-pc software (Rohlf 1994).

The localization of the IMP markers was performed by mapping them within a framework of 71 markers (mostly RFLPs) that had been used previously to construct a map of the *H. vulgare* cultivar Lina×*H. spontaneum* Canada Park population (O'Donoughue, unpublished). Mapping was performed using the computer program MAPMAKER (Lander et al. 1987). The markers were assigned to linkage groups using two-point analysis at a LOD threshold of 4 [except for group 7H which formed two groups at this threshold and were linked based on the published location of RFLP markers (Graner et al. 1991; Heun et al. 1991; Kleinhofs et al. 1993)]. Multipoint analysis with a LOD threshold of 2 was used to place the markers within the linkage groups.

Reproducibility test

To assess the reproducibility of the IMP approach, 12 entries (Lina, *H. spontaneum*, Alexis, Angora, Azhul, Ellice, Goldie, Golf, Ingri, Ingrid, Maud and Saxo) were used in three independent replications under the same conditions. PCR products of the three reactions were scored using the computer program Cross Checker (version 2.8; Dr. J.B. Buntjer, Wageningen University, The Netherlands). A given band was considered to be reproducible when it was present or absent in all three sets of reactions; otherwise it was considered to be non-reproducible.

Results

MITE identification and primer design

Although a few *Tourist* and *Stowaway* elements were previously reported to be associated with barley gene

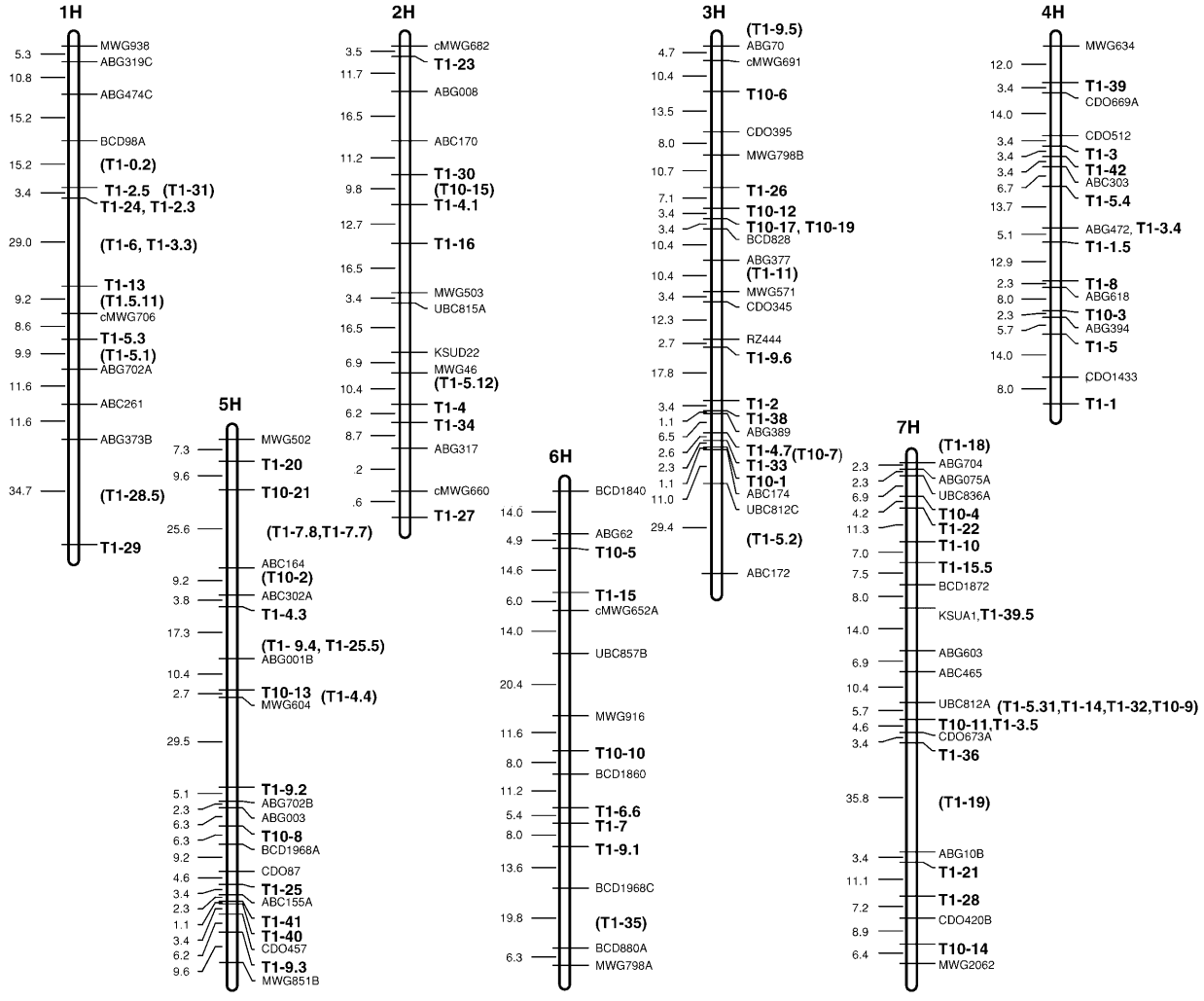

Fig. 2 Linkage map of the *H. vulgare* cv Lina×*H. spontaneum* Canada Park population. The newly placed IMP markers are shown in *bold characters*. *Loci in parentheses* could not be placed with a LOD score greater or equal to 2 but are assigned to their most-likely intervals

sequences, these sequences were re-examined to identify barley MITEs suitable for a PCR-based mapping protocol. Eighteen *Stowaway* and two *Tourist* elements were identified (Table 1). In addition a new family of MITEs, named *Barfly*, was identified with one complete element member and one partial element (Table 1). In the process of mining MITEs from barley gene sequences, several MITEs from other cereal genes were found, including two *Barfly* elements from wheat (accession number X56004, position 479–751; accession number X94693, position 187–460). *Barfly* elements share many characteristics of *Tourist*, including a *Tourist*-like TIR sequence [consensus, 5´-$G_{75}G_{75}C_{75}T_{88}G_{63}C_{75}T_{75}C_{88}A_{75}(T_{38}/C_{38})A_{88}A_{75}T_{88}G_{88}G_{88}G_{63}G_{63}A_{100}G_{63}$–3´, subscript numbers are the percentages of occurrence of each particular base at that specific position], a small size, and with a target-site preference for TAA. Yet, *Barfly* shares no overall significant sequence similarity with any reported *Tourist* element. The TIR sequences of the *Stowaway*, *Tourist* and *Barfly* elements were aligned to generate a consensus sequence for each family. The sequence diversity among the mined *Tourist* elements was deemed too high to derive a meaningful consensus sequence with an acceptable degree of degeneracy. Outwardly directed oligonucleotide primers were designed according to the consensus sequences of *Stowaway* (primer TEM-1) and *Barfly* (TEM-10) TIRs. In this way, sequences amplified by these primers were expected to lie between two adjacent MITEs within amplifiable distances.

Chromosomal localization of IMP markers

To evaluate the feasibility of using MITE sequences in PCR-based mapping and genotyping, these two primers, TEM-1 and TEM-10, were used alone in a segregation analyses involving a DH population of 88 individuals from a cross between *H. vulgare* cultivar Lina and *H. spontaneum* cultivar Canada Park. Approximately 120 and 90 clear bands were detected on a LI-COR sequencing gel with primers TEM-1 and TEM-10, respectively. The size range of the bands detected was approximately 100 bp

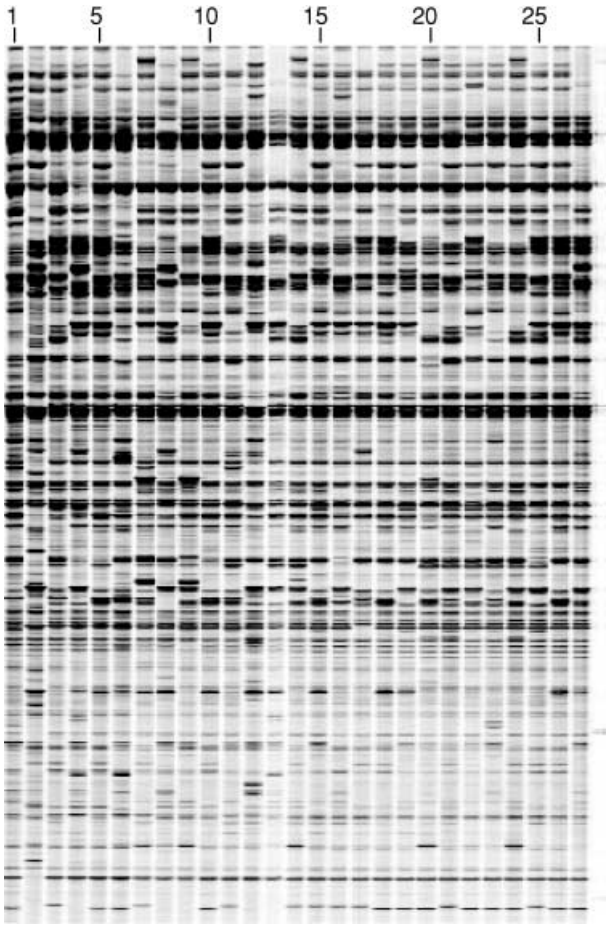

Fig. 3 Fingerprinting results obtained from 27 barley lines using the IRD700 fluorescence dye-labeled TEM-1 primer in an IMP approach. *Lanes 1–27* represent Lina, Canada Park, Alexis, Angora, Ariel, Azhul, Ellice, Express, Fillipa, Goldie, Golf, High amylose Glacier, Igri, Ingrid, Kinnan, Maud, Meltan, Mentor, Mette, Mona, Roland, Saxo, Svani, Tellus, Tofta, Trebon and Vixen, respectively. Only a section of the polyacrylamide gel is shown

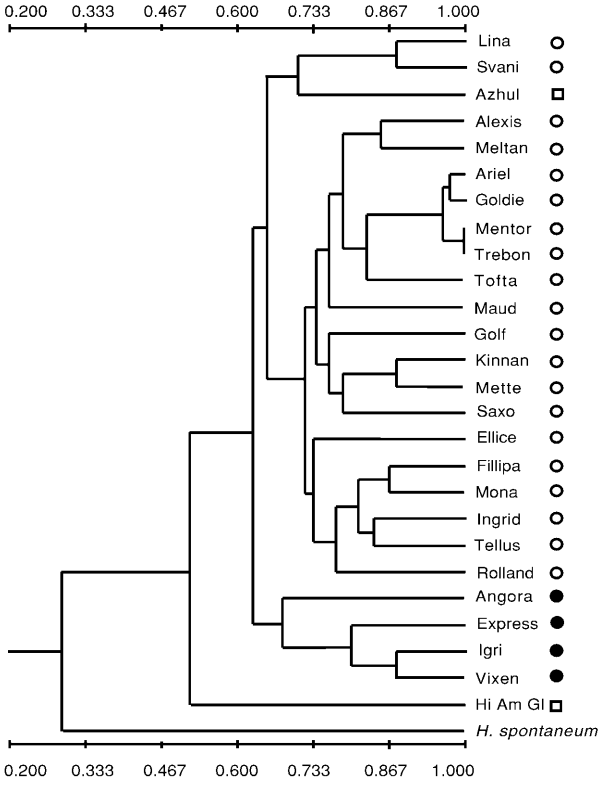

Fig. 4 Dendrogram resulting from the UPGMA clustering of the genetic similarity matrix of 27 barley lines, based on the TEM-1 and TEM-10 banding patterns. *Open circles* represent spring two-row type; *filled circles* winter two-row type; and *open squares* six-row type

to 1 kb. Part of the amplification result with TEM-1, as visualized by polyacrylamide gel electrophoresis, is shown in Fig. 1.

Seventy five and 19 polymorphic bands were generated with the TEM–1 and TEM–10 primers, respectively. Some pairs of bands exhibited co-dominant behavior (loci T1–0.2 on 1H; T1–4 and T1–16 on 2H; T10–6 on 3H; T1–8 on group 4H; and T1–36 on 7H; Fig. 2), but the remaining bands exhibited a presence/absence pattern with exactly 41 derived from the Lina parent and 41 from the *H. spontaneum* parent. Of the 70 mapped TEM-1 loci, 24 significantly deviated from the expected 1:1 segregation ratio. All 24 loci except one (T1–19 on 7H) mapped to areas where RFLP markers also exhibited distorted segregation ratios in this mapping population. Two of the 18 TEM-10 loci significantly deviated from the expected segregation ratio and these were again located in areas where RFLP loci also deviated from the expected 1:1 ratio.

In total, 88 loci were mapped. These loci covered all seven linkage groups (Fig. 2). Furthermore, the distribution of the loci showed no significant clustering other than that which would be expected around centromeric regions where recombination is typically reduced (e.g., groups 1H, 3H and 7H, Fig. 2). In fact, the distribution is similar to that found with cDNAs detecting RFLPs (L.S. O'Donoughue, unpublished). This suggests that MITEs are located in areas of the genome containing coding sequences and that it will be possible to cover the entire genome with a limited set of MITE-based primers.

Fingerprinting

A total of 27 accessions, which included the *H. vulgare* parent Lina, the *H. spontaneum* parent Canada Park and 25 other *H. vulgare* cultivars, were used to assess the usefulness of IMPs for DNA fingerprinting and genetic-similarity analyses. A total of 61 polymorphic bands were scored using the TEM-1 primer across the 27 cultivars. A section of this gel is shown in Fig. 3. A total of 28 polymorphic bands were scored using the TEM-10 primer.

Nei and Li's coefficient (Nei and Li 1979) GS matrices were generated with the combined data from the TEM-1 and TEM-10 amplifications. Dendrograms were generated by UPGMA clustering. The dendrogram of the combined data from TEM-1 and TEM-10 is shown in Fig. 4. This dendrogram clearly separates the *H. spontaneum* accession

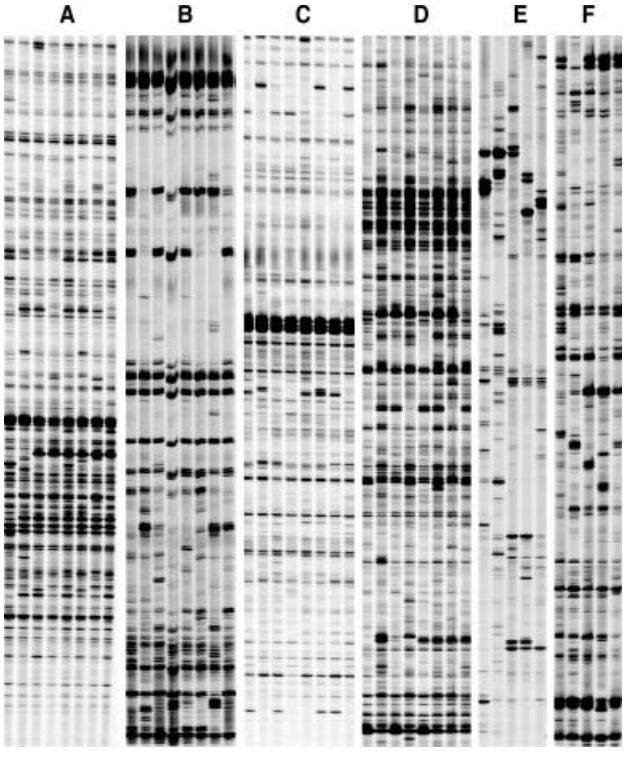

Fig. 5 Application of the IMP approach with other cereals using barley MITE-specific primers. Panels represent the following: *A* TEM-1 with wheat; *B* TEM-10 with wheat; *C* TEM-1 with oats; *D* TEM-10 with oats, *E* TEM-1 with maize and *F* TEM-10 with maize. For panels *A–D, lanes 1 and 2* are the parents and *lanes 3–8* are individuals in the segregating populations. Panels *E* and *F* indicate the results obtained from maize lines in the order of B73, Mo17, W22, Argentina and Uruguay

from the *H. vulgare* cultivars. The six-row *H. vulgare* type, high-amylose Glacier, remains separate from the two-row types. With the exception of Azhul (six-row type), the two-row cluster is subdivided into a spring-type sub-cluster and winter-type sub-cluster. It must be noted that the GS measured with both the IMP markers and the RFLP markers was calculated using only polymorphic bands and that fragments in both cases were scored as dominant markers. Therefore the GS measures should be viewed as relative, rather than absolute, measures. With a Mantel statistic of Z=0.69475, the IMP GS matrix correlated well with that based on the previous RFLP analysis. This despite the IMP GS matrix being based on only 89 bands (as opposed to 313 for the RFLP GS matrix). The positive correlation was highly significant with a probability of P=0.0020 for this value of Z to be obtained by chance alone.

IMP reproducibility

With the 12 cultivars used in the reproducibility test, 97 bands were scored for primer TEM-1. This primer showed a 87.03% reproducibility. A total of 64 bands were scored for primer TEM-10 and a lower level of reproducibility was observed (67.06%).

IMP analyses in other crop species

IMP analysis using the primers derived from barley elements was also assayed on wheat, oat and maize. Successful amplification was obtained in all three of these grass species (Fig. 5). With the wheat W7984× Opata85 RILs, approximately 100 bands were visualized with TEM-1 and about 40 with TEM-10. Approximately 20 (20%) of the TEM-1 bands segregated, whereas ten (25%) of the TEM-10 bands were polymorphic. Considerably fewer bands were detected with the oat Kanota× Ogle RILs, approximately 50 bands with TEM-1 and 15 bands with TEM-10. Of these bands, 30 (60%) with TEM-1 and five (33%) with TEM-10 were polymorphic. Approximately 100 bands were generated using either of TEM-1 or TEM-10 with the maize lines B73, Mo17, W22, Argentina and Uruguay. Whereas approximately 20 bands from the TEM-10 amplification were common, there were no common bands across the maize lines in the TEM-1 amplification, indicating a very high level of polymorphism.

Discussion

In this report, we have shown that MITEs: (1) can be readily mined from genomic gene sequences in nucleic acid sequence databases, and (2) are effective molecular anchor sequences in the development of informative genomic markers.

Transposable elements are fundamental components of eukaryotic genomes (Finnegan 1989; Pearce et al. 1996; Suoniemi et al. 1996). In plants, retrotransposons are present at high copy number and contribute the majority of sequences within the intergenic regions of some genomes (Flavell et al. 1992; Voytas et al. 1992; Hirochika and Hirochika 1993; SanMiguel et al. 1996). MITEs are also present at high copy number and are frequently found in close association with plant genes (Bureau et al. 1996). They represent a large superfamily of elements that belong to the TIR class of transposons. Most MITEs identified by computer-assisted database searches belong to the *Stowaway* and *Tourist* element families. Our search of elements associated with barley gene sequences reveals that the majority belong to either the *Stowaway* or *Tourist* family and to a newly identified family, called *Barfly*. Members of the *Barfly* family have TIR sequences and a target-site sequence preference, TAA, reminiscent of members of the *Tourist* family.

The ubiquity and dispersion of transposable elements throughout the genome suggest that they can be exploited as PCR-based mapping tools. Indeed, Nelson et al. (1989) and Sinnet et al. (1990) used *Alu*-specific primers in search of polymorphisms among different human DNA samples. *Alu* is a highly repeated short interspersed nuclear element (SINE) found in many primate genomes. *Alu*-based polymorphisms, or alumorphs, have proven to be an informative genome-analysis tool (Sinnet et al. 1990). For example, an alumorph has been found to be

linked to a gene underlying the human disease pseudo-vitamin D-deficiency rickets (Zietkiewicz et al. 1992). In plants, a *copia*-like retrotransposon, PDR1, was also successfully used to study polymorphisms and, in combination with other specific primers, to diagnose different lines in *Pisum* (Lee et al. 1990). The close association of transposable elements with genes has made possible the use of these elements as anchor points to isolate gene sequences (Nelson et al. 1989; Souer et al. 1995; Broeck et al. 1998).

More recently, different marker techniques based on transposable elements (primarily retrotransposons) have started to emerge. Waugh et al. (1997) designed a strategy termed sequence-specific amplification polymorphism (S-SAP) that combines an AFLP approach with the use of a barley retrotransposon Bare-1 primer. In this method, a ligation adapter and its corresponding primer is replaced with a segment of the LTR sequence of Bare-1. Forty eight S-SAP bands from 12 primer combinations were mapped in a DH population and were dispersed throughout the barley genome. Ellis et al. (1998) exploited the S-SAP technique successfully in pea with the PDR1 element both in mapping and in phylogenetic analysis of different subspecies and modern cultivars of *Pisum*. Kalendar et al. (1999) used the anchored-primer strategies inter-transposon amplified polymorphism (IRAP) and retrotransposon-microsatellite amplified polymorphism (REMAP) with Bare-1 to fingerprint different barley species and varieties. Another anchored-primer PCR-based assay, *copia*-SSR, was used by Provan et al. (1999) to map seven markers to four chromosome arms using a barley doubled-haploid population. Flavell et al. (1998) employed a PDR1-specific primer, as well as a primer from a known genomic sequence flanking the element, to study PDR1 insertion-site polymorphisms. A similar technique called transposon display was designed by Broeck et al. (1998) using the dTph1 element of *Petunia hybrida*, in which a 6-base restriction enzyme was chosen within the element itself and, together with a 4-base restriction enzyme, was used to anchor the two primers for the amplification reactions.

We have systematically exploited MITEs in the development of a new transposon-based mapping and fingerprinting approach in barley. This approach relies on the amplification between two adjacent MITEs and is similar in principle to methodologies such as RAPDs (Welsh and McClelland 1990; Williams et al. 1990), (SSR)-anchored PCR (Zietkiewicz et al. 1994) and *Alu*-PCR (Nelson et al. 1989). The approach is simple with no need for restriction endonuclease digestions and adapter ligations such as with the S-SAP and transposon-display techniques. A single primer can be used as well as combinations of two different MITE primers (data not shown). IMPs can be detected using both agarose (data not shown) and fluorescence-based detection approaches on acrylamide gels, though the latter is superior for revealing the full potential of the technique.

We obtained up to 120 bands with the single TEM-1 *Stowaway* primer, 75 of which segregated in our barley mapping population. This indicates that MITEs may be more abundant in the genome than the retrotransposons used previously in other transposon-based marker technologies. More *Stowaway* elements than *Barfly* or *Tourist* elements were detected in our database search. Also, the *Stowaway* TEM-1 primer amplifications resulted in the amplification of more bands than amplification with the *Barfly* TEM-10 primer. This probably indicates that *Stowaway* elements are more abundant in the barley genome than *Barfly* or *Tourist*. Some of the bands resulting from TEM-1 amplification could not be scored because of the abundance of bands and their poor resolution. We have evidence that still more loci can be detected by designing primers anchored with an additional base at the 3´ end (data not shown). This reduces the complexity of the banding pattern obtained.

In our study, use of the TEM-1 primer in the IMP approach was more reproducible than observed for the TEM-10 primer. This is most likely due to the fact that the TEM-1 primer was designed based on a consensus of considerably more element sequences (i.e., 44 barley *Stowaway* TIR sequences versus seven barley/wheat *Barfly* TIRs). The reproducibility with the TEM-10 primer may be improved by designing a more-optimal primer based on additional *Barfly* elements, as more are identified or may reflect the influence of element density on the IMP approach. Though reproducibility will most likely improve with better *Barfly* TIR primer design, we do not believe that the number of loci detected will significantly increase. Nevertheless, combining primers based on different elements is likely to provide an even better coverage of the genome. As a general recommendation, we suggest that IMP reproducibility can be optimized by selecting those elements that have a high frequency of occurrence (e.g., *Stowaway* in barley) in DNA sequence databases and avoiding those that are less frequent (e.g., *Barfly* and *Tourist* in barley). Although reproducibility is difficult to compare given the lack of quantitative estimates from other PCR-based genome-mapping protocols, we suggest that the IMP approach using frequently occurring elements is within an acceptable range of reproducibility for a PCR-based technology.

We found that the IMP markers mapped to all seven barley linkage groups, segregated normally, were distributed throughout the genome and were informative in determining genetic relationships between and within *Hordeum* species. Polymorphic bands were also observed in wheat, oat and maize lines, indicating that the IMP approach has applicability in other cereal grasses. Improvement in the level of polymorphism detected in these other species will most likely be obtained by devising primers from endogenous elements within the genomes of the target species or genus. In general, the IMP approach should be applicable to any organism.

Clearly, there is great potential in using genomic anchor sequences for the development of informative mapping and fingerprinting technologies. The exploitation of MITEs in the development of IMP serves as an ex-

cellent example of such a technology. We can expect that more MITEs and other transposable elements will be uncovered as more genomic sequences of genes are reported and as plant genome sequencing projects progress. This new information will certainly hone IMP and IMP-like approaches. In addition, new approaches utilizing several different types of genomic anchor sequences offer the possibility of a virtual cornucopia of mapping and fingerprinting tools. The development of these tools will not only accelerate the construction of detailed genome maps and facilitate molecular fingerprinting but will also have application in linkage studies of qualitative traits and QTLs, marker-assisted selection, chromosome walking, the alignment of cloned genomic fragments for genome sequencing projects, and the tagging and isolation of gene sequences.

Acknowledgements We thank Dr. Stine Tuvesson (Svalöf Weibul AB, Sweden) for kindly providing the barley DH population, Dr. Mark E. Sorrells (Cornell University) for the wheat and oat RIL populations, and Dr. Jörgen Löhde (Svalöf Weibul AB, Sweden) for the barley lines used in the fingerprinting study. We are grateful to Ms. Christiane Ménard for technical assistance. Dr. Benoit Landry (DNA LandMarks, Inc.) is gratefully acknowledged for his support and valuable discussions during the course of this study. This research was funded by a National Science and Engineering Research Council of Canada strategic grant to T.E.B. All experiments described in this study comply with the current laws of Canada. The IMP technique and other MITE-based marker technologies are covered by patent applications jointly owned by McGill University and DNA LandMarks, Inc.

References

Bostein D, White R, Skolnick M, Davis, RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphism. Am J Hum Genet 32:314–331

Broeck van den D, Maes T, Sauer M, Zethof J, Keukeleire PD, D'Hauw M, Montagu MV, Gerats T (1998) Transposon display identifies individual transposable elements in high copy number lines. Plant J 13:121–129

Bureau TE, Wessler SR (1994) *Stowaway*: a new family of inverted repeat elements associated with the genes of both monocotyledonous and dicotyledonous plants. Plant Cell 6:907–916

Bureau TE, Ronald PC, Wessler SR (1996) A computer-based systematic survey reveals the predominance of small inverted-repeat elements in wild-type rice genes. Proc Natl Acad Sci USA 93:8524–8529

Deynze van AE, Dubcovsky J, Gill KS, Nelson JC, Sorrells ME, Dvorak J, Gill BS, Lagudah ES, McCouch SR, Appells R (1995) Molecular-genetic maps for group 1 chromosomes of Triticeae species and their relation to chromosomes in rice and oat. Genome 38:45–59

Ellis THN, Poyser SJ, Knox MR, Vershinin AV, Ambrose MJ (1998) Polymorphism of insertion sites of *Ty1-copia* class retrotransposons and its use for linkage and diversity analysis in pea. Mol Gen Genet 260:9–19

Finnegan DJ (1989) Eukaryotic transposable elements and genome evolution. Trends in Genetics 5:103–107

Flavell AJ, Dunbar E, Anderson R, Pearce SR, Hartley R, Kumar A (1992) *Ty1-copia* group retrotransposons are ubiquitous and heterogeneous in higher plants. Nucleic Acids Res 20: 3639–3644

Flavell AJ, Knox MR, Pearce SR, Ellis THN (1998) Retrotransposon-based insertion polymorphisms (RBIP) for high throughput marker analysis. Plant J 16:643–650

Graner A, Jahoor A, Schondelmaier J, Siedler H, Pillen K, Fischbeck G, Wensel G, Herrmann RG (1991) Construction of an RFLP map of barley. Theor Appl Genet 83:250–256

Heun M, Kennedy AE, Anderson JA, Lapitan NLV, Sorrells ME, Tanksley SD (1991) Construction of a restriction fragment length polymorphism map for barley (*Hordeum vulgare*). Genome 34:437–447

Hirochika H, Hirochika R (1993) *Ty1-copia* group retrotransposons as ubiquitous components of plant genomes. Jpn J Genet 68: 35–46

Kalendar R, Grob T, Regina M, Suoniemi A, Schulman A (1999) IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. Theor Appl Genet 98:704–711

Kleinhofs A, Kilian A, Saghai-Maroof M, Biyashev R, Hayes P, Chen F, Lapitan N, Fenwick A, Blake T, Kanazin V, Ananiev E, Dahleen L, Kudrna D, Bollinger J, Knapp S, Liu B, Sorrells M, Heun M, Franckowiak J, Hoffman D, Skadsen R, Steffenson B (1993) A molecular, isozyme, and morphological map of the barley (*Hordeum vulgare*) genome. Theor Appl Genet 86: 705–712

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln S E, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Lee D, Ellis THN, Turner L, Hellens RP, Cleary WG (1990) A *copia*-like element in *Pisum* demonstrates the uses of dispersed repeated sequences in genetic analysis. Plant Mol Biol 15:707–722

Litt M, Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am J Hum Genet 44:397–401

Mantel NA (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220

Nei M, Li W (1979) Mathematical models for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

Nelson DL, Ledbetter SA, Corbo L, Victoria MF, Ramirez-Solis R, Webster TD, Ledbetter DH, Caskey CT (1989) *Alu* polymerase chain reaction: a method for rapid isolation of human-specific DNA sequences from complex DNA sources. Proc Natl Acad Sci USA 86:6686–6690

O'Donoughue LS, Kianian S, Rayapati J, Penner G, Sorrells ME, Tanksley SD, Phillips R, Rines H, Lee M, Fedak G, Molnar S, Hoffman D, Salas CA, Wu B, Autrique E, Van Deynze A (1995) A molecular linkage map of hexaploid oat. Genome 38: 368–380

Pearce SR, Gill H, Li D, Heslop-Harrison JS, Kumar A, Flavell AJ (1996) The *Ty1-copia* group retrotransposons in *Vicia* species: copy number, sequence heterogeneity and chromosome localisation. Mol Gen Genet 250:305–315

Provan J, Thomas WTB., Forster BP, Powell W (1999) *copia*-SSR: a simple marker technique which can be used on total genomic DNA. Genome 42:363–366

Rohlf FJ (1994) NTSYS-pc numerical taxonomy and multivariate analysis system, version 1.80, Exeter Software, New York

SanMiguel P, Tikhonov A, Jin Y-K, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer P., Edwards KJ, Lee M, Avramova Z, Bennetzen JL (1996) Nested retrotransposons in the intergenic regions of the maize genome. Science 274:765–768

Sinnet D, Deragon J-M, Simard LR, Labuda D (1990) Alumorphs–human DNA polymorphisms detected by polymerase chain reaction using *Alu*-specific primers. Genomics 7: 331–334

Souer E, Quattrocchio F, de Vetten N, Mol J, Koes R (1995) A general method to isolate genes tagged by a high copy number transposable element. Plant J 7:677–685

Suoniemi A, Anamthawat-Jonsson K, Arna T, Schulman AH (1996) Retrotransposon *BARE*–1 is a major, dispersed component of the barley (*Hordeum vulgare* L.) genome. Plant Mol Biol 30:1321–1329

Tautz D (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res 17: 6463–6471

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Voytas DF, Cummings MP, Konieczny A, Ausubel FM, Rodermel SR (1992) *copia*-like retrotransposons are ubiquitous among plants. Proc Natl Acad Sci USA 89:7124–7128

Waugh R, McLean K, Flavell AJ, Pearce SR, Kumar A, Thomas BBT, 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

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388–396

Welsh J, McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res 18:7213–7218

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18: 6531–6535

Zietkiewicz E, Labuda M, Sinnet D, Glorieux FH, Labuda D (1992) Linkage mapping by simultaneous screening of multiple polymorphic loci using *Alu* oligonucleotide-directed PCR. Proc Natl Acad Sci USA 89:8448–8451

Zietkiewicz E, Rafalski A, Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. Genomics 20:176–183