CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Sydney eScholarship

Transposon-mediated gene search: finding a needle in a haystack

Singh J[1], Lam A[2], Blosse, P[2], Bregitzer P[3], Lemaux P[2]
[1] Plant Science Department, McGill University, Ste Anne De Bellevue, QC, Canada
[2]Department of Plant and Microbial Biology University of California, Berkeley CA, 94720 USA, [3]National Small Grains Germplasm Research Facility, USDA ARS, Aberdeen, ID, USA, 83210

INTRODUCTION

Characterization of mutant alleles provides valuable insights into the genetics of normal physiological functions. Complementing the study of mutations generated randomly with chemicals or irradiation, technical advances in the field of genomics and the explosion of available gene sequences provide tools that lead to increasingly detailed characterizations of individual genes and their physiological roles.

The transposon-based mutagenesis approach has certain advantages over the other approaches for determining gene function. Transposable elements can be mobilized or immobilized on demand and the approach requires only a few initial transformants to generate large numbers of plants carrying transposed elements (transposants) at different locations. In addition, transposons can have bias for genic regions (1) – important in large genome species – and this can lead to the identification of regulatory sequences as well as introns. Introduction of the maize *Ac-Ds* transposable element system as a transposon tagging tool into heterologous species (2, 3) offers unprecedented opportunities to link genes with function by creating and characterizing mutant alleles.

The early work that enabled the use for functional genomics of the maize transposable elements, *Ac/Ds*, was first reported by McClintock (4), when she discovered that a locus of chromosome breakage, *Dissociation* (*Ds*), could move from one position to another in the maize (*Zea mays* L.) genome in the presence of another factor, *Activator* (*Ac*).

The system developed in barley relies on two parts of the maize *Ac/Ds* system, introduced into separate plants via particle bombardment into the cultivar Golden Promise. One component is a transposition-competent *Ds*-bordered, *ubiquitin*-driven *bar* gene (*Ds-bar)*; the other, a transposition-incompetent *Ac transposase* (*AcTpase*) gene driven by either the *ubiquitin* or the native *Ac* promoter. *Ds-bar* is activated to transpose by crossing the two plants. McElroy et al. (5) and Koprek et al. (6) showed transient and stable functionality, respectively, of this system in barley via *AcTpase*-mediated transpositon of *Ds-bar* cassettes. Lines with single, unique *Ds-bar* insertions (TNPs) were identified (7) and mapped using a sequence-based approach by Cooper et al. (8). Furthermore, this approach provided additional evidence that the *Ac/Ds* system functions in barley as it does in maize: there was a higher frequency of *cis*- versus *trans*-transposition; transposition occurred primarily to coding regions (frequently with homology to known genes); and transposition was accompanied by an 8 bp duplication of the sequence into which insertion occurred. Also typical of this system in its native maize, the terminal inverted repeat (TIR) of the *Ds* element was in some cases altered and such changes to the TIR appear to interfere with further transposition (7). Molecular studies and flanking sequence characterization of TNP lines reported in this paper validate *Ds* tendency to transpose into genic regions.

MATERIALS AND METHODS

Development and characterization of TNP lines

Development of the primary TNPs was accomplished by crossing transgenic barley lines expressing *Ac*TPase and barley lines containing a transposition-competent *Ds-bar* element (6). T_2 plants carrying single- or multi-copy *Ds-bar* elements were crossed with plants expressing *Ac*TPase to remobilized *Ds*. Resulting F_1 plants were selfed and F_2 populations analyzed by DNA hybridization for evidence of transposition events, based on differing band migrations (7). Lines with single, transposed *Ds-bar* copies and no *AcTPase* were made homozygous and designated secondary TNPs. Remobilization of *Ds* elements in five secondary TNPs, *i.e,* TNP-3, -13, -24, -30 and -41, was accomplished by crossing the lines with an *Ac*TPase-expressing line; F_2 populations obtained from those crosses were screened for new tertiary transposition events. Three tertiary TNPs, which had *AcTPase*, were advanced to F_3 and quaternary transpositions identified.

Generation and characterization of *Ds* flanking sequences

Plants with single- or low-copy, primary and secondary, tertiary, or quaternary *Ds* transpositions, identified as described above, were analyzed to obtain 5'- and/or 3'-flanking sequences by iPCR and TAIL-PCR as described (8). This sequence information was the basis for determining orientation of *Ds* insertions and structure of insertion sites, including TIRs and 8 bp duplications, and to characterize the genomic regions into which transposition occurred. Bioinformatic analysis was conducted as described by Singh et al., (7)

RESULTS

Development and Re-activation of *Ds* insertion lines

A total of 160 barley *Ds* insertion lines were generated in this study, which includes primary, secondary, tertiary and quaternary transposition events. Re-activation of *Ds*

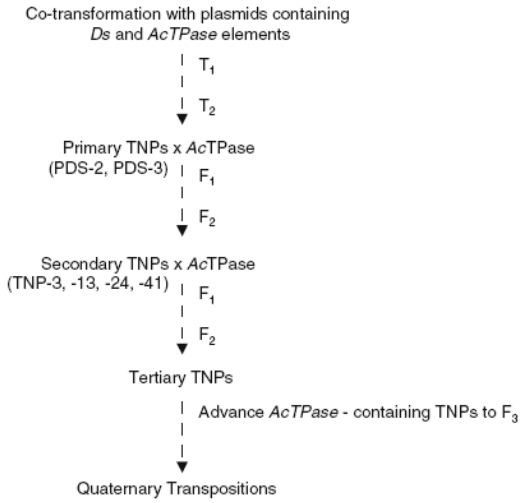

Fig. 1 Summary of *Ds* transposition and reactivation scheme

from primary to secondary and secondary to tertiary insertions was achieved by crossing homozygous *Ds* lines devoid of *AcTPase* with *Ac*TPase-expressing lines. Quaternary TNPs were generated after self-pollination of tertiary TNPs, which contained both *Ds* and *Ac*TPase. The scheme of development and re-activation of TNPs is presented in **Figure 1**.

Analysis of Ds insertion sites

Flanking sequences were obtained primarily by iPCR on >100 TNPs. The sequences from 142 lines were examined in detail to determine the nature of the regions into which *Ds* tended to insert and to determine the status of TIRs and 8 bp duplications. Analysis of flanking sequences from a total of 81 lines indicated that the TIRs of 16% of TNPs were imperfect, suggesting that aberrant insertion/excision of *Ds* occurs frequently (7). No transposition was detected among progeny of crosses using those TNPs and *Ac*TPase-expressing lines as parents.

To determine the nature of *Ds* insertion sites in the barley genome, analyses of *Ds* flanking sequences were conducted, using BLAST searches of public databases, such as those at harvEST, Gramene, TIGR and NCBI websites, looking for similarities to known proteins, genomic sequences and/or ESTs. A database of these *Ds*-flanking sequences from barley is available on the GrainGenes website (http://wheat.pw.usda.gov/BarleyTNP). Annotations of these searches are summarized in *Table 1*. Based on BLAST searches, 75 % (61 of 81) of TNP flanking sequences match ESTs or characterized gene sequences. Analysis using a gene prediction program identified 87.6% (71/81) of *Ds* insertions as being in genic regions (**Table 1**), demonstrating high-frequency *Ds* insertion either into or in close proximity to genes. Analyzing these flanking sequences with gene prediction programs, like GenScan, 87.6% (71/81) are in unique genic regions, for example, MLA1, wall-associated kinases, ubiquitin-conjugating enzyme, ATP-binding transporter, terpene synthase, ankyrin1-like protein and cytochrome P450. Similar observations were obtained in maize, where 75% of *Ac* insertions were in genic regions (1), and in rice, where 72% of *Ds* transpositions were in genic regions (9). A high frequency of *Ds* insertions into exons or close to exons is important for gene isolation and reverse genetics studies in barley and wheat, where a high percentage of the genome is composed of repetitive regions and transposable elements (10). BAC hybridization evidence indicates these *DsT* loci do not reside in repetitive DNA sequences. In general, of the regions analyzed, most *Ds* insertions appear to be in single- or low-copy genes, as evidenced by gene identification analyses of flanking DNA sequence in databases and the numbers of BAC clones hybridizing to TNP-specific probes. For mapping of Ds insertions, mainly sequence based strategy was used (data not shown). Corresponding regions of DNA sequence flanking the *Ds* insertions from the parents of the mapping population were sequenced and polymorphisms identified to use in mapping. Sequences flanking 35 *DsT*'s were also physically mapped in wheat aneuploid and deletion lines (Singh and Randhawa, unpublished). High frequency (83%) of flanking sequences mapped on all seven homoeologous groups and, more significantly, mapped to gene-rich regions of wheat.

CONCLUSIONS

Mapping and bioinformatics analysis of *Ds* flanking sequences indicate that the vast majority of *Ds* insertions (88%) are in genic regions. The bias of *Ds* towards genes is highly valuable for large genome cereals and *Ac/Ds* can be used as gene search tool. Data on multigenerational *Ds* re-activation is critical for localized saturation mutagenesis efforts, including the re-activation needed for "transposon walking", the sequential re-activation of *Ds* that can be used to identify members of clustered gene families. It is more than adequate to conduct saturation mutagenesis in barley and wheat of linked regions, setting the stage for reactivation tagging of loci proximal to mapped TNPs.

Table 1: Status of *Ds* insertion sites and annotation of flanking sequences

Ds Line	Flanking sequence Annotation, E-Value	Ds Line	Flanking sequence Annotation, E-Value
PDS 2	Barley EST HA08116u, 2e-04	TNP 54	Wheat EST wpa1c.pk01, 1.5e-07
PDS 3	Barley EST BM441822.2, 9.9e-08	TNP 64	RNA Binding Protein, 9.5e-58
TNP 1	Maize EST 3529_1_87_1_D09.y_1, 1e-07	TNP 67	ABC Transporter, 5.1e-148
TNP 3	Wall-Associated Kinase, 1.1e-34	TNP 69	Expressed Protein, 2.6e-06
TNP 6	Barley EST AV946044.1, 1.8e-09	TNP 70	ABC Transporter, 5.1e-148
TNP 11	Wall-Associated Kinase, 6.2e-73	TNP 71	Barley EST CA028167.1
TNP 12	Wheat EST CK151843.1 3.1e-17	TNP 72	Wheat EST TAS004.G09R990617, 3.5e-35
TNP 13	Wheat EST wpa1c.pk011.e7, 1.5e-07	TNP 74	Putative MLA6 Protein, 3e-19
TNP 18	F-Box Putative, 3.2e-10	TNP 79	Wheat EST BF200383.1, 2.7e-06
TNP 19	Barley EST AJ474918.1, 8.8e-19	TNP 80	Pherophorin- like Protein, 5.4e-19
TNP 22	Terpene Synthase, 2.4e-25	TNP 81	Barley EST CX632786.1, 1.5e-182
TNP 23	Wheat EST wip1c.pk005.a8, 4.5e-27	TNP 82	Putative Cytochrome p450, 2.4e-09
TNP 24	RNA Binding Protein, 9.5e-58	TNP 83	ABC Transporter, 1.4e-81
TNP 25	P450, 6.9e-52	TNP 90	Barley EST BG310371.1, 4.4e-37
TNP 26	HVSMEc0013O19f, 2.4e-35	TNP 92	*H. vulgare Mla* Locus, 5e-46
TNP 29	Wheat EST wip1c.pk005.a8, 4.5e-27	TNP 93	Wheat EST FGAS022629, 9e-42
TNP 30	Expressed Protein, 9.9e-10	TNP 94	Protein Kinase Domain, 8.8e-54
TNP 32	Phosphate Transporter *HvPTv* gene, e-134	TNP 95	Wheat EST BF200383.1, 2.7e-06
TNP 41	Ubiquitin Conjugating Enzyme, 2e-54	TNP 97	Wall-Associated Kinase, 1.1e-34
TNP 52	Wheat EST CK151843.1, 3.1e-17	TNP 98	RNA Binding Protein, 9.5e-58
TNP 53	Wheat EST TAS004.G09R990617, 3.3e-11	TNP 100	Barley EST HVSMEm0023O17f, 2.9e-07

REFERENCES

1. Cowperthwaite M, Park W, Xu Z, Yan X, Maurais SC, Dooner HK (2002) Plant Cell 14:713-726.
2. Parinov S, Sevugan M, Ye D, Yang WC, Kumaran M, Sundaresan V (1999). Plant Cell 11:2263-2270.
3. Raina S, Mahalingam R, Chen F, Federoff N (2002) Plant Mol. Biol. 50:93-110.
4. McClintock, B (1949). Mutable loci in maize. Carnegie Inst. Wash. Year Book 48:142-154.
5. McElroy D, Louwerse JD, McElroy SM, Lemaux PG (1997) Plant J. 11:157-165.
6. Koprek T, McElroy D, Louwerse J, Williams-Carrier R, Lemaux PG (2000) Plant J. 24:253-263.
7. Singh J, Zhang S, Chen C, Cooper L, Bregitzer P, Sturbaum AK, Hayes PM, Lemaux PG (2006) Plant Mol. Biol. 62 : 937-950.
8. Cooper LD, Marquez-Cedillo L, Singh J, Sturbaum AK, Zhang S, Edwards V, Johnson K, Kleinhofs A, Rangel S, Carollo V, Bregitzer P, Lemaux PG, Hayes PM (2004) Mol. Gen. Genomics 272:181-193.
9. Kolesnik T, Szeverenyi I, Bachmann D, Kumar CS, Jiang S, Ramamoorthy R, Cai M, Ma ZG, Sundaresan V, Ramachandran S (2004) Plant J. 37:301-314.
10. Wicker T, Zimmermann W, Perovic D, Paterson AH, Ganal M, Graner A, Stein N (2005) Plant J. 41:187-194.