The mechanism of gene targeting in *Physcomitrella patens*: homologous recombination, concatenation and multiple integration

Yasuko Kamisugi[1], Katja Schlink[2,4], Stefan A. Rensing[2], Gabriele Schween[2], Mark von Stackelberg[2], Andrew C. Cuming[1,*], Ralf Reski[2] and David J. Cove[1,3]

[1]Centre for Plant Sciences, Faculty of Biological Sciences, Leeds University Leeds, LS2 9JT, UK, [2]Plant Biotechnology, Faculty of Biology, University of Freiburg, Schaenzlestrasse 1, D-79104 Freiburg, Germany, [3]Department of Biology, Washington University in St Louis, St Louis, MO 63130-4899, USA and [4]Forest Genetics, Department of Plant Sciences, Life Science Center Weihenstephan, Technische Universität München, Am Hochanger 13, D-85354 Freising, Germany

ABSTRACT

The model bryophyte *Physcomitrella patens* exhibits high frequencies of gene targeting when transformed with DNA constructs containing sequences homologous with genomic loci. 'Targeted gene replacement' (TGR) resulting from homologous recombination (HR) between each end of a targeting construct and the targeted locus occurs when either single or multiple targeting vectors are delivered. In the latter instance simultaneous, multiple, independent integration of different transgenes occurs at the targeted loci. In both single gene and 'batch' transformations, DNA can also be found to undergo 'targeted insertion' (TI), integrating at one end of the targeted locus by HR with one flanking sequence of the vector accompanied by an apparent non-homologous end-joining (NHEJ) event at the other. Untargeted integration at nonhomologous sites also occurs, but at a lower frequency. Molecular analysis of TI at a single locus shows that this occurs as a consequence of concatenation of the transforming DNA, *in planta*, prior to integration, followed by HR between a single site in the genomic target and two of its repeated homologues in the concatenated vector. This reinforces the view that HR is the major pathway by which transforming DNA is integrated in *Physcomitrella*.

INTRODUCTION

Repair of double strand breaks in DNA is essential to ensure the integrity of genomes. The integration of transforming DNA into the genomes of higher eukaryotes occurs through the action of DNA double-strand break repair processes. While prokaryotes and lower eukaryotes favour the use of homologous sequences to repair chromosomal double-strand breaks (1) and to insert transforming sequences, illegitimate integration via microhomologies and non-homologous end-joining (NHEJ) is usual in higher eukaryotes.

Among plants, the moss *Physcomitrella patens* is an exception to this rule. When transforming DNA contains homology with the *Physcomitrella* genome, integration at the cognate genetic locus occurs at frequencies up to 100% (2–5), whereas in flowering plants, homologous recombination (HR) occurs at low frequency (10^{-4}–10^{-5}) (6). Since higher plants have attracted the most intense investigation of DNA repair processes, the mechanism of illegitimate recombination (IR) in plants is better characterized than that of HR. Unlike HR, IR frequently involves deletions and filler DNA insertion (7–10). Similarly, the random insertion of T-DNA sequences from *Agrobacterium tumefaciens* into plant genomes is also frequently associated with duplications or translocations of genomic sequences (11).

The prevalent models for double-strand break repair in somatic plant cells are based on single-strand-annealing (12) and the synthesis-dependent single-strand annealing mechanism (13). Both involve sequence homologies and can explain the capture of non-homologous or homologous sequences (1,14). The high rate of HR in *Physcomitrella,* compared with seed plants, could be due to (i) a different mechanism of recombination, (ii) deviations in proteins participating in recombination complexes or (iii) to a shift in the ratio of HR to IR without differences in the basic mechanisms. Unusually, we have found that whilst the majority of homologous transgene integration involves the accurate substitution of the genomic locus by the transforming DNA (targeted gene replacement-TGR), a significant proportion

---

of gene targeting events in *Physcomitrella* result in insertion of the transgenic sequence adjacent to the target locus (targeted insertion-TI). Typically, TI comprises an HR event at one end of the integrant, accompanied by an apparent NHEJ event at the other (4). To elucidate the nature of DNA integration in *Physcomitrella* in more detail, we have examined transgene integration patterns for both single constructs and for batches of multiple, independent targeting constructs, and analyzed a number of independent genomic integration sites at the DNA sequence level.

METHODS AND MATERIALS

Plant material

P.patens (Hedw.) B.S.G. was cultured as protonemal tissue, either on agar medium overlain with cellophanes (15) (Leeds) or in liquid culture (16) (Freiburg). The techniques for protoplast isolation, transformation and regeneration of stably transformed plants have been described previously (4,17–19).

Transformation vectors

Targeted loci analyzed in this study included single-gene disruptions of a 5-fatty acid desaturase gene (GenBank accession AX155059) and the *PpRac-1* gene (GenBank accession AY870928) (4), and batch transformations (codelivery of more than one construct) using cDNA clones in a minimalized plasmid vector, disrupted by an *nptII* selectable marker introduced by Tn1000 transposition (18). Linearized cDNA prepared for transformation by restriction enzyme digestion contained short (16–43 bp) non-homologous ends.

The vector RacKXH (4) was constructed in the Gateway™ vector pMBL6attR (4). A linear fragment for moss transformation was obtained by NotI digestion ('RacKXHN'), or by PCR ('RacKXHP') (Figure 3). These constructs carried 40 bp of terminal non-homology corresponding to the *AttB* sites of the Gateway™ vector.

Molecular analysis of transgenic plants

Southern blot analysis. Genomic DNA was isolated as previously described (20) and *ca.* 1 μg was digested for 5–6 h with 20 U of restriction enzyme (MBI Fermentas or New England Biolabs, Frankfurt/Main, Germany). After electrophoresis, the DNA was transferred to positively charged nylon membrane (Roche, Mannheim, Germany). Hybridization and detection were performed as described in the Roche DIG Application Manual using hybridization and blocking solutions and Anti-digoxigenin-AP conjugate from Roche and CDP-Star from Promega (Mannheim, Germany). Fluorescent bands were visualized on Hyperfilm ECL (Amersham Bioscience, Freiburg, Germany). DIG-labelled hybridization probes for detection of the *nptII* selection cassette were prepared by PCR-labeling (primers: K2E5f and K2E5r), and for detection of vector backbone, from plasmid DNA using the random-primed labeling mix from Roche and *Taq* polymerase provided by Promega.

PCR analysis of targeted loci. PCR primers used are listed in the Supplementary Information. TGR and TI loci were amplified by PCR for cloning or direct sequencing of amplicons. DNA was isolated from 14 day-old transgenic plants developed from small protonemal inocula, using the small-scale DNA isolation procedures previously described (15,20). TGR at the *PpRac-1* locus was identified by PCR using primers 35Spro2R (**p6**) and g6termF (**p7**) in conjunction with external gene-specific primers 5′racF (**p1**) and 3′rac3R (**p5**) (4). Untargeted junctions from plants in which 5′-TI had occurred were amplified by PCR with primer gRac3869S (**p8**) ('sequence **c**', the 3′-arm of the transforming cassette) in combination with primer racPro2R (**p3**) ('sequence **b**', the central, deleted region of the *PpRac-1* gene). Amplicons purified using a Qiagen Qiaquick PCR cleanup kit (Qiagen GmbH, Hilden) were sequenced using the ABI3130 capillary sequencer in the Faculty of Biological Sciences, Leeds University. Concatemer junctions in plants batch-transformed with cDNA constructs, were analysed using protocols for RACE–PCR (21,22). A 25 μl first-strand synthesis reaction contained 100 ng genomic DNA, 0.2 mM dNTPs, 1 pmol/μl primer and 0.01–0.08 U/μl *Taq* polymerase (Roche). Two reactions per transgenic line were set up, each with one of two outward primers located at the borders of Tn1000. The same PCR-cycle conditions were used for first-strand and second-strand synthesis (linear amplifications) and PCR amplification: following initial denaturation (5 min at 95°C), amplification was by 30 cycles of 0.5 min at 95°C–1.5 min at 60.7°C with a final extension of 4.5 min at 72°C. First-strand products were purified using the QIAquick PCR purification kit and eluted in 30 μl. These were incubated for 5 min at 65°C and directly cooled on ice. dC-tailing was performed in a 50 μl reaction volume with 31.5 U terminal transferase (MBI Fermentas) containing 1 mM $CoCl_2$ and 20 μM dCTP (20 min at 37°C) and stopped by heating to 72°C for 10 min. The DNA was again purified with the QIAquick PCR purification kit. Second-strand synthesis (40 μl) contained 1 pmol/μl G_{18}-primer (Thermo Electron, Dreieich, Germany), 0.2 mM dNTPs and 0.08 U/μl *Taq* polymerase. PCR amplification of the products followed immediately in a 45 μl volume by adding the respective first-strand primer to a final concentration of 1.7 pmol/μl, additional 0.03 U/μl Taq polymerase and dNTPs and $MgCl_2$ to final concentrations of 0.2 and 2.5 mM respectively. Products were cloned in pGEM-T (Promega) and sequenced on an ABI-Prism 377 sequencer (Applied Biosystems, Foster City, CA, USA) using the DYEnamic ET terminator cycle sequencing kit from GE Healthcare (Freiburg, Germany).

RESULTS

HR-mediated gene targeting results in accurate allele replacement and targeted insertion

When *Physcomitrella* is transformed with a targeting construct (a cloned gene disrupted by the insertion of a selectable marker cassette), gene targeting is detected by PCR using 'outward-pointing' primers specific to the selection cassette in combination with 'inward-pointing' 5′- and 3′-gene-specific primers corresponding to sequences external to the targeting construct. Amplification with both pairs is diagnostic of TGR resulting from accurate HR at each end of the construct; amplification with only one pair is indicative of TI resulting from accurate HR at only one end. Failure to amplify a fragment with both pairs is indicative of transgene

Table 1. Outcome of transformation with single and multiple gene fragments. Transgenic plants were analyzed by PCR to determine the nature of integration events

Construct	Transformants	HR/HR	5′-TI	3′-TI	Illegitimate
RacKXHN	130	57	39	8	26
RacKXHP	71	36	26	5	4
PpLEA-1	45	22	0	4	19
tKO25.2	12	5	4	0	3
7 cDNAs	140 loci	40	60	40	nd

The numbers of transformants of each class are provided. HR/HR indicates the number of targeted gene replacements. 5′-TI refers to targeted insertions resulting from HR only occurring in the 5′-arm of the targeting construct. 3′-TI refers to targeted insertions occurring only in the 3′-arm. Illegitimate events are transformants in which the targeting construct did not integrate at the targeted locus.

insertion at an illegitimate site (in such cases, the wild-type locus is amplifiable using the external primers). The accuracy of the HR events can be established by sequence analysis of the amplified fragments.

Initially, we performed PCR-screens in 258 transgenic lines transformed with single disruption constructs (Table 1) and analyzed the sequences of 32 recombination junctions at three loci where integration was by TGR. Five single-copy TGRs of the *PpRac-1* gene (GenBank accession AY870928) (4) were sequenced at each end following PCR-amplification of the HR-integration sites (Supplementary Figure 1A and B), as were five TGRs of a 5-fatty acid desaturase gene (GenBank accession AX155059). Additionally, twelve single-copy TGRs of the *PpLea-1* gene (GenBank accession AY870926) (4,23) were sequenced at one end (Supplementary Figure 1C). In each case, the recombination junctions showed no deviations from the sequence of the native gene, confirming that HR-mediated replacement of the wild-type gene by the disrupted construct occurred in an entirely precise manner.

Accurate gene targeting occurs following batch transformation using multiple constructs

It is also possible to achieve simultaneous targeting of multiple genes by cotransformation with batches of independent targeting constructs (17), however, the extent and accuracy of multiple targeting has not yet been systematically characterized. We therefore undertook a detailed molecular analysis of transgene integration following delivery of multiple constructs.

Efficient simultaneous delivery of multiple transforming DNAs has previously been used to create a large population of mutant plants (>50 000) by transformation with cDNA clones disrupted by the random insertion of a transposon-borne selection cassette (18,24). For efficient generation of this collection either normalized cDNA libraries, or batches of 7 or 20 disrupted cDNA clones, were delivered simultaneously to protoplasts. Since cDNA sequences typically exhibit shorter lengths of contiguous sequence homology with their genomic targets (due to the presence of introns within the genomic DNA), it was clearly important to verify that these knock-out constructs integrated into their target loci with a high degree of accuracy. We therefore estimated the numbers
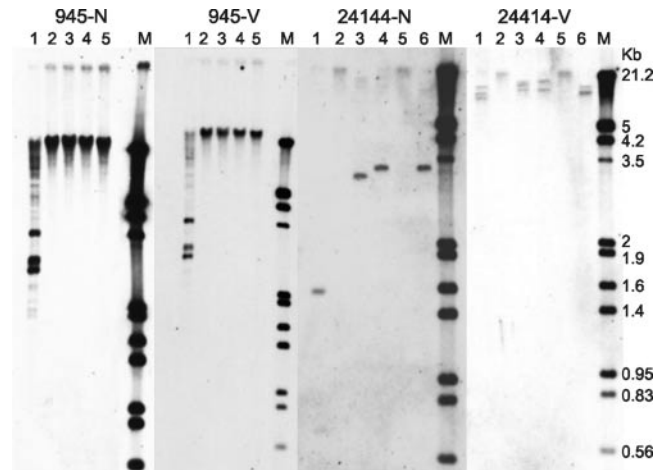
Figure 1. Batch transformation: integration of selection cassette and vector backbone. Southern blot analysis of two mutants from batch transformation with a pool of constructs. DNA (0.5–1 µg) was digested with **1**: PvuII; **2**: *Acc*651; **3**: BamHI; **4**: Eco4711; **5**: KspAI; **6**: NcoI. **945-N and 945-V:** Transgenic line 945 probed with an *nptII* (N) or vector (V) probe, respectively. **24414-N and 24414-V:** Transgenic line 24144 similarly probed. These lines demonstrate the variable range of transgene insertion: line 945 contains many transgenes, line 24414 contains only one. PvuII and NcoI cut once within the *nptII* cassette: the number of bands detected corresponds to [number of transgenic loci] × 2. The other enzymes generate an even distribution of genomic DNA fragments but do not cut within the transgenes or vector. The sizes of molecular weight markers (M) are indicated (kbp).

of integrations, and the nature of the integration events at targeted loci. The number of integration events was analysed by Southern blot analysis (Figure 1) in 146 independently regenerated transgenic lines [91 in this study and 55 in a previous study (18)]. On average, 11 copies of the selection cassette were integrated per plant. A subset of these transgenic plants [$n = 34$: 18 in this study and 16 previously (18)] were further analyzed to determine the number of integration sites. The number of integration sites per plant showed a high degree of variation, from 1 to 16 loci, with the majority of plants (59%) exhibiting transgene integration at between 3 and 7 loci (Figure 1).

In these experiments, the DNA used for transformation was linearized by restriction enzyme digestion to cut the cDNA insert out of the vector. Consequently, both *nptII*-disrupted cDNAs and vector backbone were delivered to protoplasts. The adventitious integration of vector sequence was demonstrated both by Southern blotting (Figure 1) and by the isolation and sequencing of concatemer structures that comprised cDNA and vector sequences from three 'batch mutants'.

In mammalian cells, the copy number of integrated transgenes is related to the amount of DNA delivered per transformation (25). We tested this for *Physcomitrella*. Protoplasts were transformed with 5–50 µg DNA/transformation and the numbers of *nptII* copies per plant were estimated by Southern hybridization. The efficiency of transformation increased dramatically with the quantity of DNA, reaching an optimum at 20 µg, but the number of integrated transgenes per plant was not influenced by the quantity of DNA per transformation (Table 2).

The nature of the targeting events (TGR versus TI) was determined in plants following the batchwise delivery of

Table 2. Transformation efficiency as a function of quantity of transforming DNA. Protoplasts were transformed with 5–50 µg DNA in 38 independent transformation experiments

| Quantity of DNA (µg)	5	10	15	20	25	30	50
NptII copies/plant	11.9	11.3	nd	12.1	nd	nd	11.5
Transformants/ transformation	2.2	6.6	14.3	19.1	18	19.1	20.1
Transformants/µg DNA	0.44	0.66	0.95	0.96	0.72	0.62	0.4

The average number of stable transformants obtained, per transformation is shown, as are the average numbers of transgenes integrated in 22–24 representative transgenic lines regenerated following transformation with 5, 10, 20 and 50 µg DNA, as estimated by Southern blot hybridization using an *nptII* probe.

seven different constructs, in which the selection cassette was flanked by at least 500 bp of genomic DNA on either side. The nature of targeted integration events at each locus was analyzed by PCR in a number of independently derived transgenic plants (in total, 140 targeted loci were analyzed). At 40 of these loci, accurate TGR had occurred. The remaining 100 gene targeting events comprised TIs, with HR occurring either in the 5′-flanking sequence ($n = 60$) or the 3′-flanking sequence ($n = 40$) of the targeted locus (Table 1, row 5). Nine independently regenerated transgenic plants obtained in this way were further analyzed by PCR, using primers specific to the selection cassette and to each cDNA used in the targeting constructs, in order to detect the numbers of transgenes integrated. On average, 86% of the targeting constructs were found to have become integrated into the genomes of these plants, either by HR or IR.

The accuracy of targeting events in batch transformants was determined by sequencing PCR-amplified junctions from 12 loci at which TGR had occurred. This analysis showed precise integration of the targeting cassette at the cognate locus, but also that the recombination event could occur at any point within the homologous sequence present in the targeting construct, as evidenced by loss of intron sequences occurring as a consequence of recombination between cDNA and genomic sequence (Supplementary Figure 2). In one plant only, we observed an anomalous event: the duplication and translocation of a genomic sequence of at least *ca.* 1.5 kb lying adjacent to the transgene integration site (Supplementary Figure 3). The genomic integration site of the duplicated sequence could not be determined.

How do targeted insertion events arise?

We envisage three ways in which a targeting construct might integrate via TI. These alternatives are illustrated in Figure 2B–D. Consider a vector constructed from a cloned gene '**a-b-c**' in which the central segment ('**b**') is replaced by a selectable marker cassette '**sel**'. This vector additionally is terminated by two distinct, non-homologous, cloning vector-derived sequences '**B1**' and '**B2**'. TGR will result in a precise replacement of the endogenous locus '**a-b-c**' by the sequence '**a-sel-c**' (Figure 2A). TI might occur through an HR event between one arm of the vector and its target, accompanied by an NHEJ event between the other end of the vector and the DNA double-strand break in the targeted locus. This will generate a transgenic locus '**a-sel-c-B2-a′-b-c**'
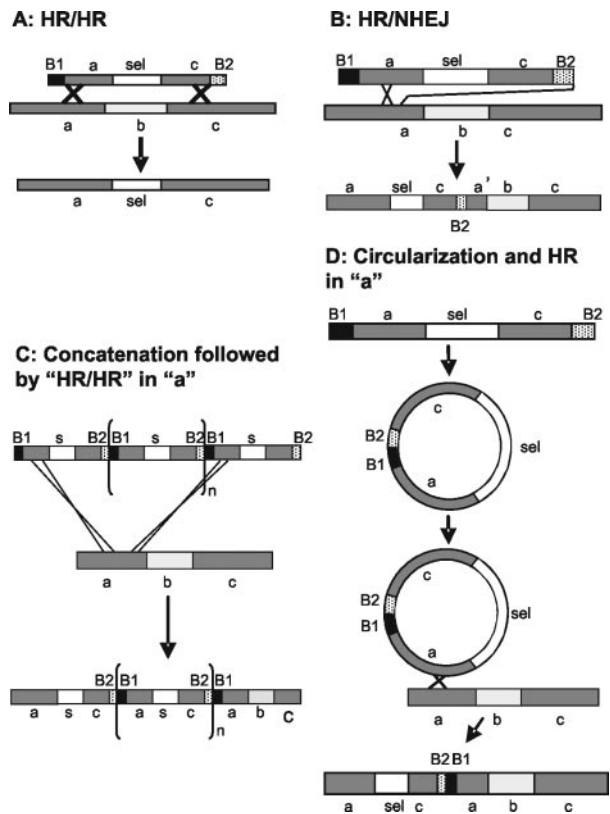
Figure 2. Possible models for integration of transforming DNA. (**A**) The transforming fragment integrates by HR in both arms of the vector, and a perfect TGR occurs. (**B**) The transforming fragment integrates by HR in one arm of of the vector: the other arm invades the breakpoint to integrate by NHEJ. (**C**) The transforming fragment undergoes head-to-tail concatenation prior to integration. Integration occurs in region '**a**' of the targeted locus by two HR events occurring in separated regions '**a**' in the concatemer. '**S**' and '**sel**' indicate the *nptII* selection cassette. '**B1**' and '**B2**' indicate the a*ttB1* and a*ttB2* sites that terminate the transforming DNA fragment. (**D**) The transforming fragment is circularized by NHEJ. Recombination at a single point in region '**a**' generates a transgenic locus indistinguishable from that in '**C**'.

(Figure 2B), where **a′** represents a terminally truncated sequence '**a**'. Alternatively, the vector may form a concatemer within the cell, and integrate by an HR event between one arm of the vector and its target, and a second HR event between the DNA double-strand break in the target and another identical homologous sequence available in the concatenated vector, as illustrated in Figure 2C. Here, the resulting transgenic locus will comprise '**a-sel-c-B2-(B1-a-sel-c-B2)$_n$-B1-a-b-c**'. Finally, if the vector becomes circularized, a single HR event between this and the target locus would generate the transgenic locus '**a-sel-c-B2-B1-a-b-c**', in which only a single copy of the vector becomes inserted (Figure 2D). These outcomes are predictable and testable by PCR amplification and sequencing of the untargeted junction sequence in TI events.

We analyzed transgenic plants in which the *PpRac-1* locus (Figure 3A) was targeted. Plants were transformed with the targeting vector 'RacKXH' (4). This comprises 1119 bp from the 5′-end of the *PpRac-1* gene (sequence '**a**') and 529 bp from the 3′-end of this gene (sequence '**c**') interrupted by an *nptII* selectable marker cassette ('**sel**') which replaced
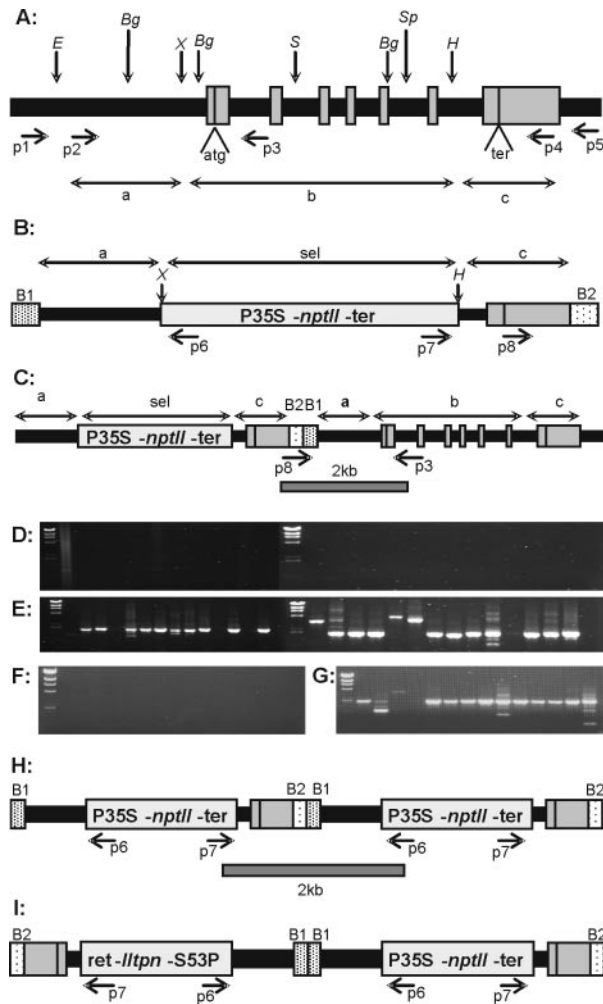
Figure 3. Targeted insertion at the *PpRac-1* locus. (**A**) Structure of the *PpRac-1* gene. Exons shown as shaded boxes, and non-coding DNA (5′- and 3′-flanking sequences and introns) as a solid line. Restriction sites for BglII (Bg), EcoRI (E), HindIII (H), SalI (S), SphI (Sp) and XbaI (X) are indicated. PCR primers p1 and p5 are external primers used to confirm gene targeting in conjunction with the 'outward' primers p6 and p7 (**B**). Primers p2 and p4 were used to amplify the *PpRac-1* gene and create the targeting fragment RacKXH. (**B**) The targeting construct RacKXH. The selection cassette replaces the central XbaI–HindIII fragment. (**C**) The outcome of a TI event of the type shown in Figure 2C. 'B1' and 'B2' denote the *AttB1* and *AttB2* sites flanking the transforming fragment. Primer p3 is specific for the region of the *PpRac-1* gene deleted in the construction of RacKXH ('region **b**' in Figure 2), and is used in conjunction with primer p8, corresponding to 'region **c**', for amplification of 3′-NHEJ junctions, indicated by the grey bar (**D** and **E**). DNA from transgenic plants in which integration occurred at non-targeted sites. (**D**) Amplification using primer p6. (**E**) Amplification using primers p6 and p7 together. (**F** and **G**) DNA from transgenic plants in which 5′-TI at the *PpRac-1* locus occurred. (**F**) Amplification with primer p6. (**G**) Amplification with primers p6 and p7 together. (**H**) Predicted structure of concatenated transforming DNA by 'head-to-tail' ligation. (**I**) Predicted structure of concatenated transforming DNA by 'head-to-head' ligation.

a central XbaI–HindIII fragment (sequence '**b**'). The terminal Gateway™ vector *AttB1* and *AttB2* sequences comprised the non-homologous sequences '**B1** and **B2**' at either end of the fragment (Figure 3B). The transforming fragment was prepared either from its plasmid vector by digestion with NotI ('RacKXHN') or by PCR amplification ('RacKXHP'). In the first case, the transformation mixture contained

equimolar quantities of both the disrupted *PpRac-1* gene and the plasmid vector backbone. Additionally, the fragments terminated in NotI cohesive ends. In the latter case, protoplasts were transformed with a highly pure DNA fragment comprising the disrupted *PpRac-1* gene that retained the terminal a*ttB* sites, but with non-cohesive ends.

We obtained 201 transgenic plants using these fragments, of which 93 exhibited accurate TGR. Among the remainder, 65 had undergone TI in the 5′-arm of the targeting construct, 13 had undergone TI in the 3′-arm of the targeting construct and 30 contained DNA integrated ectopically at unknown sites in the genome (4). We analyzed the DNA in these plants by PCR to determine two features. (i) The nature of concatenation between transforming fragments and (ii) the nature of the non-targeted junction between the transforming DNA and its target locus, in plants exhibiting 5′-TI.

### Concatenation of transforming DNA: single targeting vectors

Using the 'outward-pointing' primers 'p6' and 'p7', corresponding to the *nptII* selection cassette (Figure 3B), it is possible to determine whether concatenates form, and their relative orientation. 'Head-to-tail' concatemers will be amplified using both primers in conjunction (Figure 3H). 'Head-to-head' and 'tail-to-tail' concatemers will be amplified using a single primer only (e.g. Figure 3I).

We first analyzed the 30 plants in which the DNA had integrated ectopically. Use of a single primer failed to generate an amplification product (Figure 3D), whereas the combination of primers 'p6' and 'p7' resulted in amplification of a product from 23 plants, and in 20 of these, the fragment was of the size predicted for a 'head-to-tail' concatemer (Figure 3E and H). Three plants yielded larger amplicons, probably resulting from DNA rearrangement in the generation of the transgenic locus. Similarly, when we analyzed plants in which 5′-TI had occurred, 13 out of 14 plants tested exhibited the characteristic amplification pattern of head-to-tail concatemers (Figure 3G and H). Head-to-head and tail-to-tail concatemers were not detected (Figure 3F and I). We therefore conclude that transgenic loci containing multiple copies of transforming DNA comprise head-to-tail concatemers of the transforming fragment.

Using primers external to the targeted locus ('p1' and 'p5' in Figure 3A), we were able to amplify single-copy integrants that resulted from accurate TGR (4). However, no single-copy integrants resulting from TI could be identified in the *PpRac-1* locus.

Concatenation of transforming DNA: multiple targeting vectors

Although tandem, head-to-tail concatemers are the only structures detected when a single transforming fragment is delivered to protoplasts, this is not the case when multiple independent fragments are delivered in a batch transformation. Isolation and sequence analysis of single transgenic loci at which concatenated cDNA-based vectors had integrated by IR revealed concatemers between different cDNA constructs in all possible orientations (Figure 4). We found six loci at which concatemers arose from HR between identical constructs (data not shown). Eight loci contained
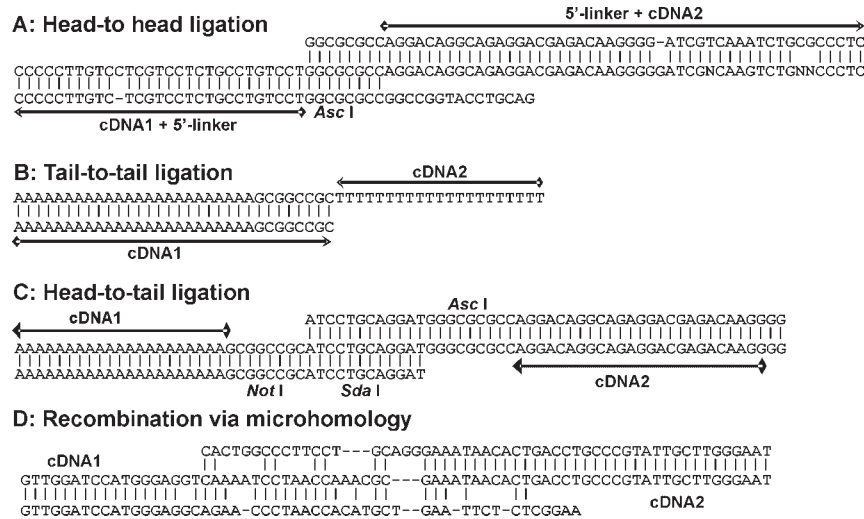
Figure 4. Configuration of cDNA concatemers at TI loci. Examples of concatenated cDNAs in transgenic loci amplified from plants multiply transformed with batches of multiple disrupted cDNAs. Arrows indicate the directions of the reading frames. Concatenation resulted in the loss of a terminal SdaI restriction site present in the original constructs, adjacent to the 5′-end AscI and 3′-end NotI sites, respectively. (**A**) cDNAs joined 'head-to-head' have a single AscI site between them. (**B**) cDNAs joined 'tail-to-tail' with a single NotI site between the poly(A)-tails. (**C**) cDNAs joined 'head-to-tail' identified by the relative dispositions of the poly(A) tail of cDNA1, the 3′ NotI site and the AscI site at the 5′-end of cDNA2. (**D**) Two independent sequences have joined, probably by recombination via microhomologies. The microhomologous sequences in the concatemer are separated by several hundred base pairs from the ends of each of the two constructs from which the concatemer derived.

concatemers that were derived from end-to-end ligation, of which three were in head-to-head (Figure 4A), three in tail-to-tail (Figure 4B) and two in head-to-tail orientation, respectively (Figure 4C). Additionally 3 loci contained concatemers that were probably formed by recombination between microhomologous sequences in different cDNAs (Figure 4D). We routinely observed short deletions at the ends of each concatenated transgene with an average length of 10 bp ($n = 10$). We therefore conclude that concatenation of transforming DNA by NHEJ is a frequent occurrence during protoplast transformation.

It is possible that concatemers could have formed by annealing between complementary cohesive ends of DNA molecules in the highly concentrated DNA mixture prior to transformation, rather than *in planta*, following DNA delivery. This was tested in cDNA batch transformation experiments. To discriminate between the two possibilities, the DNA was either used directly for transformation or heated to 60°C for 10 min, to denature concatemers formed by annealing of cohesive ends, then chilled on ice prior to transformation. The numbers of *nptII* cassettes and integration loci were estimated by Southern blot analysis (Table 3). Neither the number of *nptII* cassettes, the number of integration loci, nor the transformation efficiency showed a significant difference between treatments. Since the integration of concatemers is also a feature of plants transformed with the blunt-ended PCR fragment RacKXHP, we conclude that concatenation occurs *in planta*.

Sequence analysis of untargeted junctions

Plants exhibiting TI in the 5′-arm of the *PpRac-1* gene (Figure 3C) were selected for analysis of the untargeted (3′) integration junction, the shorter length of the 3′ targeting sequence enabling the PCR amplification and subsequent sequencing of this junction using the primers 'p3' and 'p8'

Table 3. Transformation with heat-denatured DNA protoplasts were transformed with batches of disrupted cDNA clones (50 µg DNA) linearised by digestion with SdaI

	Transformation efficiency	*nptII* cassettes	Integration loci
DNA (untreated)	30.4	14.8	Range 1–20 Median 12
DNA (60°C/chilled)	32.4	12.9	Range 2–17 Median 10
Two-sided *t*-test: $t =$	0.55	0.204	1.65
Probability $P =$	0.58	0.84	0.1
$N =$	Untreated: 48 Heat-treated: 24	9	9

DNA was used directly, or heated at 60°C for 10 min and chilled on ice before transformation to prevent concatenation, *in vitro*, via annealing of cohesive ends. Transformation efficiency was determined as the mean number of transgenic plants per transformation. The mean numbers of integrated *nptII* cassettes per plant and the numbers of integration loci per plant were estimated by Southern blot analysis of transgenic lines ($n = 9$).

(Figure 3A and C). Because primer 'p3' corresponds to the central region of the *PpRac-1* gene (sequence 'b'), deleted from the transforming construct, only the 3′-untargeted junction will be amplified using this primer combination.

When 19 transgenic lines carrying 5′-TIs of the RacKXHN fragment were subjected to amplification with these primers, 10 yielded a PCR product. Similarly, of the 26 transgenic lines carrying 5′-TIs of the RacKXHP fragment, 21 yielded a PCR product. All but one of these latter amplicons were electrophoretically identical, the remaining fragment being slightly longer. This fragment, together with nine other RacKXHP amplicons and all ten RacKXHN amplicons were sequenced.

All but 2 of the 20 amplified junction sequences contained the '**c***AttB2-AttB1***-a-b**' configuration consistent with

A: RacKXHN junction sequences

```
              Region "c"           AttB2              Xba I   Not I   Xba I              AttB1                  Region "a"
Predicted  AAATTGGGATCACCCAGCTTTCTTGTACAAAGTGGTTGATCTAGAGC-GGCGC-TCTAGATCAACAAGTTTGTACAAAAAAGCAGGCTAGCCGCCTCCAA
 T8/3-56   AAATTGGGATCACCCAGCTTTCTTGTACAAAGTGGTTGATCTAGAGC-GGCGC-TCTAGATCAACAAGTTTGTACAAAAAAGCAGGCTAGCCGCCTCCAA
 T8/3-20   AAATTGGGATCACCCAGCTTTCTTGTACAAAGTGGTTGATCTAGA--------------TCAACAAGTTTGTACAAAAAAGCAGGCTAGCCGCCTCCAA
 T8/3-63   AAATTGGGATCACCCAGCTTTCTTGTACAAAGTGGTTGATCTAGA--------------TCAACAAGTTTGTACAAAAAAGCAGGCTAGCCGCCTCCAA
 T8/3-64   AAATTGGGATCACCCAGCTTTCTTGTACAAAGTGGTTGATCTAGA--------------TCAACAAGTTTGTACAAAAAAGCAGGCTAGCCGCCTCCAA
 T8/3-29   AAATTGGGATCA------------------------------------------------------------------------------------CCAA
 T8/3-61   AAATTGGGATCACCCAGCTTTCTTGTACAA-----------------------------------------------AAAAGCAGGCTAGCCGCCTCCAA
 T8/3-62   AAATTGGGATCACCCAGCTTTCTTGTACAA-----------------------------------------------AAAAGCAGGCTAGCCGCCTCCAA
 T8/3-89   AAATTGGGATCACCCAGCTTTCTTGTACAA-----------------------------------------------AAAAGCAGGCTAGCCGCCTCCAA
 T8/3-45   AAATTGGGATCACCCAGCTTTCTTGTACAAAGTGGTTGATCTAGAGC-GGC---+455bp Vector--TTGTACAAAAAAGCAGGCTAGCCGCCTCCAA
 T8-3-44   AAATTGGGATCACCCAGCTTTCTTGTACAAAGTGGTTGATCTAGAGC-GGCGC
                                                                       \_ + entire vector _/
                                                     GC-GGCGC-TCTAGATCAACAAGTTTGTACAAAAAAGCAGGCTAGCCGCCTCCAA
```

B: RacKXHP junction sequences

```
              Region "c"           AttB2                           AttB1               Region "a"
Predicted  AAATTGGATCACCCAGCTTTCTTGTACAAAGTGGTCCCCGGGGACAAGTTTGTACAAAAAAGCAGGCTATCCGCCTCCAA
 T15/5-1   AAATTGGATCACCCAGCTTTCTTGTACAAA--------------------------AAAGCAGGCTATCCGCCTCCAA
 T15/5-19  AAATTGGATCACCCAGCTTTCTTGTACAAA--------------------------AAAGCAGGCTATCCGCCTCCAA
 T15/5-22  AAATTGGATCACCCAGCTTTCTTGTACAAA--------------------------AAAGCAGGCTATCCGCCTCCAA
 T15/5-23  AAATTGGATCACCCAGCTTTCTTGTACAAA--------------------------AAAGCAGGCTATCCGCCTCCAA
 T15/5-31  AAATTGGATCACCCAGCTTTCTTGTACAAA--------------------------AAAGCAGGCTATCCGCCTCCAA
 T15/5-48  AAATTGGATCACCCAGCTTTCTTGTACAAA--------------------------AAAGCAGGCTATCCGCCTCCAA
 T15/5-49  AAATTGGATCACCCAGCTTTCTTGTACAAA--------------------------AAAGCAGGCTATCCGCCTCCAA
 T15/5-52  AAATTGGATCACCCAGCTTTCTTGTACAAA--------------------------AAAGCAGGCTATCCGCCTCCAA
 T15/5-24  AAATTGGATC--------------------------------------------------ATCCGCCTCCAA
 T15/5-55b AAATTGGATCACCCAGCTTTCTTGTACAAAGTGGTC------GACAAGTTTGTACAAAAAAGCAGGCTATCCGCCTCCAA
                                           \_ AAGCTAATTCGAGCTCCACCGCGGTGGCGGCCGCTCTAGAGATC _/
```

Figure 5. Sequences of the untargeted junction in *PpRac-1* loci exhibiting 5′-TI. Primers 'p3' and 'p8' were used to amplify the region '**c-?-b**' in plants identified as having undergone 5′HR/3′-insertion following transformation with the single genomic constructs 'RacKXHN' (**A**) and 'RacKXHP' (**B**). Sequences of the junctions are shown aligned with the predicted outcome of the event illustrated in Figure 3C. (A) For RacKXHN, Line T8/3–56 retains a near-identical match with the predicted outcome. Lines T8/3–20, –63 and –64 are deleted between the microhomologous XbaI sites, while lines T8/3–61, –62 and –89 are deleted between microhomologous sites in the *AttB2* and *AttB1* sequences. Line T8/3–29 is deleted for the *AttB* sites. Lines T8/3–44 and T8/3–45 contain insertions derived from vector sequences. For T8/3–44, this insertion comprises the entire vector. (B) For RacKXHP transformants, Lines T15/5–1, –19, –22, –23, –31, –48, –49 and –52 have identical deletions between the microhomologous sequence 'AAA' in the *AttB* sequences. Line T15/5–24 is deleted for the *AttB* sites between the microhomologous sequence 'ATC'. Line T15/5–55b contains a short insertion of 'filler' DNA of unknown provenance.

formation of the untargeted junction by HR as illustrated in Figure 2C and D. However, all of the junction sequences were characterized by local rearrangements between the *AttB* sites.

In the lines transformed by RacKXHN, one transformant contained a near-perfect match with the predicted concatemer junction sequence (Figure 5A: line T28/3–56). Three lines contained a small deletion in the microhomologous XbaI–NotI–XbaI sequence located between the *AttB* sites (Figure 5A: lines T28/3–20, 3–63, 3–64), whilst a further three exhibited a deletion in a region of microhomology between the two *AttB* sites themselves, but still carried the recognizable signatures of the *AttB1* and *AttB2* motifs (Figure 5A: lines T8/3–61, 3–62 and 3–89). Two lines contained additional sequences derived from the vector backbone. In one case, this comprised the entire vector (Figure 5A: line T8/3–45). In the other line, however, the structure of the NHEJ junction was more complex (Figure 5A: line T8/3–44). The amplified fragment contained a 455 bp insertion between the *AttB2* and *AttB1* sites, made up of two sequences of 212 and 243 bp derived from the plasmid vector backbone. These were not contiguous sequences within the source plasmid. Instead, they appear to result from a 1790 bp deletion. We suggest that this arises from a recombination event between two microhomologous CA dinucleotides that flank the deleted intervening sequence. Because in the generation of these transformants, no effort was made to purify the targeting fragment free of its vector, it is likely that T8/3–44 and T8/3–45 derived from an initial extrachromosomal concatenation event between the targeting fragment and the free vector sequence. The tenth line analyzed (Figure 5A: line T28/3–29) was deleted for the *AttB2* and *AttB1* sites, but still retained the '**c-a-b**' sequences in the orientation predicted in Figure 2C and D.

In the lines transformed with RacKXHP, 9 of the 10 fragments contained the '**c**-*AttB2*-*AttB1*-**a-b**' configuration. The tenth (Figure 5B: line T15/5–24) was deleted for the *AttB* sites between a flanking microhomology, but retained the characteristic '**c-a-b**' sequence arrangement. Of the nine fragments containing *AttB* sites, eight displayed a small deletion between regions of microhomology on the *AttB2* and *AttB1* sequences (Figure 5B: lines T15/5–1, 5–19, 5–22, 5–23, 5–31, 5–48, 5–49 and 5–52) whilst the ninth (Figure 5B: line T15/5–55b, corresponding to the larger, amplified fragment) contained a small insertion of 'filler DNA' of unknown origin, characteristic of NHEJ (7–10).

These data are consistent with a model for integration by TI in which two HR events occur between the target locus and a tandemly repeated targeting construct formed by concatenation of the transforming DNA (Figure 2C).

DISCUSSION

In *Physcomitrella*, multiple loci may be independently targeted within a single transformation experiment, using mixtures of targeting constructs. In such cases, each locus is targeted with the same frequency as if the single construct is used. However, the probability of obtaining multiply targeted mutants is a function of the multiplied frequencies observed for individual constructs (17). Multiple transformation was developed as means by which large numbers of targeted mutants could be efficiently generated, and our analysis of multiply transformed plants confirms that the majority of targeting constructs (86%) delivered in a batchwise manner become integrated in the transgenic lines so generated. However, phenotypic analysis of such mutants requires that several caveats be observed. In particular, it is imperative that the mutant phenotype be shown not to derive from an

adventitious integration of vector DNA, elsewhere within the genome, nor from an ectopic insertion of a targeting cassette. As with all analysis of mutant strains, the standard precaution of genetic analysis of mutants should be routinely undertaken, by crossing to a wild-type strain in order to demonstrate Mendelian segregation of the mutant phenotype with a targeted locus. Complementation of the mutant strain by further transformation with a wild-type allele should provide additional confirmation that a particular targeted gene is responsible for the phenotype in question.

Integration of targeting DNA into the homologous genomic locus frequently occurs via HR at both ends. However, transforming DNA can integrate at multiple genomic loci (from 3 to 7 for batch transformation) due to additional IR. Additionally, multiple copies of transgenes are integrated (ave = 11). We ascribe this to the concatenation of transforming DNA being the rule, rather than the exception: we have shown previously that between 40 and 85% of TGR loci generated by single transformation constructs contain multiple copies of the targeting cassette (4).

Concatenation is most frequent in TI loci. Significantly, it appears that such events do not arise by a combination of HR and NHEJ, as first postulated (4), but arise wholly by HR-mediated integration of concatenated DNA. It is known that concatenated copies of plasmid DNA insert at one or a small number of loci when circular molecules are delivered to eukaryotic cells (25–28), and linear DNA delivered to mammalian cells can undergo end-to-end ligation, by NHEJ, all possible orientations of ligated fragments being detectable (25). Such a process is the likely cause of the incorporation of nonspecific 'carrier' DNA in transgenic loci in plant cells, following its co-delivery with a circular transforming plasmid (29). Thus, after entering the cell, transforming DNA can undergo both fragmentation and religation events prior to its insertion into the genome.

In *Physcomitrella,* a typical transformation experiment generates, in addition to stable (integrated) transformants, a larger number of unstable transformants, that maintain transgenes only so long as selection is maintained. These greatly outnumber stable transformants, when circular DNA is delivered to protoplasts. Unstable transformants contain extrachromosomal concatemers of the transforming DNA (30) and the generation of such concatemers may occur prior to and preferentially over transgene integration for circular molecules.

Delivery of linear DNA reduces the frequency of unstable transformants (17), but evidently concatenation still occurs. This must occur *in planta*, since prior heat-denaturation of potential cohesive ends prior to DNA delivery did not prevent the recovery of transformants containing concatenated transgenes (Table 2). Recombination between a linear concatemer and the target locus would generate TI events by the model indicated in Figure 2C. Although recombination between a circularized fragment of transforming DNA would also have the same outcome (Figure 2D), rates of transformation by circular DNA are orders of magnitude lower than those resulting from delivery of linear DNA in *Physcomitrella* (17) and consequently we do not favour a model involving circular intermediates.

The occurrence of small deletions between the terminal sequences ('**c-a**') that are found at the untargeted junction in TI events are characteristic of NHEJ events, as is the occurrence of insertions via DNA sequence microhomologies (10,31). We observe just such insertions and deletions in nearly all the untargeted junctions analyzed, strongly supporting the view that TI involves ligation of fragments by NHEJ prior to homologous integration into the genome.

We propose that TI events involve the following series of events. Following uptake of massive numbers of linear fragments by the cell, their ends would be perceived by the DNA damage repair machinery as the apparently catastrophic occurrence of double-strand breaks. The most rapid mechanism by which these could be 'repaired' is by NHEJ between these fragments. A priori, we would expect the ends to be joined at random, so that concatemers containing head-to-tail, head-to-head and tail-to-tail junctions should be formed in equal number. Such randomly joined concatemers are indeed found integrated in plants batch-transformed with multiple, heterogeneous cDNA targeting constructs. However, only head-to-tail concatemers of the transforming DNA are detected in plants transformed with single genomic targeting constructs. End-to-end ligation of identical sequences to form a large inverted repeat probably generates a transforming molecule either (i) not readily detectable by PCR, or (ii) not suitable for TI or (iii) a transgenic locus in which the inverted repeats are unstable, and are excised.

The model presented in Figure 2C requires a concatemer to be the substrate for TI in the target locus. The mechanism we suggest for such insertion is elaborated in Figure 6, and is based on current models for HR-mediated DNA repair (32). One end of the concatemer (region '**a**') initiates HR (Figure 6B), following recruitment of DNA repair proteins resulting in resection of the targeting construct to generate 3′-ssDNA (presumably complexed with *PpRad51* proteins (33) and *Rad52* epistasis group homologues) that can initiate invasion of the homologous genomic target (32). Strand scission of the targeted locus (Figure 6C) would then generate a genomic 3′-ssDNA that could, in turn, invade the concatenated targeting construct (Figure 6D). For a long concatemer, invasion by genomic DNA of a nearby repeat of region '**a**' in the transforming DNA might be sterically more favourable than invasion of the genomic region '**c**' by the distant terminus of the targeting construct. We note that the length of the concatemer is not fixed, but if it were only a dimer then only a single copy of the targeting construct would be inserted, and consequently no PCR products would be obtained using the outward-facing primers '**p6**' and '**p7**' (Figure 3H). Since in 13 of the 14 plants we analyzed, a PCR product was obtained (Figure 3G), the insertion substrate must have been at least trimeric. This may reflect either (i) that dimers are structurally unsuitable for insertion, (ii) that dimers are preferentially integrated as multicopy TGRs (4), or (iii) may be a stochastic reflection of the population of available concatemers. If multiple subunit concatemers were very common, the chance of the second recombination event occurring in the concatemer subunit adjacent to the first would be low, and our sample may be too small to identify such an event.

Overall, the observation that the overwhelming majority of targeted transformation events in *Physcomitrella* involve only HR, supports the contention that HR represents the predominant route for transgene insertion in this species.
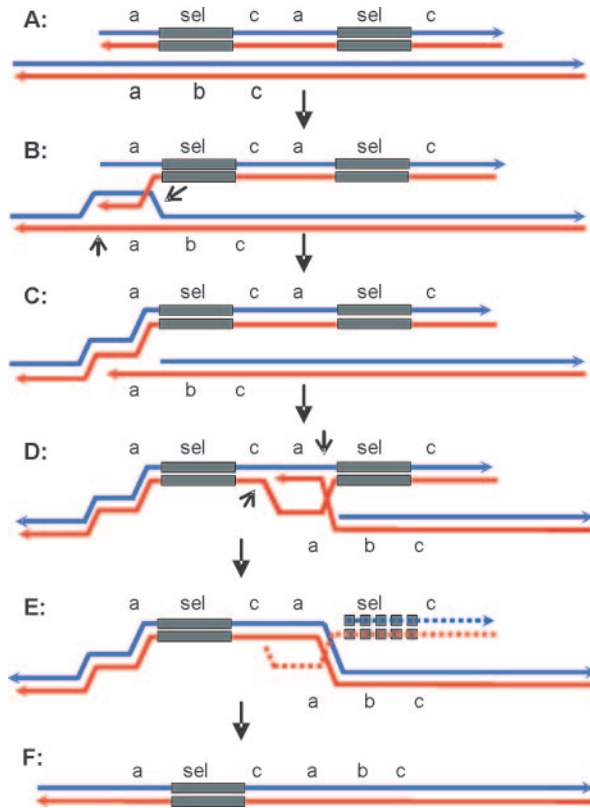
Figure 6. Proposed mechanism for TI. (**A**) The concatenated targeting construct (here shown as a dimer, containing two selection cassettes-'**sel**', to reduce the complexity of the diagram) aligns with the targeted locus-'**a-b-c**'. (**B** and **C**) One end of targeting construct is recognized by the cellular DNA repair machinery as a DNA ds-break and is resected to generate a 3′-ss overhang. This invades the genomic target sequence. Scission of the genomic target strands at the points indicated by arrowheads permits extension of the invading strand by copying the invaded template sequence. Resolution of the structure at this end of the targeted sequence is completed by joining of the upper (blue) strands. (**D** and **E**) The ds-break formed by the strand scission of the genomic locus initiates repair by invading the concatenate at its homologous sequence ('**a**'). Scission of the invaded DNA at the points arrowed permits strand extension of the invading genomic ssDNA and resolution of the recombination complex in the same way as occurred at the other end of the targeting construct. The residual DNA fragment (dashed lines) is degraded. (**F**) The outcome of construct-initiated HR at one sequence '**a**', and target-initated HR at the next sequence '**a**' is the generation of the transgenic TI locus '**a-sel-c-a-b-c**'.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

REFERENCES

1. Gorbunova,V. and Levy,A.A. (1999) How plants make ends meet: DNA double-strand break repair. *Trends Plant Sci.*, **4**, 263–269.
2. Reski,R. (1998) *Physcomitrella* and *Arabidopsis*: the David and Goliath of reverse genetics. *Trends Plant Sci.*, **3**, 209–210.
3. Schaefer,D.G. (2001) Gene targeting in *Physcomitrella patens*. *Curr. Opin. Plant Biol.*, **4**, 143–150.
4. Kamisugi,Y., Cuming,A.C. and Cove,D.J. (2005) Parameters determining the efficiency of gene targeting in the moss *Physcomitrella patens*. *Nucleic Acids Res.*, **33**, e173.
5. Cove,D. (2005) The moss *Physcomitrella patens*. *Annu. Rev. Genet.*, **39**, 339–358.
6. Britt,A.B. and May,G.D. (2003) Re-engineering plant gene targeting. *Trends Plant Sci.*, **8**, 90–95.
7. Gheysen,G., Villarroel,R. and VanMontagu,M. (1991) Illegitimate recombination in plants—a model for T-DNA integration. *Genes Dev.*, **5**, 287–297.
8. Salomon,S. and Puchta,H. (1998) Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. *EMBO J.*, **17**, 6086–6095.
9. Takano,M., Egawa,H., Ikeda,J.E. and Wakasa,K. (1997) The structures of integration sites in transgenic rice. *Plant J.*, **11**, 353–361.
10. Gorbunova,V. and Levy,A.A. (1997) Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. *Nucleic Acids Res.*, **25**, 4650–4657.
11. Tax,F.E. and Vernon,D.M. (2001) T-DNA-associated duplication/translocations in *Arabidopsis*. Implications for mutant analysis and functional genomics. *Plant Physiol.*, **126**, 1527–1538.
12. Lin,F.L., Sperle,K. and Sternberg,N. (1984) Model for homologous recombination during transfer of DNA into mouse L-cells—role for DNA ends in the recombination process. *Mol. Cell. Biol.*, **4**, 1020–1034.
13. Nassif,N., Penney,J., Pal,S., Engels,W.R. and Gloor,G.B. (1994) Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Mol. Cell. Biol.*, **14**, 1613–1625.
14. Puchta,H. (1999) Double-strand break-induced recombination between ectopic homologous sequences in somatic plant cells. *Genetics*, **152**, 1173–1181.
15. Knight,C.D., Cove,D.J., Cuming,A.C. and Quatrano,R.S. (2002) Moss gene technology. In Gilmartin,P.M. and Bowler,C. (eds.), *Molecular Plant Biology volume 2*. Oxford University Press, Oxford, pp. 285–301.
16. Hohe,A. and Reski,R. (2002) Optimisation of a bioreactor culture of the moss *Physcomitrella patens* for mass production of protoplasts. *Plant Sci.*, **163**, 69–74.
17. Hohe,A., Egener,T., Lucht,J.M., Holtorf,H., Reinhard,C., Schween,G. and Reski,R. (2004) An improved and highly standardised transformation procedure allows efficient production of single and multiple targeted gene-knockouts in a moss, *Physcomitrella patens*. *Curr. Genet.*, **44**, 339–347.
18. Egener,T., Granado,J., Guitton,M.-C., Hohe,A., Holtorf,H., Lucht,J.M., Rensing,S.A., Schlink,K., Schulte,J., Schween,G., Zimmerman,S., Duwenig,E., Rak,B. and Reski,R. (2002) High frequency of phenotypic deviations in *Physcomitrella patens* plants transformed with a gene-disruption library. *BMC Plant Biology*, **2**, 6.
19. Hohe,A., Decker,E.L., Gorr,G., Schween,G. and Reski,R. (2002) Tight control of growth and cell differentiation in photoautotrophically growing moss (*Physcomitrella patens*) bioreactor cultures. *Plant Cell Rep.*, **20**, 1135–1140.
20. Schlink,K. and Reski,R. (2002) Preparing high-quality DNA from moss *Physcomitrella patens*. *Plant Mol. Biol. Rep.*, **20**, 423a–423f.
21. Rougeon,F., Kourilsky,P. and Mach,B. (1975) Insertion of a rabbit beta-globin gene sequence into an *Escherichia coli* Plasmid. *Nucleic Acids Res.*, **2**, 2365–2378.
22. Land,H., Grez,M., Hauser,H., Lindenmaier,W. and Schutz,G. (1981) 5′-terminal sequences of eukaryotic messenger-RNA can be cloned with high-efficiency. *Nucleic Acids Res.*, **9**, 2251–2266.
23. Kamisugi,Y. and Cuming,A.C. (2005) The evolution of the abscisic acid-response in land plants: comparative analysis of group 1 LEA gene expression in moss and cereals. *Plant Mol. Biol.*, **59**, 723–737.
24. Schween,G., Egener,T., Fritzowsky,D., Granado,J., Guitton,M.C., Hartmann,N., Hohe,A., Holtorf,H., Lang,D., Lucht,J.M., Reinhard,C., Rensing,S.A., Schlink,K., Schulte,J. and Reski,R. (2005) Large-scale analysis of 73 329 *Physcomitrella* plants transformed with different

gene disruption libraries: production parameters and mutant phenotypes. *Plant Biol.*, **7**, 228–237.

25. Folger,K.R., Wong,E.A., Wahl,G. and Capecchi,M.R. (1982) Patterns of integration of DNA micro-injected into cultured mammalian-cells—evidence for homologous recombination between injected plasmid DNA-molecules. *Mol. Cell. Biol.*, **2**, 1372–1387.

26. Wernars,K., Goosen,T., Wennekes,L.M.J., Visser,J., Bos,C.J., Vandenbroek,H.W.J., Vangorcom,R.F.M., Vandenhondel,C.A.M.J.J. and Pouwels,P.H. (1985) Gene amplification in *Aspergillus nidulans* by transformation with vectors containing the amds gene. *Curr. Genet.*, **9**, 361–368.

27. Deshayes,A., Herreraestrella,L. and Caboche,M. (1985) Liposome-mediated transformation of tobacco mesophyll protoplasts by an *Escherichia coli* plasmid. *EMBO J.*, **4**, 2731–2737.

28. Schaefer,D., Zryd,J.P., Knight,C.D. and Cove,D.J. (1991) Stable transformation of the moss *Physcomitrella patens. Mol. Gen. Genet.*, **226**, 418–424.

29. Jongsma,M., Koornneef,M., Zabel,P. and Hille,J. (1987) Tomato protoplast DNA transformation-physical linkage and recombination of exogenous DNA-sequences. *Plant Mol. Biol.*, **8**, 383–394.

30. Ashton,N.W., Champagne,C.E.M., Weiler,T. and Verkoczy,L.K. (2000) The bryophyte *Physcomitrella patens* replicates extrachromosomal transgenic elements. *New Phytologist.*, **146**, 391–402.

31. Haber,J.E. (2000) Partners and pathways—repairing a double-strand break. *Trends Genet.*, **16**, 259–264.

32. Krogh,B.O. and Symington,L.S. (2004) Recombination proteins in yeast. *Annu. Rev. Genet.*, **38**, 233–271.

33. Markmann-Mulisch,U., Hadi,M.Z., Koepchen,K., Alonso,J.C., Russo,V.E.A., Schell,J. and Reiss,B. (2002) The organization of *Physcomitrella patens RAD51* genes is unique among eukaryotic organisms. *Proc. Natl Acad. Sci. USA*, **99**, 2959–2964.