Chromosome Research 1995, **3**, 335–345

Analysis and chromosomal localization of retrotransposons in sugar beet (*Beta vulgaris* L.): LINEs and *Ty1-copia*-like elements as major components of the genome

T. Schmidt, S. Kubis & J.S. Heslop-Harrison

Received 8 January 1995; received in revised form 10 February 1995
Accepted for publication by T.D. Allen 13 February 1995

DNA sequences of the reverse transcriptase gene of long terminal repeat (LTR) and non-LTR (non-viral) retrotransposons have been isolated and cloned from the genome of sugar beet (*Beta vulgaris*). Both retrotransposon types are highly amplified in sugar beet and may account for 2–5% of the genome. The BNR1 family, representing the first non-viral retrotransposon reported from a dicotyledonous species, shows homology to the mammalian L1 family of long interspersed repeated sequences (LINEs) and to retrotransposable elements from maize and lily. Sequences of the Tbv family are homologous to the *Ty1-copia* class of LTR retrotransposons. The BNR1 and Tbv retrotransposon families are characterized by sequence heterogeneity and are probably defective. The deduced peptide sequences were used to investigate the relation to other retroelements from plants, insects and mammals. Fluorescence *in situ* hybridization was used to investigate the physical distribution and revealed that both retrotransposon families are present on all sugar beet chromosomes and largely excluded from chromosomal regions harbouring the 18S–5.8S–25S rRNA genes. The BNR1 family is organized in discrete clusters, while the Tbv family of *Ty1-copia*-like retrotransposons shows a more uniform distribution along chromosome arms and is absent from some chromosomal regions. These contrasting distributions emphasize the differences in evolutionary amplification and dispersion mechanisms between the two types of retrotransposons. The *in situ* results of both elements reflect significant features of a higher order structure of the genome, as it is known for both short interspersed repeated sequences (SINEs) and LINEs in human.

Key words: *Beta vulgaris*, *in situ* hybridization, LINE, LTR retrotransposons, non-LTR (non-viral) retrotransposons, *Ty1-copia*

Introduction

Mobile genetic elements proliferating by reverse transcription of RNA intermediates have been described as retrotransposons and have been found in all eukaryotic organisms examined so far. Because of their structure, two classes of retrotransposons are distinguished, those flanked by long terminal repeats (LTR) and non-LTR retrotransposons. Non-LTR retrotransposons have also been called non-viral retrotransposons and contain open reading frames (ORFs), like LTR retrotransposons, encoding polypeptides with reverse transcriptase activity and DNA-binding capacity. A poly-A tail on the 3′ end and truncation at the 5′ end are characteristic features of non-viral retrotransposons. Well-studied examples include the elements R1Bm and R2Bm from *Bombyx moori* and *ingi* from *Trypanosoma brucei* (Burke *et al.* 1987, Kimmel *et al.* 1987, Xiong & Eickbush 1988, Luan *et al.* 1993). Long interspersed repeated elements (LINEs) are the best-known non-viral retrotransposons and have been detected dispersed through much of the genomes of mammalian species (Singer 1982). Approximately 100 000 copies of the L1Hs element, a typical LINE sequence, have been found in human genomes (Hutchinson *et al.* 1989). Only two non-viral retrotransposons have been described in plants so far. Cin4 is a moderately repeated sequence in the *Zea mays* genome (Schwarz-Sommer *et al.* 1987). A second non-viral retrotransposon discovered in plants is *del*2 from *Lilium speciosum* (Leeton & Smyth 1993). Both cin4 and *del*2 share many structural characteristics with mammalian LINEs. Furthermore, *del*2 is a particularly abundant element and contributes considerably to the large genomes of lily species (Bennett & Smith 1976, Leeton &

T. Schmidt (corresponding author), S. Kubis and J.S. Heslop-Harrison are at the Karyobiology Group, Department of Cell Biology, John Innes Centre, Colney Lane, Norwich NR4 7UH, UK. Tel: (+44) 1603 452571; Fax: (+44) 1603 451704; Email: schmidt@ bbsrc.ac.uk.

Smyth 1993). The amplification of *del2* indicates clearly that non-viral retrotransposons may form major components of plant genomes.

The two main types of viral or LTR retrotransposons, the *Ty1-copia-* and *Ty3-gypsy*-like elements, are not or only very distantly related to non-viral retrotransposons (Xiong & Eickbush 1990), and were originally discovered in *Saccharomyces cerevisiae* and *Drosophila melanogaster*. LTR retrotransposons are closely related to retroviruses but lack the *env* gene, which codes for a coat protein that enables a virus to escape from a eukaryotic cell. The *Ty1-copia-* and *Ty3-gypsy*-like retrotransposon types differ from each other in the order of genes within the internal domain. The first *copia*-like retrotransposons detected in plants were Bs1 from maize (Johns *et al.* 1985, 1989), Ta1 from *Arabidopsis thaliana* (Voytas & Ausubel 1988) and Tnt1 from *Nicotiana tabacum* (Grandbastien *et al.* 1989). They were discovered after transposition as inserted sequences causing restriction fragment length polymorphisms or gene inactivation. However, the majority of LTR retrotransposons in plants are defective, and there are only a few examples of their autonomous activation and transposition (Pouteau *et al.* 1991a, Hirochika 1993). Based on the polymerase chain reaction (PCR) using primers derived from conserved regions of the reverse transcriptase gene (*RT* gene), it was shown that *Ty1-copia* retrotransposons are ubiquitous and exist in most species investigated so far (Flavel *et al.* 1992a, Voytas *et al.* 1992). Recently, extensive studies have been performed, resulting in a detailed knowledge of the structural features of *copia*-like elements in plants (Hirochika *et al.* 1992, Manninen & Schulman 1993) and their relationship to other transposable elements. Copy numbers of LTR retrotransposons vary from a few to some hundred or thousand per genome (for review see Grandbastien 1992), and it has been shown that many related, but heterogeneous, retrotransposon families may exist in a genome (Flavell *et al.* 1992b).

Sequence repetition is a common feature of plant genomes, and numerous families of repeated DNA sequences contribute to the bulk of the nuclear DNA, leading to variation in genome size and complexity over several orders of magnitude (Bennett & Smith 1991). The crop *Beta vulgaris*, with cultivated forms such as sugar beet, fodderbeet, beet root and leaf beet, has a haploid genome size of about 750 Mbp (Arumuganathan & Earle 1991). Tetraploid, pentaploid and even octoploid forms have been observed in wild beet species, whereas many cultivars are diploid ($2n = 2x = 18$) or triploid. Estimation of the proportion of repetitive DNA revealed that the genome of sugar beet contains some 60% repeated sequences (Flavell *et al.* 1974). Previous studies have shown that most of the repetitive DNA in sugar beet is organized as satellite DNA in tandem arrays of 149 bp, 150–159 bp and 327 bp (Schmidt *et al.* 1991, Schmidt & Heslop-Harrison 1993). In addition, it has been found that the genomes of *Beta* species contain many different microsatellite families consisting of di-, tri- or tetranucleotide motifs (Schmidt *et al.* 1993) An objective of our work is the molecular and cytological characterization of the genome structure of sugar beet, and studies on the repetitive DNA fraction have now been extended to DNA sequences with a dispersed genomic organization.

Here we report that both LTR and non-LTR retrotransposons exist in the sugar beet genome. We have analysed conserved regions of *Ty1-copia* and LINE-like retrotransposons by PCR and show by fluorescence *in situ* hybridization the chromosomal organization of these retrotransposons in *B. vulgaris*.

Materials and methods

Isolation, restriction enzyme digestion and hybridization of nucleic acids

Genomic DNA was isolated from fresh leaf material of *Beta vulgaris* L. cv Rosamona (Quedlinburger Saatgut GmbH, Germany) as described previously (Schmidt & Heslop-Harrison 1993). Eight to ten micrograms of DNA was digested with an excess of restriction enzyme, separated on 1.1% agarose gels and transferred onto positively charged nylon membrane using standard techniques (Sambrook *et al.* 1989). Total RNA from leaves was isolated according to Napoli *et al.* (1990). Probes were labelled using the non-radioactive chemiluminescence system and Southern and Northern hybridization steps were carried out according to the manufacturer's protocol (Amersham).

Polymerase chain reaction and sequence analysis

PCR was carried out on a Perkin Elmer thermal cycler in 50-µl reaction volumes using 20 pmol of each primer and 50–100 ng of genomic DNA as template. Amplification of domains from the reverse transcriptase gene of *Ty1-copia* retrotransposons was performed following the method of Flavell *et al.* (1992b). The following primers were used for the amplification of non-viral retrotransposons: 5′-AA(A/G)CNTT(C/T)GA(C/T)AG-3′ and 5′-GC(G/A)TC(G/A)TCNGC(G/A)TA-3′. They were kindly provided by Dr H. Hirochika (National Institute of Agrobiological Resources, Tsukuba, Japan) and derived from the cin4 element of *Z. mays* (Schwarz-Sommer *et al.* 1987). After a 1-min initial denaturation step at 94°C, the reaction was subjected to 35 cycles, each comprising 1 min at 94°C, 1 min at 38°C and 1 min at 72°C, followed by a final elongation step of 5 min at 72°C. PCR products were gel-purified and cloned into pUC18. After transformation into *Escherichia coli* XL1-Blue cells, positive clones were screened by colony hybridization using labelled genomic DNA as probe. Both strands of the insert were sequenced on an automated sequencer (Pharmacia) using the dideoxy chain-termination procedure.

Chromosome preparation and *in situ* hybridization

Plasmids containing retrotransposon sequences were PCR labelled with digoxigenin-11-dUTP (Boehringer Mannheim) or biotin-11-dUTP (Sigma). Chromosome preparation and *in situ* hybridization were performed as described previously (Schmidt *et al.* 1994). Briefly, fixed root tips were digested in an enzyme mixture containing cellulase and pectinase and squashed on slides. The hybridization mixture containing the probe was added to pretreated slides and covered with a coverslip. Slides were denaturated in an Omnislide *in situ* hybridization machine (Hybaid) at 70°C for 8 min, and the temperature was then gradually decreased to 37°C. After hybridization slides were washed and the detection reaction performed. Digoxigenin-labelled probes were detected with FITC (fluorescein isothiocyanate)-conjugated sheep anti-digoxigenin antibody (Boehringer Mannheim), while biotin-labelled probes were detected with streptavidin conjugated with CY3 (Sigma). Slides were counterstained with DAPI (4',6-diamidino-2-phenylindole), mounted in antifade solution and examined with a Leitz epifluorescence microscope with filter sets A, I2/3 and N2.1. Photographs were taken on colour print film, digitized to photo CD and printed using Adobe Photoshop after contrast optimization of the whole image.

Computer analysis

For homology search within the EMBL protein database the FASTA program of the GCG package (release 83) was used. The alignment was manually optimized. Relatedness of deduced peptide sequences was calculated with the Molecular Evolutionary Genetics Analysis program MEGA (version 1.0) (Kumar *et al.* 1993). Each putative peptide sequence was entered as an operational taxonomic unit and the pairwise deletion option was chosen to compute gaps and insertions. The estimation was then conducted using the neighbour joining algorithm. Nucleotide sequence alignments were performed with the program PILEUP of the GCG package using a gap penalty of 1.0 and a gap length penalty of 0.1.

Results

Isolation and characterization of non-viral retrotransposons

The isolation of retrotransposon sequences from sugar beet was based on the amplification of parts of the reverse transcriptase gene. The primers were short and degenerated, and hence a low annealing temperature was used for the PCR. Each reaction gave several size classes of PCR products ranging from 250 to 800 bp, which were separated and purified by gel electrophoresis and separately cloned into pUC18. Positive clones were selected according to their signal strength after colony hybridization using genomic DNA as a probe to find members of any families of putative non-viral retrotransposons representing a substantial component of the sugar beet genome. After sequencing of the PCR products, predicted peptide sequences were used to query nucleic acid and protein database searches. Three clones showed a striking homology to the retrotransposase domain of cin4, a LINE-like non-viral retrotransposon from *Z. mays*. These clones were designated BNR1-1, BNR1-2 and BNR1-6. The epithet BNR was chosen for *b*eet *n*on-*v*iral *r*etrotransposon. The sequences of BNR1-1, BNR1-2 and BNR1-6 are 310 bp, 324 bp and 317 bp long, respectively (Figure 1a). The inserts of BNR1-1 and BNR1-2 show 91% nucleotide homology and 92% identity on the level of the predicted peptide sequence. A higher divergence was observed for BNR1-6, with 59% nucleotide identity and 52% peptide homology with respect to BNR1-1, suggesting that BNR1-6 belongs to a related subfamily. Putative stop codons were found in all reading frames of the BNR1 clones. Since the stop codons were present at different positions, they might be the result of individual mutation events. However, one stop codon was found at the same position in BNR1-1 and BNR1-2, indicating the probable loss of transposition capability of a group of BNR1 sequences. Figure 2a shows the alignment of the deduced peptide sequences of BNR1-1, BNR1-2 and BNR1-6 with reverse transcriptase regions of the plant retrotransposons cin4 and *del*2 and members of the L1 family from human and mouse. In general, most of the identity was found at positions which were conserved in the majority of the compared retroelements. The introduction of frameshifts was necessary in all BNR1 clones to optimize the alignment. A remarkable conservation of the motif R-G-X_1-R-Q-G-D/C-P-L-S-P-X_1-L-F was observed in half of the compared peptide sequences, and seven amino acid residues were found to be conserved in plant LINEs only. The best alignment was achieved with the reverse transcriptase of the cin4 element. Members of the BNR1 family share 40.8–42.8% of the amino acid residues with the cin4 reverse transcriptase domain. The homology of BNR1 to the ORF2 found in *del*2 from *L. speciosum* is slightly lower than to the ORF2 of the L1 element from human and mouse, but still in a similar range, and varies from 21.4% to 27.5%.

Southern hybridization to sugar beet DNA digested with different restriction enzymes allowed the investigation of the genomic organization of the BNR1 family (Figure 3a). Hybridization to *Dra*I-digested DNA resulted in a smear over the whole track, indicating the presence of BNR1 in many different genomic loci, presumably as an interspersed DNA sequence. A similar pattern, but with stronger hybridization to DNA fragments between 9 kb and 23 kb, was observed in *Eco*RI and *Bam*HI digests. The occurrence of some stronger fragments indicates the clustered organization of BNR1 elements. The intensity of the hybridization signals suggests that the BNR1 sequence is highly amplified in the genome of sugar beet. Northern ana-

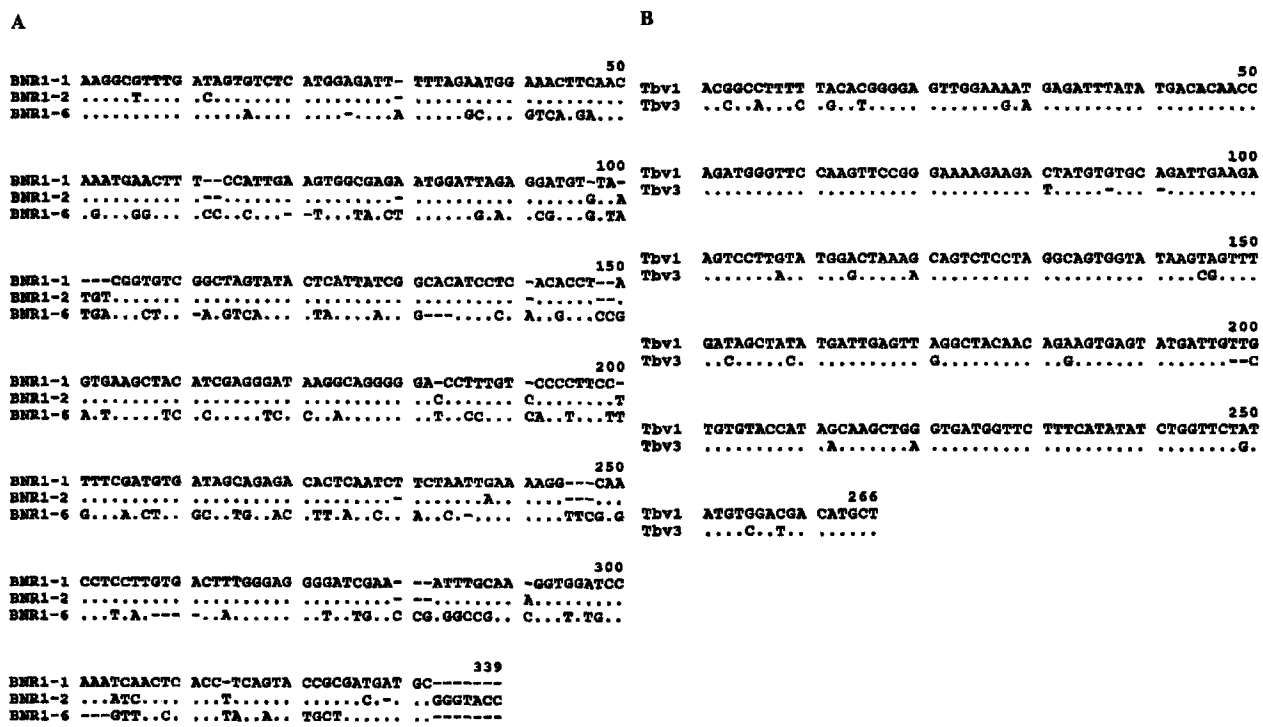

Figure 1. Nucleotide alignment of non-viral retrotransposon subclones BNR1-1, BNR1-2 and BNR1-6 (**A**) and LTR-retrotransposon subclones Tvb1 and Tbv3 (**B**). Identical nucleotides are shown by dots. Gaps introduced for optimal alignment are shown by dashes. The nucleotide sequences appeared in the EMBL, GenBank and DDBJ Nucleotide Sequence Database under the accession numbers Z38073 (BNR1-1), Z38074 (BNR1-2), Z38075 (BNR1-6), Z38076 (Tbv1) and Z38077 (Tbv3).

lysis did not reveal the existence of BNR1 transcripts (not shown).

Isolation and characterization of LTR retrotransposons

In addition to the LINE-like BNR1 family, the sugar beet genome contains LTR retrotransposons of the *Ty1-copia* type, as shown by PCR using primers specific for a conserved region within the *RT* gene. After PCR the expected products of approximately 260 bp were purified and cloned. Several inserts were subjected to sequencing and two clones, designated Tbv1 and Tbv3, were analysed in detail. Tbv1 and Tbv3 are 266 bp and 262 bp long and show 90% homology at both nucleotide and predicted peptide sequence levels (Figure 1b).

We assume that most of the *Ty1-copia* retrotransposons are defective in sugar beet since putative stop codons were found in both Tbv clones. Furthermore, the introduction of frameshifts was necessary in Tvb3 to enable an alignment with peptide sequences of reverse transcriptase domains from other LTR retrotransposons. We compared Tbv1 and Tbv3 with the tobacco element Tnt1 (Grandbastien *et al.* 1989), Ta1 from *A. thaliana* (Voytas & Ausubel 1988), TosRT1 from *Oryza sativa* (Hirochika *et al.* 1992), M143 from *Solanum tuberosum* (Flavell *et al.* 1992b) and retrotransposons from *Spinacea oleracea* (Hirochika & Hirochika 1993). The *copia* element from *D. melanogaster* was also included. Sequence similarities are shown in Figure 2b. Divergence was observed between all examined peptide sequences, although short invariant regions or single residues were found. The reverse transcriptase that shares the highest degree of similarity to the Tbv family is TosRT1 (74.6% homology), a *Ty1-copia*-like sequence of rice. Spinach and sugar beet are closely related species belonging to the same family Chenopodiaceae. However, the *Ty1-copia* retrotransposons from spinach and beet exhibited a considerable heterogeneity in the reverse transcriptase domain and they are less than 50% homologous to each other (45.5% and 48.1% amino acid identity).

Southern experiments demonstrated the repetitive nature of the *Ty1-copia* retrotransposons in sugar beet (Figure 3b). The Tbv family hybridizes with many differently sized *Dra*1 fragments, resulting in a smear with the strongest signals between 3 kb and 9 kb. The hybridization to *Eco*RI- and *Bam*HI-digested DNA was strongest with fragments from 10 kb up to 23 kb. In addition, a 2.8-kb *Eco*RI fragment was detected, indicating the conservation of an internal part of the Tbv elements. The observed hybridization patterns are indicative of a dispersed distribution of these retroelements, and the strength of hybridization reflects the high copy number of the Tbv family within the beet genome. Northern analysis and a PCR assay using cDNA templates failed to reveal Tbv transcripts or amplification products.

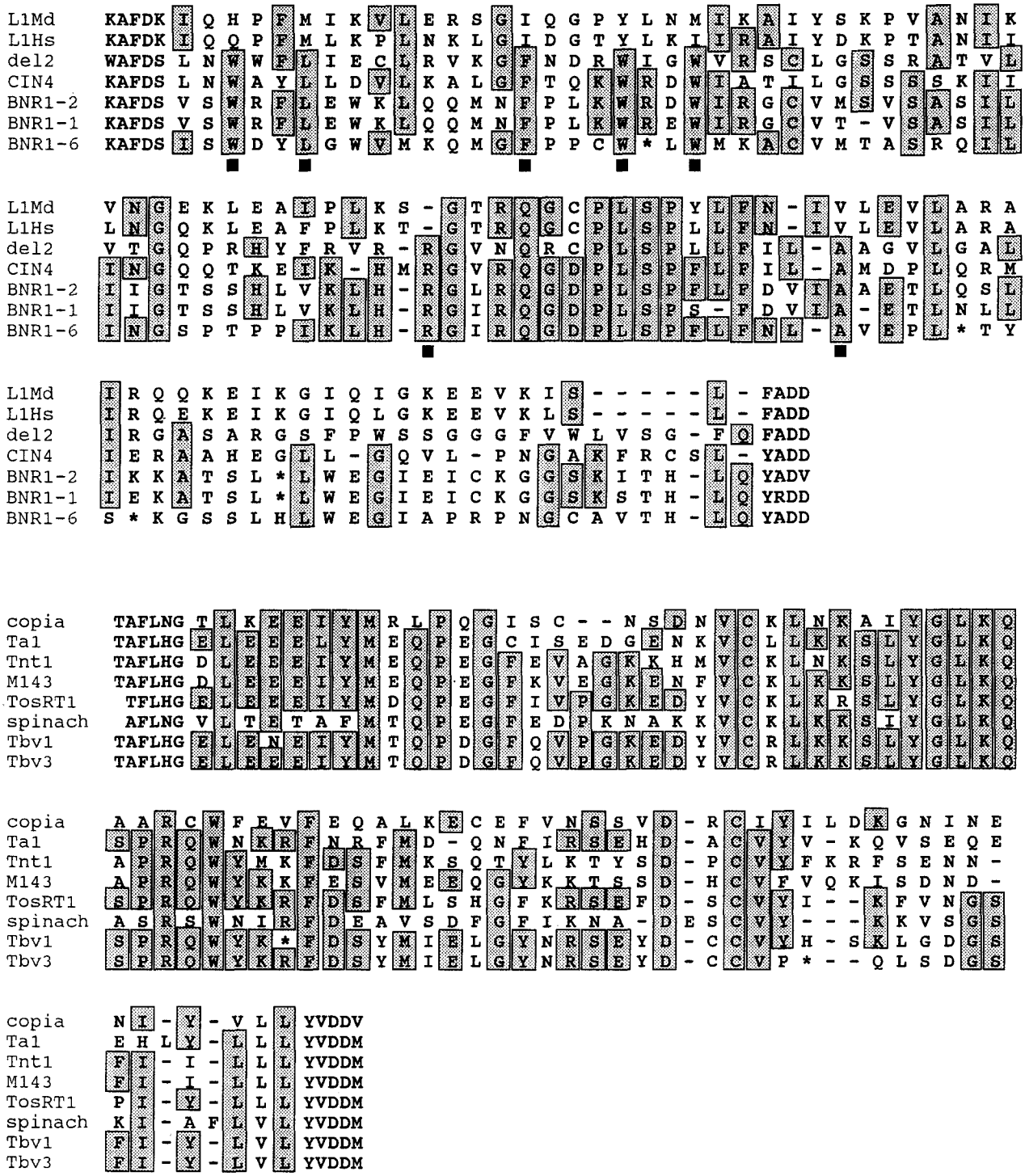

Figure 2. Alignment of predicted peptide sequences of different retrotransposons. Dashes show gaps which were introduced to optimize the alignment. Stop codons are marked by asterisks. **A** Alignment of sugar beet elements BNR1-1, BNR1-2 and BNR1-6 with reverse transcriptases of non-viral retrotransposons from mouse (L1Md), human (L1Hs), *L. speciosum* (*del2*), *Z. mays* (cin4). Homologous amino acid residues present in at least two non-viral retroelements and at least one sugar beet sequence are boxed and shaded. Residues conserved in all plant elements are indicated by filled squares. **B** Alignment of sugar beet sequences Tbv1 and Tbv3 with reverse transcriptases of LTR retrotransposons from *D. melanogaster* (*copia*), *A. thaliana* (Ta1), *N. tabacum* (Tnt1), *S. tuberosum* (M143), *O. sativa* (TosRT1) and *S. oleracea* (spinach). Homologous amino acids conserved in at least four *Ty1-copia* retrotransposons including at least one sugar beet sequence are boxed and shaded.

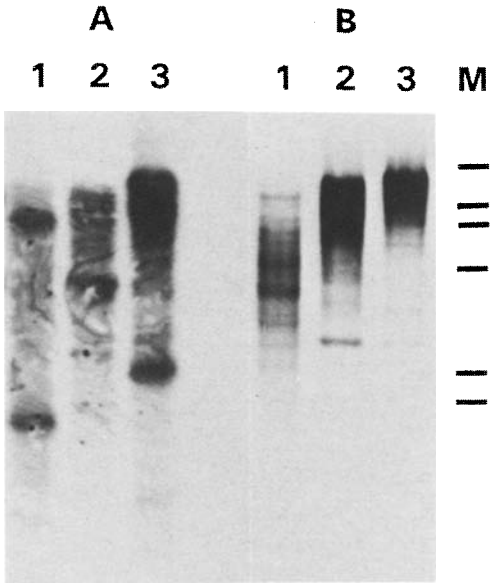

Figure 3. Genomic organization of retrotransposon families in *B. vulgaris*. Southern blots of genomic DNA digested with *Dra*I (1), *Eco*RI (2) and *Bam*HI (3) were probed with the BNR1 sequence (**A**) and the Tbv sequence (**B**). DNA size marker (m) from top: 23.1, 9.4, 6.6, 4.4, 2.3 abd 2.0 kb, respectively

Relationship between retrotransposons

Based on the number of amino acid differences, we calculated the relatedness between reverse transcriptases of different non-LTR and LTR retrotransposons. Primer sequences were excluded from the calculation. The resulting dendrogram is shown in Figure 4. It has two main branches corresponding to two different types of retroelements, namely LTR and non-LTR retrotransposons. All non-LTR or non-viral retrotransposons examined are ordered in the lower branch of the tree. The closest relation of the BNR1 family was found to cin4 from *Z. mays*. BNR1-1 and BNR1-2 form a subgroup; the considerable divergence of BNR1-6 resulted in an extra branch. The second plant non-LTR retrotransposon, *del*2, was found to be less related to the BNR1 family than the LINE group represented by L1 families from human and mouse. It was placed more distantly from the BNR1 group than the human and mouse LINEs, reflecting the fact that several residues were found to be conserved in mammalian LINE sequences and clones of the sugar beet BNR1 family, but not in *del*2.

The heterogeneity of *Ty1-copia*-like reverse transcriptase sequences is illustrated in the dendrogram. Both Tbv sequences are grouped on one arm of the tree. Within the compared LTR retrotransposons, relatively low homology of Tbv1 and Tbv3 was found to *copia* from *D. melanogaster* (37.1 and 38.4%). Therefore, *copia* was placed separately from the remaining plant reverse transcriptase sequences, but still in the same half of the tree. The spinach sequence was ranked next to *copia* and demonstrates the observed divergence to Tbv1 and Tbv3. Allocated on separate branches, but closely related to the sugar beet sequences, are Ta1 from *A. thaliana* and the retrotransposon sequence M143 from potato and Tnt1 from tobacco. M143 and Tnt1 are similar enough to be placed in one group.

Both the Tbv and the BNR1 family show the maximum possible separation within the dendrogram. Although the BNR1 and the Tbv sequences have analogous functional characteristics in the reverse transcriptase domain, both groups of retrotransposons represent different evolutionary lineages consistent with the analysis of Xiong & Eickbush (1990). Therefore, it can be concluded that the genome of sugar beet or of a sugar beet progenitor was invaded by different retrotransposable sequences represented today by the BNR1 and Tbv family.

Chromosomal localization of retrotransposons in *B. vulgaris*

Fluorescence *in situ* hybridization was used to investigate the physical distribution of retrotransposon families along the chromosomes. *In situ* hybridization showed that the BNR1 and Tbv family were present on both arms of all but one pair of beet chromosomes (Figure 5f–h). Both elements were largely excluded from the centromeric, subtelomeric and nucleolar organizer regions of the beet genome, but double *in situ* hybridization showed only low tendency for the two

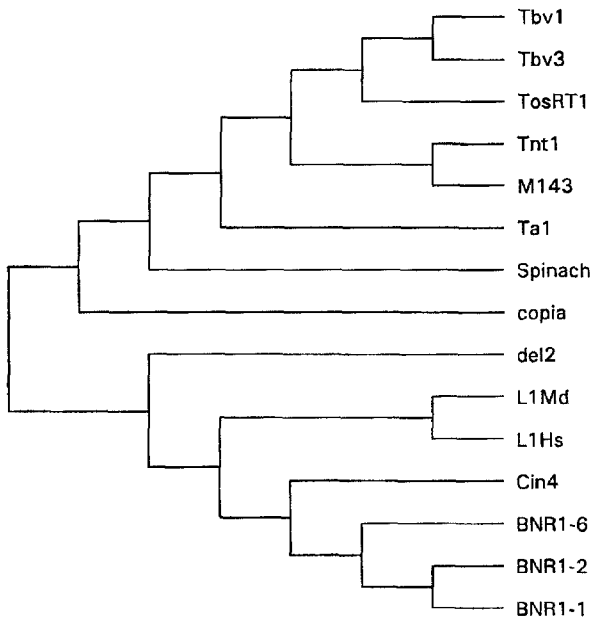

Figure 4. Dendrogram showing the relation of the sugar beet retrotransposon families BNR1 and Tbv to other retroelements from plants, mammals and insects (for abbreviations refer to Figure 2). The relatedness of putative peptide sequences was estimated using a neighbour-joining algorithm and resulted in an unrooted tree with two major branches representing viral and non-viral retroelements. The calculation was performed with the program Molecular Evolutionary Genetics Analysis (version 1.0).

elements to be co-localized. However, several DAPI-positive blocks of heterochromatin contained both BNR1 and Tbv sequences. Most copies of the BNR1 element were present in discrete clusters seen as bands of hybridization signals on both chromatids of the chromosomes (Figure 5a, b, f & g). There were more than 50 such clusters on the diploid chromosome complement, and five or more sites on some chromosome arms. At interphase, many of the BNR1 hybridization signals coincided with brightly DAPI staining points, away from the nucleoli, although there were also many unassociated DAPI positive sites (Figure 5c–e). At metaphase, almost no hybridization signals of BNR1 and Tbv sequences were observed on the short arm of chromosome 1, which consists mainly of tandem arrays of the 18S–5.8S–25S rRNA genes (Figure 5f-h). In general, the Tbv family showed a more uniform distribution along chromosome arms than BNR1 elements, although bands of hybridization signal were still detected.

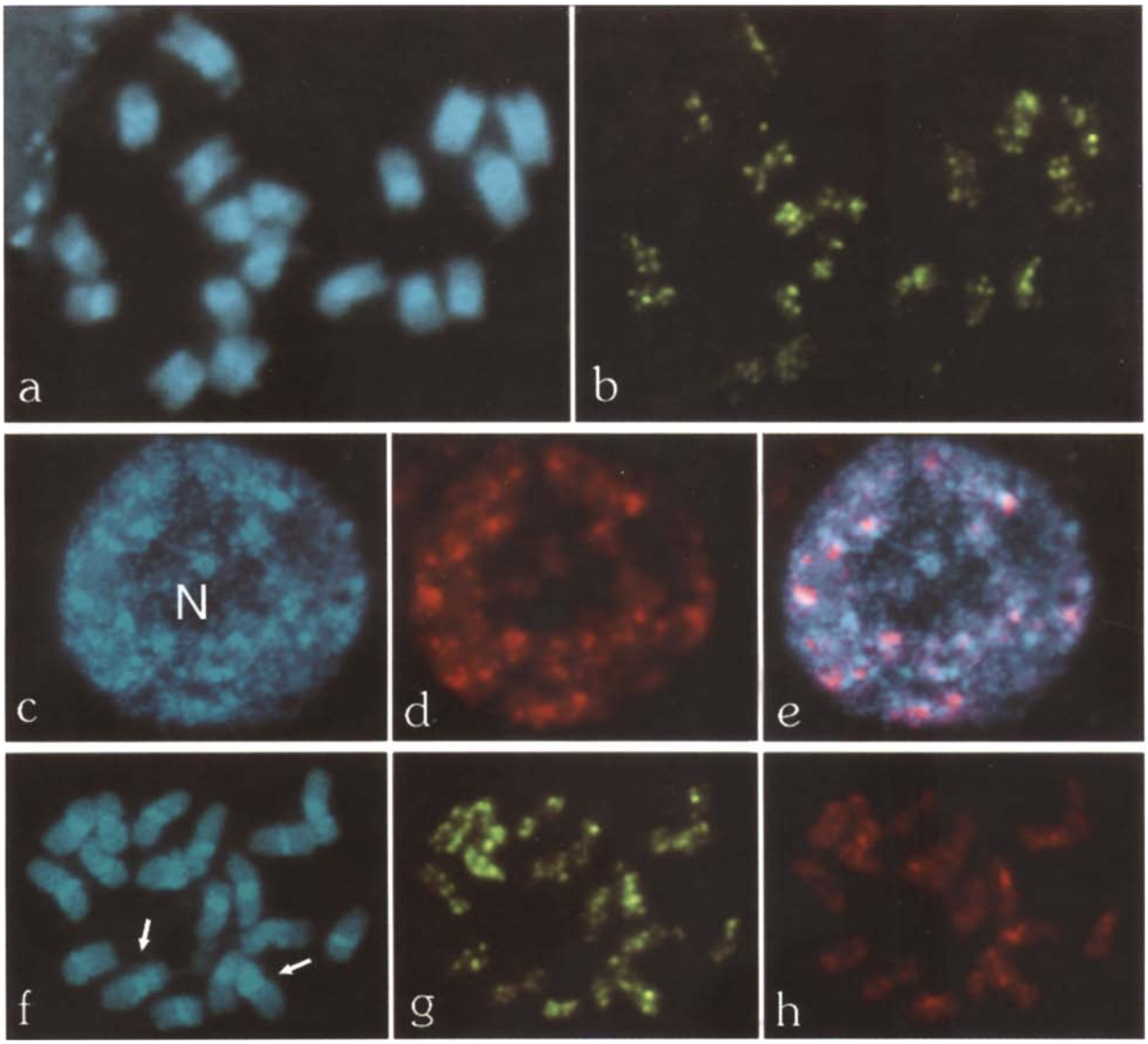

Figure 5. Localization of retrotransposon sequences along chromosomes of sugar beet root tips by fluorescence *in situ* hybridization. **a, c & f** The chromosomes at metaphase or interphase were stained with DAPI (blue). **b & d** Clustered organization of the non-viral retrotransposon family BNR1 at metaphase (yellow–green fluorescence) and interphase (red fluorescence); N = nucleolus. **e** Computerized overlay of **c** and **d** showing limited co-localization of LINEs with DAPI-positive regions. **g & h** Simultaneous detection of the non-viral retrotransposon BNR1 (yellow–green fluorescence) and the LTR retrotransposon Tbv (red fluorescence) on metaphase chromosomes. The nucleolar organizing region on the short arm of chromosome 1 (arrowed in **f**) contains only few BNR1 and Tbv sequences.

Discussion

We have isolated repetitive DNA elements of sugar beet (*B. vulgaris*) with homology to the internal region of two retrotransposon classes, namely viral and non-viral retrotransposons (Figure 1). The elements were localized on sugar beet chromosomes using fluorescence *in situ* hybridization, and the predicted peptide sequences of the reverse transcriptase domain were used to investigate the relation to conserved regions of other retroelements from plants, insects and mammals (Figures 2 & 5).

Recently, transposable elements with structural homologies to the *Ty1* and *copia* retrotransposons from yeast and *D. melanogaster* have been isolated from plant genomes and described in detail (Moore *et al.* 1991, Hirochika *et al.* 1992, Manninen & Schulman 1993). Several PCR-based studies have been undertaken to investigate the general occurrence of this class of retroelements in plants (Flavell *et al.* 1992a, Hirochika & Hirochika 1993). Voytas *et al.* (1992) showed that LTR retrotransposons of the *Ty1-copia* type are ubiquitous in plant genomes and present in nine out of ten plant divisions, including bryophytes, lycopods, ferns, gymnosperms and angiosperms. However, only a few investigations have been performed to gain information about the genomic organization and distribution of LTR retrotransposons in plants. A barley retrotransposon-like sequence, BIS-1, which constitutes at least 5% of the barley genome, shows quite uniform hybridization along all barley chromosome arms (Moore *et al.* 1991). A less uniform distribution with absence, or presence at a reduced density, at some chromosomal regions, in particular at centromeres, was observed for the *copia*-like Tbv family of sugar beet. A similar distribution over all chromosomes with reduced copy numbers at some chromosomal loci was observed for a family of *Ty1-copia* retrotransposons from *Pinus elliottii* Engelm. *elliottii* (A. Kamm *et al.*, in preparation). The widespread distribution over all chromosomes with locus-dependent variation reflects presumably the general organization of *Ty1-copia* retrotransposon families in plant genomes.

In contrast, relatively little is known about non-viral retrotransposons or LINEs in plants, and the chromosomal distribution of plant non-viral retroelements has not yet been investigated by *in situ* hybridization. Only two such elements, differing enormously in redundancy, have been isolated so far. Cin4 was observed as an insertion causing the inactivation of the maize A1 gene, and Schwarz-Sommer *et al.* (1987) have shown that the maize genome (5500 Mbp) contains only 50–100 copies of cin4. The lily retroelement *del*2 is a highly amplified LINE, and Leeton & Smyth (1993) estimated that 250 000 copies of *del*2 exist in the large genome of *L. speciosum*. The genome of sugar beet is about 45 times smaller (750 Mbp) than that of *L. speciosum*. However, our results show that the relatively small genome of *B. vulgaris* contains a high proportion of non-viral retrotransposons.

In situ hybridization revealed that the LINE-like family BNR1 is non-randomly distributed over all chromosomes of sugar beet. The genomic organization of BNR1 elements is characterized by the occurrence of distinct clusters which are clearly visible on both chromatids (Figure 5a, b, f & g). The clustered organization of BNR1 sequences in many discrete regions was also indicated from Southern blot hybridization, which resulted in the detection of several strongly hybridizing fragments superimposed on a smear (Figure 3a). Molecular and cytological studies have shown that mammalian L1 elements are not evenly distributed over all chromosomes and predominantly integrate into late-replicating regions of the genome (Chen & Manuelidis 1989, Taruscio & Manuelidis 1991). This genomic distribution pattern is highly specific, and L1Md elements have been used as probes for *in situ* hybridization to identify individual mouse chromosomes (Boyle *et al.* 1992). Korenberg & Rykowski (1988) demonstrated that the LINE L1Hs family, with SINEs (short interspersed repeated sequences) or Alu sequences the main class of dispersed DNA in human, is not randomly scattered. Sequences of L1Hs were detected in clusters along human metaphase chromosomes and occur predominantly in Giemsa- and Quinacrine-positive bands which are rich in adenine and thymine residues. The genomic organization of the *B. vulgaris* BNR1 family is similar to that of the L1Hs sequences in human, which were found to be enriched in numerous regions. However, a predominant occurrence in AT-rich regions of the sugar beet genome was not observed, although co-localization with some DAPI-positive regions was visible (Figure 5e). There are a number of AT-rich satellite sequences in beet (T. Schmidt, in preparation) which have a characteristic genomic location in most of the DAPI-positive bands, and it is possible that the BNR1 retrotransposons are excluded from the large genomic regions occupied by these tandemly repeated sequences. This finding is consistent with the observation in human that regions containing the centromeric alpha satellite repeat and the satellite DNAs I, II and IV show a strongly reduced number of LINEs (Korenberg & Rykowski 1988).

Examination of interphase nuclei revealed the general exclusion of BNR1 retroelements from the NOR (Figure 5c–h). This absence of both the BNR1 and Tby family, as shown by double *in situ* hybridization, from the short arm of chromosome 1 harbouring the 18S–5.8S–25S rRNA genes is remarkable. Similarly, the absence of the LINE family L1 from rDNA sites of some human chromosomes has been reported (Korenberg & Rykowski 1988). However, the exclusion from regions consisting of ribosomal genes is not a general feature of non-viral retrotransposons since the mobile elements R1 and R2 insert specifically into the 28S ribosomal RNA gene of many insect species by a mechanism which has been recently discovered for the R2 retro-

transposition (Jakubczak *et al.* 1991, Luan *et al.* 1993). Nevertheless, the exclusion of retroelements from rDNA loci and genomic regions occupied by major tandem repeats at centromeres and blocks of intercalary DAPI-positive heterochromatin is a clear indication of constraints on where BNR1 clusters or Tbv elements can occur, indicating that retrotransposon distribution patterns reflect a higher order of genome organization and nuclear architecture in sugar beet. Specific genomic regions or target sites might be less accessible for retrotransposons resulting from modulation of chromatin structure or DNA–protein assembly (Pryciak & Varmus 1992), which is genetically determined (Dorn *et al.* 1993).

Knowledge about genomic regions which are either preferred or less accessible sites for retroelement integration might have implications for genome mapping and gene isolation by transposon tagging (Meyer *et al.* 1994, Pouteau *et al.* 1991b). The rapid progress in plant genome mapping and gene isolation has led to an understanding of function and developmental regulation of genes and their products (Meyerowitz 1994) which is far ahead of our knowledge of genome organization and architecture. Most plant genomes consist of sequence motifs which are repeated up to hundreds to hundreds of thousands of times, and detectable diversification in sequence, copy number and genomic localization occurs over a plant breeding time scale as well as over evolutionary time during speciation. Many different types contribute to the amount of the repeated DNA, and the study of genome organization of repetitive DNA is essential to gain information about the nuclear architecture of plant genomes and to understand their gross structures and organization. Previously, we have analysed several tandemly repeated sequences and microsatellites from sugar beet (Schmidt *et al.* 1991, 1993, Schmidt & Heslop-Harrison 1993). By comparison with the data from our satellite DNA analyses, we estimate that BNR1 and *Ty1-copia* retroelements occur in a similar order of magnitude and might account for up to 2–5% of the sugar beet genome. The retroelement families BNR1 and Tbv represent a major fraction of the dispersed DNA sequences, contribute considerably to the amount of repetitive DNA and are important for the understanding of the genome structure of *B. vulgaris*.

High copy number sequences cannot be mapped easily by segregation analysis, and it is difficult to assess transposition and mobility of the two retrotransposon families studied here. In Northern analyses, no transcripts have been found in RNA from green leaves, and PCR with sugar beet cDNA as template failed to generate specific products. Furthermore, we observed putative stop codons and/or disrupted reading frames, and it was necessary to introduce frameshifts in four of the five retrotransposon sequences analysed to find the correct alignment with known LTR and non-LTR retrotransposons. Therefore, it is likely that most of BNR1 and Tbv copies of *B. vulgaris* are defective in terms of retrotransposition, and we assume that only a small number of both retroelements are capable of generating new copies. These findings are consistent with data from other retroelements. In human and rodents, LINEs of the L1 family are one of the most abundant sequences. However, the vast majority of these sequences is transcriptionally inactive, and it is assumed that most of the members were derived from only a few 'master genes' (for review see Deininger *et al.* 1992). According to the 'master gene model', a few genes or elements generate many copies which are themselves pseudogenes or, in the particular case of LINEs, defective elements. Mutations within the master or source element during evolution give rise to the amplification of derivates and to the formation of subfamilies. This evolutionary model can be applied to the LINE family BNR1 of sugar beet, and suggests that BNR1-6 is in fact a member of a diverged subfamily.

Only a few examples of active LTR retrotransposons have been reported in plants (reviewed by Flavell 1992), and the transcription and transposition of *Ty1-copia* retrotransposons Tnt1 and Tto1 from tobacco has been observed only under tissue culture conditions (Pouteau *et al.* 1991b, Hirochika 1993). Spreading of retrotransposon sequences at the DNA level followed by fixation, as observed for other repetitive sequences and caused by recombination events such as gene conversion or unequal crossing over, has been discussed as an alternative mode of retroelement amplification (Wichmann *et al.* 1992). However, those events have not been observed yet, although the incorporation of a cellular gene into a maize LTR retrotransposon indicates the involvement of retrotransposons in recombination events (Bureau *et al.* 1994).

A positive correlation between the very low rates of transcription and transposition and sequence heterogeneity of plant *Ty1-copia* retrotransposons has been proposed by Flavell (1992). The *B. vulgaris* subclones Tbv1 and Tbv3 share 90% homology at both nucleotide and predicted peptide sequence level, but they show more than 50% divergence from the *S. oleracea Ty1-copia* sequence reported by Hirochika & Hirochika (1993). Hence, their *Ty1-copia* sequences were distantly arranged within the dendrogram (Figure 4), although both species belong to the family Chenopodiaceae. Assuming a vertical transmission of *Ty1-copia* retroelements, this implies that a considerable sequence diversification occurred during the speciation of spinach and beet. Similarly, no cross-hybridization between the *A. thaliana* retroelement Ta1 and genomic DNA from *B. vulgaris* has been detected in Southern experiments (Grandbastien *et al.* 1989), indicating that, despite maintenance of the overall structure, *Ty1-copia* retrotransposons are subject to sequence divergence leading to species-specific variants. It is likely that the heterogeneity is mainly caused during the synthesis of new copies by reverse transcription. Reverse transcriptases lack a proofreading function, and thus generate nucleotide changes at a rate up to 10^6 higher than observed for DNA

replication of cellular genes (Holland *et al.* 1982, Gojobori & Yokoyama 1985). Moreover, it has often been observed that several heterogeneous retrotransposon subfamilies exist within the genome of sugar beet, and some of them might show a greater homology to the spinach sequence. Konieczny *et al.* (1991) have investigated nine families from *A. thaliana* related to the originally isolated retrotransposon Ta1, and Flavell *et al.* (1992b) have examined more than 30 *Ty1-copia* clones from potato which clearly could be grouped into six subfamilies. In contrast to yeast and *D. melanogaster*, plant *Ty1-copia* retrotransposons are very heterogeneous, and copies with higher homology are often found in genomes of different species. This suggests strongly the possibility of horizontal transfer of retrotransposons between species (Kidwell 1992).

The heterogeneity within the BNR1 family of non-viral retrotransposons is also considerable. BNR1-6 shows less than 60% homology to BNR1-1 and BNR1-2, which themselves are more than 90% identical, indicating the existence of subfamilies. However, a striking homology of all BNR1 elements to non-viral retrotransposons of plants and to the L1 family of long interspersed repeated sequences from human and mouse was observed. Nevertheless, significant differences between plant and mammalian LINEs were found. We noted several amino acid residues within the peptide sequences which were conserved in all plant LINEs, and, furthermore, many additional sites showing only little and hence probably random variation.

The BNR1 family represents the first non-viral retrotransposon isolated from a dicotyledonous plant. Non-viral retrotransposons were found in monocotyledonous plants species only, and Leeton & Smyth (1993) showed that counterparts of the lily retrotransposon *del2* are present in several other monocotyledonous species. We assume that, in addition to LTR retrotransposons, non-viral retrotransposons or LINEs are ubiquitous in plants and contribute to the amount of excess DNA observed in their nuclear genomes.

Acknowledgements

T. Schmidt was supported by grants from the Deutscher Akademischer Austauschdienst and currently holds an EU fellowship (ERB4001GT931647). S. Kubis acknowledges financial support by a COMETT fellowship. We thank Dr H. Hirochika (National Institute of Agrobiological Resources, Tsukuba, Japan) for providing non-LTR retrotransposon primers.

References

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

Bennett MD, Smith JB (1976) Nuclear DNA amounts of angiosperms. *Proc R Soc Lond B* **274**: 227–274.

Bennett MD, Smith JB (1991) Nuclear DNA amounts of angiosperms. *Proc R Soc Lond B* **334** 309–345.

Boyle AL, Feltquite DM, Dracopoli NC, Housman DE, Ward DC (1992) Rapid physical mapping of cloned DNA on banded mouse chromosomes by fluorescence *in situ* hybridization. *Genomics* **12**: 106–115.

Bureau TE, White SE, Wessler SR (1994) Transduction of a cellular gene by a plant retroelement. *Cell* **77**: 479–480.

Burke WD, Calalang CC, Eickbush TH (1987) The site-specific ribosomal insertion element type II of *Bombyx moori* (R2Bm) contains the coding sequence for a reverse transcriptase-like enzyme. *Mol Cell Biol* **7**: 2221–2230.

Chen TL, Manuelides L (1989) SINEs and LINEs cluster in distinct DNA fragments of Giemsa band size. *Chromosoma* **98** 309–316.

Deininger PL, Batzer MA, Hutchinson II CA, Edgell MH (1992) Master genes in mammalian repetitive DNA amplification. *Trends Genet* **8**: 307–311.

Dorn R, Krauss V, Reuter G, Saumweber H (1993) The enhancer of position-effect variegation of *Drosophila*, E(var)3-93D, codes for a chromatin protein containing a conserved domain common to several transcriptional regulators. *Proc Natl Acad Sci USA* **90**: 11376–11380.

Flavell AJ (1992) *Ty1-copia* group retrotransposons and the evolution of retroelements in the eukaryotes. *Genetica* **86**: 203–214.

Flavell RB, Bennett MD, Smith JB (1974) Genome size and the proportion of repeated nucleotide sequence DNA in plants. *Biochem Genet* **12**: 257–269.

Flavell AJ, Dunbar E, Anderson R, *et al.* (1992a) *Ty1-copia* group retrotransposons are ubiquitous and heterogenous in higher plants. *Nucleic Acids Res* **20**: 3639–3644.

Flavell AJ, Smith DB, Kumar A (1992b) Extreme heterogeneity of *Ty1-copia* group retrotransposons in plants. *Mol Gen Genet* **231**: 233–242.

Gojobori T, Yokoyama S (1985) Rates of evolution of the retroviral oncogene of Moloney murine sarcoma virus and of its cellular homologues. *Proc Natl Acad Sci USA* **82**: 4198–4201.

Grandbastien MA (1992) Retroelements in higher plants. *Trends Genet* **8**: 103–108.

Grandbastien MA, Spielmann A, Caboche M (1989) Tnt1, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. Nature 337: 376–380.

Hirochika H (1993) Activation of tobacco retrotransposons during tissue culture. *EMBO J* **12**: 2521–2528.

Hirochika H, Fukuchi A, Kikuchi F (1992) Retrotransposon families in rice. *Mol Gen Genet* **233**: 209–216.

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

Holland J, Spindler K, Horodyski F, Grabau E, Nichol S, VandePol S (1982) Rapid evolution of RNA genomes. *Science* **215**: 1577–1585.

Hutchinson CA, Hardies SC, Loeb DD, Shehee WR, Edgell MH (1989) LINEs and related retroposons: long interspersed repeated sequences in the eukaryotic genome. In: Berg DH, Howe MM, eds. *Mobile DNA*. Washington, DC: American Society of Microbiology, pp 593–617.

Jakubczak JL, Burke WD, Eickbush TH (1991) Retrotransposable elements *R1* and *R2* interrupt the rRNA genes of most insects. *Proc Natl Acad Sci USA* **88**: 3295–3299.

Johns MA, Mottinger J, Freeling M (1985) A low copy number, *copia*-like transposon in maize. *EMBO J* **4**: 1093–1102.

Johns MA, Babcock MS, Fuerstenberg SM *et al.* (1989) An

unusually compact retrotransposon in maize. *Plant Mol Biol* **12**: 633–642.

Kidwell MG (1992) Horizontal transfer. *Curr Opinion Gen Dev* **2**: 868–873.

Kimmel BE, ole-Moiyoi OK, Young JR (1987) *Ingi*, a 5.2-kb dispersed sequence element from *Trypanosoma brucei* that carries half of a smaller mobile element at either end and has homology to mammalian LINEs. *Mol Cell Biol* **7**: 1465–1475.

Konieczny A, Voytas DF, Cummings MP, Ausubel FM (1991) A superfamily of *Arabidopsis thaliana* retrotransposons. *Genetics* **127**: 801–809.

Korenberg JR, Rykowski MC (1988) Human genome organization: Alu, Lines, and the molecular structure of metaphase chromosome bands. *Cell* **53**: 391–400.

Kumar S, Tamura K, Nei M (1993) *MEGA: Molecular Evolutionary Genetics Analysis*, version 1.0. University Park, PA: Pennsylvania State University.

Leeton PRJ, Smyth DR (1993) An abundant LINE-like element amplified in the genome of *Lilium speciosum*. *Mol Gen Genet* **237**: 97–104.

Luan DD, Korman MH, Jakubczak JL, Eickbush TH (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* **72**: 595–605

Manninen I, Schulman AH (1993) *BARE-1*, a *copia*-like retroelement in barley (*Hordeum vulgare* L.) *Plant Mol Biol* **22**: 829–846.

Meyer C, Pouteau S, Rouze P, Caboche M (1994) Isolation and molecular characterization of dTnp1, a mobile and defective transposable element of *Nicotiana plumbaginifolia*. *Mol Gen Genet*: 194–200.

Meyerowitz EM (1994) Plant developmental biology: green genes for the 21st century. *BioEssays* **16**: 621–625.

Moore G, Cheung W, Schwarzacher T, Flavell R (1991) BIS 1, a major component of the cereal genome and a tool for studying genomic organization. *Genomics* **10**: 469–476.

Napoli C, Lemieux C, Jorgensen R (1990) Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous gene in trans. *Plant Cell* **2**: 279–289.

Pouteau S, Huttner E, Grandbastien MA, Caboche M (1991a) Specific expression of the tobacco Tnt1 retrotransposon in protoplasts. *EMBO J* **10**: 1911–1918.

Pouteau S, Spielman A, Meyer C, Grandbastien MA, Caboche M (1991b) Effects of Tnt1 tobacco retrotransposon insertion on target gene transcription. *Mol Gen Genet* **228**: 233–239.

Pryciak PM, Varmus HE (1992) Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration target site selection. *Cell* **69**: 769–780.

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Schmidt T, Heslop-Harrison JS (1993) Variability and evolution of highly repeated DNA sequences in the genus *Beta*. *Genome* **36**: 1074–1079.

Schmidt T, Jung C, Metzlaff M (1991) Distribution and evolution of two satellite DNAs in the genus *Beta*. *Theor Appl Genet* **82**: 793–799.

Schmidt T, Boblenz K, Metzlaff M, Kaemmer D, Weising K, Kahl G (1993) DNA fingerprinting in sugar beet (*Beta vulgaris*) – identification of double-haploid breeding lines. *Theor Appl Genet* **85**: 653–657.

Schmidt T, Schwarzacher T, Heslop-Harrison JS (1994) Physical mapping of rRNA genes by fluorescent *in situ* hybridization and structural analysis of 5S rRNA genes and intergenic spacer sequences in sugar beet (*Beta vulgaris*). *Theor Appl Genet* **88**: 629–636.

Schwarz-Sommer Z, Leclercq L, Göbel E, Saedler H (1987) Cin4, an insert altering the structure of the *A1* gene in *Zea mays*, exhibits properties of non-viral retrotransposons. *EMBO J* **6**: 3873–3880.

Singer M (1982) Highly repeated sequences in mammalian genomes. *Int Rev Cytol* **76**: 67–112.

Taruscio D, Manuelidis L (1991) Integration site preferences of endogenous retroviruses. *Chromosoma* **101**: 141–156.

Voytas DF, Ausubel FM (1988) A *copia*-like transposable element family in *Arabidopsis thaliana*. *Nature* **336**: 242–244.

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

Wichmann HA, Van Den Bussche RA, Hamilton MJ, Baker RJ (1992) Transposable elements and the evolution of genome organization in mammals. *Genetica* **86**: 287–293.

Xiong Y, Eickbush TH (1988) The site-specific ribosomal DNA insertion element R1Bm belongs to a class of non-long-terminal-repeat retrotransposons. *Mol Cell Biol* **8**: 114–123.

Xiong Y, Eickbush TH (1990) Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J **9**: 3353–3362.