Applications of retrotransposons as genetic tools in plant biology

Amar Kumar and Hirohiko Hirochika

Retrotransposons are mobile genetic elements that accomplish transposition via an RNA intermediate that is reverse transcribed before integration into a new location within the host genome. They are ubiquitous in eukaryotic organisms and constitute a major portion of the nuclear genome (often more than half of the total DNA) in plants. Furthermore, they are dispersed as interspersed repetitive sequences throughout most of the length of all host chromosomes. These unique properties of retrotransposons have been exploited as genetic tools for plant genome analysis. Major applications are in determining phylogeny and genetic diversity and in the functional analyses of genes in plants. Here, recent advances in molecular markers, gene tagging and functional genomics technologies using plant retrotransposons are described.

The major challenge in the era of genome-wide sequencing research is to find ways to use the vast amount of sequence data to improve our understanding about the biology of organisms. In this context, genetic tools such as molecular markers and insertional mutagenesis (functional analysis of genes) methodologies are playing an important role.

Retrotransposons (RTNs) are mobile genetic elements that are ubiquitous throughout the plant kingdom and constitute a major portion of the nuclear genomes (often more than half of the total DNA) of plants (reviewed in Ref. 1). Furthermore, they are distributed as interspersed repetitive sequences almost throughout the length of all host chromosomes. RTNs have achieved this unique situation (i.e. high copy number and genome-wide dispersion) because of their inherent replicative mode of transposition within the host genome. They transpose via an RNA intermediate that is reverse transcribed into extrachromosomal DNA and inserted into the genome by the encoded reverse transcriptase RNaseH and integrase enzymes. This replicative mode of transposition also allows RTNs to generate genetic diversity by altering the size and organization of the host genomes and by generating insertional mutations of genes. Moreover, transposable elements consisting of both Class I (RNA-based RTNs) and Class II (DNA-based transposons) play the major role in maintaining genome plasticity in plants. This article highlights recent advances in the field of molecular markers, gene tagging and functional genomics technologies using RTNs as genetic tools in plants.

Molecular marker technology

Molecular marker technology is playing a vital role in plant biology, including DNA fingerprinting, genetic linkage mapping and phylogenetic relationship studies, and in molecular breeding (reviewed in Refs 2,3). Several DNA-based marker technologies have been developed to detect polymorphisms by assaying subsets of the total amount of DNA in a genome.

We describe a newly developed RTN-based marker technology in plants that is based on fundamentally different biological processes to the commonly used marker technologies. In the case of RTN-based markers, polymorphisms are generated by a unique biological process of retrotransposition, resulting in insertions of RTNs into new sites without losing the parental copies. Furthermore, the consequences of retrotransposition range from the alteration of a few hundred base pairs to a few kilobases at the site of insertions. By contrast, molecular marker technology based on single nucleotide polymorphisms (SNPs) and simple sequence repeats (SSRs) are based on random small-scale changes (i.e. one or a few tens of nucleotides). Moreover, most RTN insertions are reversible, whereas SNP and SSR polymorphisms are irreversible, which limits their use for determining parental lineage data in any study of phylogenetic relationships. Other advantages to using RTN sequences over other molecular markers are that they are ubiquitous, are present in high copy numbers and as highly heterogeneous populations, are widely dispersed within euchromatic regions of chromosomes and show insertional polymorphism both within and between species in plants.

Several types of RTN-based markers have been developed (Fig. 1). One of the most useful RTN-based marker systems in revealing large numbers of highly polymorphic markers is sequence-specific amplification polymorphism (SSAP). SSAP is a multiplex amplified fragment length polymorphism (AFLP)-like technique that displays individual RTN insertions as bands on a sequencing gel. Moreover, SSAP-based markers appear to be better for estimating phylogenetic relationships in plants compared with the conventional AFLP-based makers because one of the primers is based on specific RTN sequences (Fig. 1).

Recently, heterogeneity of the internal structure within a family of the PDR1 elements has been exploited as SSAP markers for studying genetic diversity in the genus *Pisum* using the retrotransposon internal variation polymorphisms (RIVP) method (Fig. 1). Several RTN-based marker systems use a simple PCR method to generate polymorphic products. For example, a PCR method has also been used to amplify internal sequences of
A PCR-based approach has also been developed to detect individual RTN insertions using primers derived from the element and its flanking DNA\(^{11,12}\). This retrotransposon-based insertion polymorphism (RBIP) method is a co-dominant marker system, where the different allelic states (i.e. presence and absence of the RTN insertion) at a locus can be revealed. The main advantage of this RBIP method is that it has the potential to be fully automated for high-throughput marker analysis.

Recent development of a rapid technique for isolation of LTR RTN sequences\(^{13}\), together with easy access to transposon sequences from the vast amount of genomic sequence information\(^{14}\), have extended applications of RTN-based marker technology in plants. Indeed, phylogenetic and biodiversity studies are being performed in several crops (barley, broad beans, maize, pea, potato, tomato, rape, rice, soybean, sugar beet and tobacco) and their related wild species using these RTN-based markers.

**Biodiversity and phylogeny**

Biodiversity and phylogenetic studies are essential for preserving both land races and their related wild species, which often are disappearing rapidly\(^2\). RTN-based markers are especially suitable for studying phylogenetic relationships and genetic diversity within and between species. For example, active members of the RTN family produce new insertions in the genome, leading to polymorphism. The new insertions can then be detected and used to establish the temporal sequence of insertion events, helping to determine phylogenies. These genetic properties have recently been exploited to study biodiversity and phylogeny in the genera *Brassica*\(^{12}\), *Hordeum*\(^{5,8,10,15}\), *Oryza*\(^{16,17}\) and *Pisum*\(^{6,18,19}\).

Short interspersed repetitive elements (SINEs)-based markers have been extremely useful in phylogeny and genetic diversity studies in eukaryotic organisms, including humans and whale\(^{20}\). This is mainly because SINE insertions are independent and irreversible in spite of their repetitive nature. In plants, the rice *p-SINE* element members have been used to evaluate relationships among the Asian cultivated rice and their related species in the *Oryza* genus\(^{18}\) and for classification of rice strains with AA genomes\(^{17}\). Similarly, the *S*\(_{Br}\) SINE elements have been also used to study genetic diversity in *Cruciferae* species and to determine the extent to which introgression has occurred between cultivated and wild species in *Brassica*\(^{12}\).

Another important recent application of RTN-based markers has been to show that genomic diversity within wild *Hordeum spontaneum* populations is likely to be generated by both active amplification and losses of the *BARE1* family of RTNs (Ref. 10). More interestingly, a correlation was found between the *BARE1* copy number and genome size in wild barley inhabiting the sharply differing microclimates in Evolution Canyon near Mount...
Carmel, Israel. This indicates a molecular mechanism linking the influence of environmental factors to the generation of genomic diversity in plants.

A multi-RTN approach has been used recently to estimate phylogenetic relationships within and between species in legume19 and cereal15 plants (Fig. 2). This approach is highly informative because each element has its own transpositional history (i.e. the time and rate of transpositions). In this recent study, multi-RTN-based markers were more polymorphic and informative than the conventional AFLP-based markers6,19. For example, PDR1 appears to be almost inactive in Pisum abyssinicum accessions, but there is some internal divergence of the element as judged from the PDR1 gag trees (Fig. 2). Furthermore, there is little variation in P. abyssinicum as judged by either the gag tree or the insertion site trees, based on the PDR1 and Tps19 elements6. By contrast, Tps12a and Tps12b have been more active within P. abyssinicum accessions19. Within P. abyssinicum the variation for the gag tree is probably similar to the variation as judged by the insertion site polymorphism. What is different is that the distance between P. sativum and P. abyssinicum is much shorter for the gag tree than the insertion site tree, indicating that PDR1 is absent in P. abyssinicum and that all the family variants in P. abyssinicum are present in P. sativum. A few new variants have arisen in the P. sativum lineage, and in this lineage transposition has continued generating many insertion sites that are absent from P. abyssinicum6,19 (Fig. 2).

There are small differences between the four element insertion site trees: presumably these reflect differences in transposition history, as discussed earlier19. It appears that the advantage of the RTN-based markers is that they partition the variance in the observable patterns of band differences in a phylogenetically meaningful way. The marker system allows partitioning to be exploited with RTNs in a manner not possible with other markers. Furthermore, the combination of both internal (i.e. RIVP) and insertion site (i.e. SSAP) data was complementary and provided ‘an extended depth of phylogenetic relationships’ within Pisum species6,19.

**Frequency and timing of retrotransposition**

The SSAP method also provides an efficient technique for evaluating retrotransposition history and behaviour in natural and manipulated plant populations. It has been used to study the insertional polymorphism of the Ty1-copia group RTNs in plants4,5,7,10,15,18,19. These studies have shown that BARE-1 insertions, for example, are highly polymorphic in barley, wheat, rye and oat, suggesting that the transpositional activity of this RTN has persisted for millions of years in several cereal species18. A similar study has shown that a few RTNs have been transpositionally active in the recent past in several Pisum species13,18,19 and in diploid Avena species21.

Other RTN-based marker systems such as RBIP, IRAP and REMAP should also be useful to evaluate the transposition history of RTNs in plants. In addition, RTN-based markers, especially SSAP and RBIP, provide an ideal tool for evaluating to what extent RTN-induced genomic variability can be generated under abiotic (e.g. tissue cultures, environmental stresses) and biotic (e.g. pathogen infections) conditions in plants.

**Linkage mapping and markers linked to agronomic traits**

Molecular breeding is dependent on good genetic linkage mapping data and markers linked to agronomic traits. SSAP markers have also been used to construct linkage maps in barley, oat and pea using the LTR-specific sequences of the BARE-1 and PDR1 RTNs, respectively4,5,18,21. Interestingly, both BARE-1 and PDR1 markers were found to be better for constructing linkage genetic maps in barley and pea, respectively4,5,18.

With respect to the production of linkage maps for agronomically important traits, several quantitative trait loci (QTLs) have been mapped onto seven linkage groups of barley using BARE-1–SSAP analysis. For example, QTLs for grain hardness and β-glucan content were mapped onto chromosome 1H; β-glucan content, wort maltose and grain nitrogen content onto chromosome 2H; and height, ferment quality and resistance to leaf rust onto chromosome 5H (B. Thomas et al., unpublished). In barley, other RTN-based marker systems such as IRAP and REMAP in conjunction with SSAP have been used to map a resistance locus to Net blotch, which is caused by the fungus Pyrenophora teres, onto barley chromosome 3H (Ref. 22). Recently, PDR1–SSAP analysis has been used to clarify the complex inheritance patterns of pea blight resistance, which is caused by Pseudomonas
Table 1. Retrotransposon-based markers linked to useful genes

<table>
<thead>
<tr>
<th>RTN</th>
<th>Plant</th>
<th>Linkage</th>
<th>Property of the gene</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>BARE-1</td>
<td>Barley</td>
<td>0.28 cM from the Mla locus</td>
<td>Mla, a member of a multigene family that confers resistance to powdery mildew, Erysiphe graminis</td>
<td>42</td>
</tr>
<tr>
<td>Tgm</td>
<td>Soybean</td>
<td>0.2 cM from the Rps 1-k locus</td>
<td>Rps 1-k, a member of a multigene family that confers resistance to Phytophthora sojae</td>
<td>43</td>
</tr>
<tr>
<td>Tnd1</td>
<td>Tobacco</td>
<td>Linked in coupling to an R-gene</td>
<td>A resistance gene to Black root rot, Chalara elegans</td>
<td>44</td>
</tr>
<tr>
<td>Tlp1</td>
<td>Tomato</td>
<td>Linked in coupling to the Hero locus</td>
<td>Hero, a resistance gene to potato cyst nematodes, Globodera rostochiensis</td>
<td>a</td>
</tr>
<tr>
<td>Tos3</td>
<td>Rice</td>
<td>0.4 cM distance to the Pit gene</td>
<td>Pit, a resistance gene to Magnaporthe grisea</td>
<td>b</td>
</tr>
<tr>
<td>PDR1</td>
<td>Pea</td>
<td>~0.5 cM from sym 19</td>
<td>sym 19, a gene likely to be involved in the early stages of nodule formation in pea</td>
<td>45</td>
</tr>
</tbody>
</table>

*K. Ernst et al., unpublished.

*H. Hirochika et al., unpublished.

*syringae pv. pisi*, in crossed populations between *P. abyssinicum* (resistant line) and *P. sativum* (susceptible line). The inheritance and molecular data indicated that one major recessive gene and a set of modifier genes controlled resistance. Furthermore, the presence of the three *P. abyssinicum* bands and the absence of the three *P. sativum* bands were always associated with resistance, providing the opportunity for exploiting these markers in pea breeding (M. Elvira-Recuenco et al., unpublished). Because of the extensive genomic coverage of RTNs, the most immediate applications of RTN-based markers will be in targeted high-density mapping of important genomic regions and as a fingerprinting technique. Agronomically important genes that are linked to RTN-based markers are listed in Table 1.

Finally, a wider application of RTN-based markers, as with any other marker systems, will depend upon factors such as easy accessibility, wider genomic coverage, dispersion within euchromatic and heterochromatic regions, and a high level of informativeness. Moreover, RTN sequences are now easily accessible from any plant genome, widely dispersed throughout the length of all chromosomes, commonly found in the euchromatic region, and as molecular markers are highly informative. Hence, RTN-based markers compare well against most other types of existing marker systems.

**Gene tagging and functional analysis of genes**

Screening of mutants induced by insertion elements, followed by isolation of causative genes using these sequences as molecular tags (‘gene tagging’) is one of the most important methods for plant molecular biology. Many genes involved in important plant-specific processes have been cloned using this method. Reflecting the recent rapid progress of sequencing projects, a systematic method for assigning biological functions to sequenced genes becomes more important. For this reverse genetic analysis as well as classical gene tagging, large populations of tagged insertional mutants are being produced and used in model plants such as *Arabidopsis* and maize (reviewed in Refs 23–25). Insertion elements used in these systems are class II transposable elements, such as *Ac, En/Spm* and *Mutator* of maize, and T-DNA. Provided that clones of active RTNs are available, those elements can also be used for gene tagging, as has been shown in yeast and *Drosophila*. Based on unique features of plant RTNs, a PCR-based method for cloning active RTNs has been developed, which can make the tagging system available to many plant species. Indeed, RTN-based tagging and reverse genetic systems are being developed in another important model plant, rice, whose genomic sequencing project has been initiated.

Plant RTNs have unique features that make them an ideal genetic tool for gene tagging in plants, giving them the advantages over traditional insertional elements:

- **RTN-mediated insertion mutations are stable** because they transpose via a replicative mode.
- **Transposition target sites are unlinked** with the site of the original copy, so that it is easy to generate a large collection of random insertions for saturation mutagenesis.
- **Transposition can be easily and strictly regulated**, activated by tissue culture and inactivated by regeneration.
- **They are highly mutagenic as some of them normally transpose into gene-rich regions**.
- **The original copy number can be low**, so that it is easy to identify the transposed copy responsible for the specific mutation.
- **Endogenous active RTNs can be isolated by using the simple reverse transcription-PCR (RT-PCR) method**.

However, some of the features might also be disadvantageous. For example, transposition via a replicative mode means that no revertants can be obtained, although revertants are useful for confirming that a gene is tagged and can be obtained from the mutants induced by class II elements.

**Active RTNs**

Although RTNs are ubiquitous and abundant in plants, especially in large genomes, only a small fraction of them seem to be active. Considering the frequency of transposition observed in the original or heterologous hosts, three Ty1-copia RTNs, Tnt1 and Tto1 in tobacco

http://plants.trends.com
Table 2. A list of tagged and functionally analysed genes

<table>
<thead>
<tr>
<th>Mutant phenotype</th>
<th>Mutated genes</th>
<th>Strategy</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dwarf (gibberellin-deficient)</td>
<td>ent-Kaurene synthase</td>
<td>Forward</td>
<td>a</td>
</tr>
<tr>
<td>Semidwarf (brassinolide-insensitive)</td>
<td>No homology</td>
<td>Forward</td>
<td>b</td>
</tr>
<tr>
<td>Viviparous (abscisic acid-deficient)</td>
<td>Zeaxanthin epoxidase</td>
<td>Forward</td>
<td>c</td>
</tr>
<tr>
<td>Pale green</td>
<td>TatChomologue</td>
<td>Forward</td>
<td>c</td>
</tr>
<tr>
<td>Brittle</td>
<td>Cellulose synthase (OsCesA7)</td>
<td>Forward</td>
<td>d</td>
</tr>
<tr>
<td>Drooping leaf</td>
<td>YABBY family</td>
<td>Forward</td>
<td>e</td>
</tr>
<tr>
<td>Deficient in internode elongation</td>
<td>Homeobox gene (OSH15)</td>
<td>Reverse</td>
<td>39</td>
</tr>
<tr>
<td>Deficient in lateral root elongation</td>
<td>Ring finger protein</td>
<td>Reverse</td>
<td>f</td>
</tr>
<tr>
<td>Altered photomorphogenesis unique to rice</td>
<td>Phytchrome A (phyA)</td>
<td>Reverse</td>
<td>g</td>
</tr>
<tr>
<td>Dwarf (gibberellin-deficient)</td>
<td>P450</td>
<td>Reverse</td>
<td>c</td>
</tr>
<tr>
<td>Semidwarf</td>
<td>ERECTA-like receptor kinase</td>
<td>Reverse</td>
<td>c</td>
</tr>
</tbody>
</table>

| a | S. Takeda et al., unpublished. |
| b | M. Yamazaki et al., unpublished. |
| c | K. Agrawal et al., unpublished. |
| d | K. Tanaka et al., unpublished. |
| e | H. Hirano et al., unpublished. |
| f | T. Lu et al., unpublished. |
| g | M. Takano et al., unpublished. |

and Ttos17 in rice, seem suitable for tagging. Tnt1 has been isolated as an insertion in the nitrate reductase (nix) gene, in which mutations were induced by culturing protoplasts. The frequency of nix mutations induced by Tnt1 is relatively high (1–3 × 10⁻⁵) and comparable to that of mutations (10⁻⁴–10⁻⁶ per locus per chromosome) induced by maize transposable elements used for tagging in maize. However, tagging with Tnt1 is difficult in tobacco because of its allotetraploid genome. It might be feasible in the diploid tobacco (Nicotiana plumbaginifolia), in which active Tnt1 is also present. By assuming that transcription of only active RTNs is activated under stress conditions, Tto1 has been isolated using PCR, with cDNA prepared from cultured cells as the template. Both Tnt1 and Tto1 can be used in Arabidopsis as discussed later. The rice RTN Tos17 has also been isolated using the RT-PCR method and shown to be active only during tissue culture. Furthermore, Tnt1 can be isolated by the RT-PCR method, indicating the general applicability of this method. All three active RTNs are regulated mainly at the transcriptional level. Interestingly, transcription of Tnt1 and Tto1 was induced in response to diverse biotic and abiotic stresses.

RTNs as insertional mutagens

The transposition of Tto1 introduced into Arabidopsis can be activated by tissue culture as in tobacco. On average, 3.2 transposed copies of Tto1 were observed in plants regenerated from explants cultured for only one week. Between five and 30 transposed Tos17 copies were detected in all the plants regenerated from three-to-nine-month-old cultures, whereas the original copy number of Tos17 in the japonica variety is two. These results show that the frequency of transposition is high enough for mutagenesis. However, for efficient mutagenesis, the specificity of integration target sites as well as the frequency of transposition must be considered. Commonly used class II elements, such as Ac/Ds and Spm elements, are known to transpose into linked sites. Thus, the original transposon must be placed near the target gene for efficient tagging. Given that one of the important goals of insertional mutagenesis is saturation mutagenesis, this feature must be a drawback. Considering the mechanism of transposition, RTNs are expected to transpose into unlinked sites. This was shown with Tto1 (Ref. 34), Tos17 (Ref. 35) and Tnt1 (Ref. 36; H. Lucas et al., pers. commun.). Another important feature determining target specificity is the selection of specific loci in the genome. Extensive studies on yeast RTNs have shown that all five RTNs have a strong bias for integration target sites. For example, the target site of Ty3 is restricted to the promoter region of genes transcribed by RNA polymerase III, whereas Ty5 prefers silenced chromatin regions. All the available data on target sites of recent transposition of Tto1, Tnt1 and Tos17 indicate that plant RTNs prefer low-copy, gene-rich regions. This preference is especially important in genomes rich in non-genic sequences, such as repetitive sequences which, for example, account for >80% of the rice genome. The Mutator element shows a similar preference to gene-rich regions, which can be a major factor for the high frequency of induced mutations in maize.

Forward genetics

Cloning of genes by gene tagging has succeeded in rice using Tos17. Several important genes have been cloned by this method (Table 2). The Tos17 copy causing the specific mutation can be identified by a simple DNA gel-blot analysis and the causative gene can be isolated by using inverse-PCR (IPCR) or thermal asymmetric interlaced-PCR (TAIL-PCR). A straightforward strategy to confirm that the cloned genes are causative genes is to use a complementation test. Another important strategy is the analysis of allelic mutants isolated by using the PCR-screening method. Tagging was confirmed in several genes (H. Hirochika et al., unpublished). Although gene tagging with Tos17 is powerful for cloning important genes, one major problem has been noted: the low efficiency of tagging (5–10% (Ref. 29)). This indicates that tissue culture-induced mechanisms other than Tos17 insertions cause a high frequency of untagged mutations. To increase the tagging efficiency, it would be possible to avoid tissue culture-induced mutations by modifying the promoter of RTNs or by finding new conditions under which RTNs can be activated. Considering the frequency of transposition and the target site specificity, Tto1 and Tnt1 should also be used for cloning genes.

Reverse genetics

Two strategies are employed to screen mutants for reverse genetic studies (Fig. 3). One is the PCR-
Review

Pools giving a PCR product (pool A, pool 2 and pool 0). Screening. The mutant individual can be addressed on the matrix by a combination of X, Y and Z DNA sequences. Another strategy is PCR-screening of mutants. Insertions in a gene of interest (gene A) are identified by PCR with gene A-specific and insertion-specific primers using DNA samples sequences are catalogued in databases. Another strategy is PCR-screening of mutants. Insertions in the flanking sequences. To carry out large-scale sequencing of the genomic DNA flanking sequences, a systematic approach to find suitable for reverse genetics, it might not be efficient approach for reverse genetics, it might not be suitable for the analysis of many genes. Considering large-scale functional genomics in the post-sequencing era, a systematic approach to find mutants for a large number of genes is needed. One of the possible approaches is the cataloguing of insertion mutants by sequencing the genomic DNA flanking sequence insertions\(^{21}\), using IPCR or TAIL to amplify the flanking sequences. To carry out large-scale sequencing of the Tos17-flanking sequences, TAIL- and suppression-PCR were adopted. By combining these two PCR methods, ~95% of flanking sequences have been amplified. As of September 2000, more than 20 000 flanking sequences from 2387 lines have

Fig. 3. Identification of RTN insertions for reverse genetics. (a) Two PCR-based strategies to identify insertions in a gene of interest (gene A). One strategy is random sequencing of insertion flanking sequences. Flanking sequences are amplified from DNA samples prepared from every individual by thermal asymmetric interlaced-, inverse- and suppression-PCR, and directly sequenced. These sequences are catalogued in databases. Another strategy is PCR-screening of mutants. Insertions in gene A are identified by PCR with gene A-specific and insertion-specific primers using DNA samples prepared from a pool of mutant lines. The mutant individual can be identified within a pool by using a three-dimensional pooling system. (b) Three-dimensional pooling of DNAs of insertion mutant lines. A population of 960 lines was arrayed in a 96-well format. Seeds were pooled in a three-dimensional matrix, and DNA was extracted from shoots derived from each pool of seeds and subjected to PCR-screening. The mutant individual can be addressed on the matrix by a combination of X, Y and Z DNA pools giving a PCR product (pool A, pool 2 and pool 0).

Screening of mutants, the other is the random sequencing of insertion flanking sequences.

The PCR-screening method was first developed in Drosophila\(^{21}\) and has since been used in Caenorhabditis elegans, petunia, maize and Arabidopsis (reviewed in Ref. 38). The principle of PCR-screening of mutants is as follows. Four combinations of two gene-specific primers and two insertion-specific primers in forward and reverse orientations are used in separate reactions to detect insertions in a gene of interest, irrespective of their position and orientation. Because PCR is sensitive, one insertion in a gene of interest can be identified among a pool of thousands of mutants. To identify mutant individuals, two- or three-dimensional DNA-pooling systems are generally adopted. By using a two-dimensional pooling system, a mutant of the homeobox gene (OSH15) was found in a small population consisting of 529 Tos17 induced mutant lines\(^{26}\). Based on the phenotype of the mutant, OSH15 was shown to be involved in the development of internodes. A total of 11 809 lines carrying 84 975 Tos17 insertions were arrayed in a 96-well format and pooled in a three-dimensional matrix (Fig. 3b). This mutant population has been screened for mutants of 47 genes and mutants of 15 genes identified (H. Hirochika et al., unpublished). These include mutants of genes involved in signal transduction and genes for transcription factors. Functionally analysed genes are listed in Table 2. To make this system generally applicable to any gene of interest, mutations induced by Tos17 insertion must be saturated. The success rate (15 out of 47) of PCR-screening described here suggests that at least 37 000 lines are required for saturation mutagenesis.

Several problems associated with features inherent in each insertion element have been noted when PCR-screening was carried out. For example, a proportion of mutations induced by Ac (Ref. 40) or Mutator (Ref. 25) were not inherited in the next generation, because the mutations detected were induced in somatic cells. Another problem is that only a fraction of mutants can be detected by PCR-screening because the primers cannot bind to T-DNA owing to frequent deletions induced at its ends\(^{41}\). Tissue culture-induced activation of RTNs seems most suitable for reverse genetic studies because transposition can be regulated strictly and rearrangements of RTNs are only rarely induced\(^{34,35}\). Because of the multiple copy nature of RTN insertions, the number of mutant lines required for saturation mutagenesis can be reduced.

Although PCR-screening seems to be the most efficient approach for reverse genetics, it might not be suitable for the analysis of many genes. Considering large-scale functional genomics in the post-sequencing era, a systematic approach to find mutants for a large number of genes is needed. One of the possible approaches is the cataloguing of insertion mutants by sequencing the genomic DNA flanking sequence insertions\(^{21}\), using IPCR or TAIL to amplify the flanking sequences. To carry out large-scale sequencing of the Tos17-flanking sequences, TAIL- and suppression-PCR were adopted. By combining these two PCR methods, ~95% of flanking sequences have been amplified. As of September 2000, more than 20 000 flanking sequences from 2387 lines have
been determined (A. Miyao et al., unpublished). These are classified into 8435 independent flanking sequences, and ~38% of the sequences showed homology to known genes. This strategy should be useful not only for the identification of mutants but also for the discovery of new genes. Similar analysis using Tto1 has been carried out in Arabidopsis34. If the entire genomic sequencing is completed as in Arabidopsis, the disrupted genes can be readily identified and the insertion sites can be mapped by comparing the flanking sequences with the genomic sequence. Thus, flanking sequence databases will become a powerful tool for reverse genetic studies.

Finally, although many insertion mutant lines have been produced in Arabidopsis by using class II elements and T-DNA, mutant libraries induced by RTNs are expected to complement other mutant libraries induced by other insertion elements, because each insertion element seems to have some extent to have different target-site specificity25.

Conclusions

Clearly, RTNs are one of the most fluid of genomic components, varying greatly in copy number and genomic localization and are therefore responsible for not only generating mutations in plant genomes but also for significantly increasing their size. However, another intriguing aspect is that these small mobile elements can themselves be used as genetic tools to analyse plant genomes and thereby improve our knowledge about the biology of plants. There are significant applications for RTNs in determining phylogeny, assessing genetic diversity and in the functional analyses of plant genes. The availability of the complete genomic sequences of several plant species, together with refinement of the information tools for data mining of mobile elements, should greatly facilitate a wider application of RTNs as genetic tools in plant biology in the near future. Finally, it is envisaged that RTN-based markers and insertional mutagenesis systems should add to the already powerful collection of available genetic tools in plant biology.

References

4 Kumar, A. et al. (1997) The Ty1-copia group of retrotransposons in plants: genomic organization, evolution, and use as molecular markers. Genetica 100, 205–217
29 Sato, Y. et al. (1999) Loss-of-function mutations in the rice homeobox gene OSH15 affect the architecture of internodes resulting in dwarf plants. EMBO J. 18, 992–1002
32 Fei, F. et al. (1999) The Mia (powdery mildew) resistance cluster is associated with three NBS-LRR gene families and suppressed
In December 2000, a participant on the Internet list server Bryonet (bryonet-1@mtu.edu) asked where one could find the latest classification of mosses. An answer came back quickly—Bryophyte Biology. Indeed, this is the source to consult for an overview of areas as varied as systematics, morphology, chemistry, development, ecology and bryogeography. This book supplements the New Manual of Bryology¹, bringing up-to-date many of the same topics and adding some current ones. However, with fewer than 500 pages, Bryophyte Biology cannot duplicate the Manual’s encyclopedic coverage.

In their preface, the editors state two ‘rather disparate needs’ that they hope this volume will meet: a replacement for Schofield’s out-of-print textbook² and an ‘accessible but inclusive reference on the biology of bryophytes’ for ecologists and molecular biologists. Their compromise is clearly in favor of advanced students of bryology and mature researchers rather than the novice. The specialized vocabulary is well beyond most undergraduates. There is no glossary and few illustrations are provided. Yet other editorial goals, such as ‘rich citation of both current and classic research’, are well met in each chapter.

Most chapters are written by authorities in their fields. For example, editor Jonathan Shaw reviews his own impressive and eclectic research (23 papers cited) on population ecology, population genetics and microevolution, along with that of 100 other workers. M.C.F. Proctor highlights the unique ecophysiological characteristics of bryophytes. David Cove summarizes the molecular genetics of model organism Physcomitrella patens. Michael Christianson contrasts the role of plant growth regulators in bryophytes and tracheophytes, and then presents case studies on the conversion of protenemal filaments from chloronema to caulonema and on initiation of shoot buds from protonema. A pair of chapters, one on peatland ecosystems by Dale Vitt and the other on the role of bryophytes in carbon budgets by K.P. O’Neill, are apropos given concerns about global climate change and increasing human pressures on boreal and arctic lands. Authors review recent advances in their subjects, but there is little new information here that is not scattered elsewhere in the literature.

My own area being systematics, I found that the first three chapters on classification and structure of hornworts, liverworts and mosses provided valuable updating for a bryologist who lately has spent more time among the taxonomic thorns of the Rose Family. As would be expected in any edited work, each set of authors approaches their chapter differently. The description of moss morphology is the easiest to digest, yet a novice student would find it difficult to form a clear image of the structures being described. There are only two plates of drawings and seven photographs. Although the two scanning electron micrographs are sharp, the others have lost resolution through digitization of images. William Buck and Bernard Goffinet present a classification of the mosses that is firmly grounded in tradition but with numerous innovations. They annotate each taxon above genus level with a short description. By necessity, these diagnoses are so brief that for clues on the arrangement of taxa you will need to browse the classic literature as well as newly published phylogenetic analyses based on molecular evidence³. Whereas moss structure is described at the level of cells and above, Karen Renzaglia and Kevin Vaughn describe hornwort anatomy and development in ultrastructural detail. Not being taxonomists, they contrast four alternative hornwort classifications recently published by others. The morphological description of liverworts is focused on their defining characteristics, such as oil bodies and elaters, rather than on an overview of the entire range of hepatic variation. Barbara Crandall-Stotler and Raymond Stotler include a cladistic analysis of 61 characters for 34 liverwort taxa plus five outgroup taxa. They defend their revised classification on the basis of this cladistic analysis and references to molecular studies. Within each phylum the classification system is elaborated down to the level of genus. You can locate the placement of your favorite genus through the index at the end of the volume, but no synonymy is given.

We are thematically reminded throughout the volume that in spite of their many unique features, these diminutive plants have much in common with other land plants. Continued dialogue and collaboration between bryologists and those working on tracheophytes will be mutually beneficial.

Joseph R. Rohrer
Dept of Biology, University of Wisconsin–Eau Claire, Eau Claire, WI 54702-4004, USA.
e-mail: jrohrer@uwec.edu

References
2 Schofield, W.B. (1985) Introduction to Bryology, Macmillan

43 Bhattacharyya, M.K. et al. (1997) A copia-like retrotransposon Tymr closely linked to the Rps1-k allele that confers race-specific resistance of soybean to Phytophthora sojae. Plant Mol. Biol. 34, 255–264
44 Kenward, K.D. et al. (1999) Isolation and characterization of Tnd-1, a retrotransposon marker linked to black root resistance in tobacco.

Book Review