Plant Mol Biol (2012) 78:393–405
DOI 10.1007/s11103-011-9872-z

Survey of sugar beet (*Beta vulgaris* L.) *hAT* transposons and MITE-like *hATpin* derivatives

Gerhard Menzel · Carmen Krebs · Mercedes Diez · Daniela Holtgräwe · Bernd Weisshaar · André E. Minoche · Juliane C. Dohm · Heinz Himmelbauer · Thomas Schmidt

Received: 31 August 2011 / Accepted: 20 December 2011 / Published online: 13 January 2012

Abstract Genome-wide analyses of repetitive DNA suggest a significant impact particularly of transposable elements on genome size and evolution of virtually all eukaryotic organisms. In this study, we analyzed the abundance and diversity of the *hAT* transposon superfamily of the sugar beet (*B. vulgaris*) genome, using molecular, bioinformatic and cytogenetic approaches. We identified 81 transposase-coding sequences, three of which are part of structurally intact but nonfunctional *hAT* transposons (BvhAT), in a *B. vulgaris* BAC library as well as in whole genome sequencing-derived data sets. Additionally, 116 complete and 497 truncated non-autonomous BvhAT derivatives lacking the transposase gene were in silico-detected. The 116 complete derivatives were subdivided into four BvhATpin groups each characterized by a distinct terminal inverted repeat motif. Both BvhAT and BvhATpin transposons are specific for species of the genus *Beta* and closely related species, showing a localization on *B. vulgaris* chromosomes predominantely in euchromatic regions. The lack of any BvhAT transposase function together with the high degree of degeneration observed for the BvhAT and the BvhATpin genomic fraction contrasts with the abundance and activity of autonomous and non-autonomous *hAT* transposons revealed in other plant species. This indicates a possible genus-specific structural and functional repression of the *hAT* transposon superfamily during *Beta* diversification and evolution.

Keywords *Beta vulgaris* · Transposable element · *hAT* · *hATpin* · FISH

G. Menzel · C. Krebs · M. Diez · T. Schmidt (✉)
Institute of Botany, Dresden University of Technology, 01062 Dresden, Germany
e-mail: thomas.schmidt@tu-dresden.de

D. Holtgräwe · B. Weisshaar
Center for Biotechnology, University of Bielefeld, 33594 Bielefeld, Germany

A. E. Minoche · J. C. Dohm · H. Himmelbauer
Centre for Genomic Regulation (CRG) and UPF, 08003 Barcelona, Spain

A. E. Minoche · J. C. Dohm · H. Himmelbauer
Max Planck Institute for Molecular Genetics, Ihnestr. 63-73, 14195 Berlin, Germany

Introduction

Transposable elements (TEs) form the most abundant class of dispersed repeats in plant genomes, which can make up to 80% of the genomic fraction in some grass species (Vicient et al. 1999). By their mobility and accumulation, TEs significantly influence gene expression as well as the size variability of plant genomes, thereby increasing genetic diversity (Bennetzen 2005). Thus, TE activity can be regarded as a major driving force for gene and genome evolution (Feschotte and Pritham 2007).

According to their mechanism of transposition TEs are divided into two classes (Finnegan 1989). Retroelements (class I) are copied via an RNA-intermediate, and the 'cut and paste' excision and genomic reintegration of class II DNA-transposons is mediated by an element-encoded transposase. Common features of 'cut and paste' transposons are a transposase gene and terminal inverted repeats (TIR) flanked by a short target site duplication (TSD) generated upon transposition. Based on sequence similarities and structural conservations of TIRs, TSDs and catalytic transposase domains, the six DNA transposon superfamilies *hAT*, Tc1/*mariner,* CACTA, *PIF*/*Harbinger,*

Mutator and *P* are distinguished in plants (Wicker et al. 2007).

Non-autonomous miniature inverted-repeat transposable elements (MITEs) have been shown to represent derivatives of most class II transposon superfamilies in diverse plant species (Benjak et al. 2009). Similar to DNA transposons, MITEs possess TIRs and TSDs, but do not encode a transposase due to a partial or complete deletion of the corresponding open reading frame (ORF). Thus, MITE mobilization depends on the recognition of the TIRs by the transposase of related autonomous elements (Yang et al. 2006). In contrast to the low copy number of these autonomous founder elements, MITEs are often highly amplified within plant genomes (Feschotte et al. 2002).

The most prominent system of autonomous and non-autonomous plant transposons are the *hAT* transposon *Activator* (*Ac*) and its derived *Dissociation* (*Ds*) element, identified and intensively analyzed in maize (McClintock 1947; Kunze and Weil 2002). Specific features of *hAT* transposons are a TSD of eight base pairs and an ORF coding for a transposase harbouring six amino acid domains conserved across plant, animal and fungal species (Rubin et al. 2001).

Plant *hAT* transposons are predominately assigned to a family designated as *Ac/Tam3*, according to the corresponding *hAT* transposons isolated from *Zea mays* and *Antirrhinum majus* (Fedoroff et al. 1983; Hehl et al. 1991). The accelerating analysis of whole plant genomes has significantly increased the information on the abundance, variability and evolutionary history of plant *hAT* families (Holligan et al. 2006; Benjak et al. 2008; Du et al. 2010; Cavallini et al. 2010).

Within the genus *Beta*, the section *Beta* comprises all *B. vulgaris* cultivars such as sugar beet, fodder beet, garden beet and leaf beet. Sugar beet (*Beta vulgaris*) has a relatively small genome of 758 Mbp (Arumuganathan and Earle 1991). Numerous *B. vulgaris* repetitive sequence families such as satellite DNA, retrotransposons and transposons have been extensively studied (Jacobs et al. 2004; Menzel et al. 2006; Menzel et al. 2008; Heitkam and Schmidt 2009). With the genome sequencing of *B. vulgaris* being in progress, a comprehensive identification and annotation of sugar beet repeats is necessary, as its genome was estimated to consist of at least 63% repetitive sequences (Flavell et al. 1974; Menzel et al. 2008).

In this paper, we report on the identification and structural classification of autonomous *B. vulgaris hAT* transposons designated BvhAT as well as non-autonomous MITE-like BvhATpin elements. The abundance and genomic organization of BvhAT and BvhATpin families was revealed in species of the section *Beta* and species of related genera of the Amaranthaceae family. Moreover, the physical organization of both autonomous and non-autonomous *hAT* families on *B. vulgaris* chromosomes was visualized by fluorescent in situ hybridization (FISH).

Materials and methods

Plant material and DNA isolation

Plants of *Beta vulgaris* ssp. *vulgaris* 'KWS 2320' (sugar beet), 'Brigadier' (fodder beet), 'Rote Kugel' (garden beet) and 'Lukullus' (leaf beet), *Beta vulgaris* ssp. *maritima*, *Beta vulgaris* ssp. *adanensis* as well as the species *Beta patula*, *Beta macrocarpa*, *Beta corolliflora*, *Beta macrorhiza*, *Beta nana*, *Patellifolia procumbens*, *Patellifolia patellaris* and *Spinacia oleracea*, and plants of the species *Chenopodium quinoa*, *Arabidopsis thaliana*, and *Zea mays* were grown under greenhouse conditions. Genomic DNA was isolated from seedlings or young leaves using the CTAB (cetyltrimethyl/ammonium bromide) standard protocol (Saghai-Maroof et al. 1984). DNA of *Pinus elliotii* ssp. *elliotii* was provided by RL Doudrick (USDA Southern Research Station, Ashville, USA).

PCR protocols

For the generation of hybridization probes, PCR reactions with 50 ng template DNA were performed in 50 µl containing final concentrations of 0.2 mM dNTPs, 50 mM KCl, 1.5 mM $MgCl_2$, 10 mM Tris–HCl (pH 9.0), and 1 unit of Taq DNA polymerase (Promega, Madison, WI, USA). Standard PCR conditions were 94°C for 3 min, followed by 35 cycles of 94°C for 1 min, primer-specific annealing temperature for 30 s, 72°C for 45 s to 90 s and a final incubation at 72°C for 5 min.

The amplification of *hAT* transposase fragments from *B. vulgaris* genomic DNA was performed according to De Keukeleire et al. (2004) by the degenerated primer pair 5′-CA(C/T)GTI(A/C)GITG(C/T)IIITG(C/T)CA(C/T)AT(A/C/T)(C/T)T-3′ and 5′-AAIGCI(C/G)I(C/T)TCI(C/G)(A/T)IGC(A/C/G/T)-AC(A/C/G/T)GT-3′ (I: Inositol) with an annealing temperature of 40°C.

After gel electrophoresis, PCR fragments were purified with the QIAquick gel extraction kit (Qiagen, Hilden, Germany) and ligated into the vector pGEM-T (Promega, Madison, WI, USA).

Filter hybridizations

For the isolation of full length *hAT* transposons, a high density filter with 9216 clones from a sugar beet BAC library comprising of 50.304 clones (Gindullis et al. 2001) was screened with ^{32}P-labelled probes. Overnight hybridizations were performed at 55°C in 5 x SSPE (20× SSPE contains 3 M sodium chloride, 200 mM NaH_2PO_4, and 20 mM

EDTA [pH 7.4]) with 5× Denhardt solution (100× Denhardt contains 2% polyvinylpyrrolidone, 2% bovine serum albumine, and 2% Ficoll 400) and 0.2% SDS. Posthybridization washings were twice at 60°C in 2× SSC/0.1% SDS for 10 min. Identified BACs were purified with the NucleoBond Xtra Maxi kit (Macherey–Nagel, Düren, Germany).

For Southern hybridization, 2 µg genomic DNA was restricted with different enzymes, separated on an 1.1% agarose gel and transferred onto Hybond-XL nylon membranes (GE Healthcare, Chalfont, UK) using alkaline transfer. Southern hybridization was performed with ^{32}P-labelled probes using standard protocols (Sambrook et al. 1989). Filters were washed as described above. Signals were detected by autoradiography.

Sequence analyses

Selected BACs were sequenced by a bi-directional outward primer walking strategy, starting with primers derived from a BAC-specific transposase fragment. BAC and plasmid DNAs were sequenced with a CEQ 8000 capillary sequencer (Beckman Coulter, Fullerton CA, USA) according to the manufacturer's instructions. Sequences were aligned by the MegAlign option of the Lasergene 8.0 software (DNAStar, Madison, Wisconsin, USA) using CLUSTAL with default parameters.

For the in silico identification of *hAT* transposons and derived *hATpin* elements, we used sequence data of BAC and fosmid clones from the double haploid sugar beet cultivar 'KWS 2320' (Lange et al. 2008; Dohm et al., submitted), and from sugar beet hybrid 'US H20' BAC clones (McGrath et al. 2004). Collectively, 20.4 Mb of 25.874 'US H20' BAC end sequences and 70 Mb from 76.203 KWS2320 fosmid end sequences were used. This data, including redundancies, comprised approximately 0.11× coverage of the sugar beet genome. The search was complemented with 631 Mb of a preliminary sugar beet genome assembly (version RefBeet 0.1.1 from October 2009), a partial coverage of the 'KWS 2320' genotype. The non-public assembly RefBeet 0.1.1. has been calculated from 9.3 Gigabases of Roche/454 single read data; this draft is remained unedited, and no integration with the physical and genetic maps of sugar beet has been performed yet. RefBeet 0.1.1 comprises approximately 340,000 contigs with an N50 contig size of approximately 4,000 bp. For access to the current version of the genome sequence draft, which is in progress and can be made available on a collaborative basis, please contact Heinz Himmelbauer (Heinz.Himmelbauer@crg.es). The sequence data were searched with *B. vulgaris*-specific transposon-derived queries using FASTA (ftp://ftp.ebi.ac.uk/pub/software/unix/fasta/fasta36/). Matching sequences with an e-value threshold of 1×10^{-10} were further analyzed with the softwares BioEdit 7.0 (Hall 1999) and Geneious 4.7 (Drummond et al. 2009).

Open reading frames of autonomous *hAT* transposons were deduced using the GeneWise algorithm (Birney et al. 2004).

Neighbor-joining consensus trees revealing the divergence of *hAT* transposase sequences were generated by the algorithm of the MEGA4.0 software (Tamura et al. 2007).

Sequence logos of *hAT* transposon-specific sequence motifs were generated by WebLogo v2.8.2 (Crooks et al. 2004) via http://weblogo.berkeley.edu/.

Accession numbers

Representative sequence data from this study have been deposited in the EMBL database under the accession numbers FR871854–FR871858.

Fluorescent in situ hybridization

Young leaves of *Beta vulgaris ssp. vulgaris* plants were used for the preparation of mitotic chromosomes. The material was synchronized for 2 h in 2 mM 8-hydroxyquinoline and fixed in methanol:glacial acetic acid (3:1). The leaves were macerated in an enzyme mix and nuclei were dropped on slides as described by Schwarzacher and Heslop-Herrison (2000) with modifications (Desel 2002). The probes were labelled with biotin-16-dUTP and digoxigenin-11-dUTP using a standard PCR. Hybridization and detection was performed according to Schwarzacher and Heslop-Herrison (1991) modified for sugar beet by Schmidt et al. (1994). The chromosome preparations were counterstained with DAPI (4′, 6′-diamidino-2-phenylindole) and mounted in antifade solution (Citifluor).

Observation of the slides was carried out with a Zeiss Axioplan2 fluorescence microscope using the filter sets 15 (Cy3), 09 (FITC) and 02 (DAPI). Images were acquired with the Applied Spectral Imaging (Migdal Ha'Emek, Israel) v. 3.3 software coupled with a high resolution CCD camera (ASI BV300-20A).

Results

Identification of autonomous *hAT* transposons in the *B. vulgaris* genome

Emerging resources of genomic sequencing data from sugar beet genomic clones and a sugar beet whole-genome draft assembly enabled genome-wide in silico analyses of the abundance and diversity of *hAT* transposon families. In order to identify structurally complete autonomous *hAT* transposons, *B. vulgaris*-specific *hAT* query sequences were isolated. Therefore, conserved fragments coding for the *hAT*-specific transposase domains B, C, D and E (Rubin et al. 2001) were amplified by PCR. Sequencing of a

randomly chosen 971 bp fragment revealed significant similarities to plant *hAT* transposases, in particular to the *hAT* transposon *Tag2* of *Arabidopsis thaliana* (Henk et al. 1999). Subsequently, the 971 bp transposase fragment was used as probe for the hybridization of a high density filter containing 9216 BAC clones, representing a 1.5× genome coverage. Fourteen hybridization signals of variable intensity were detected. Based on their strong hybridization signals, three *hATs* were sequenced by primer walking along BAC sequences. Software-assisted analyses identified terminal inverted repeats (TIRs) of 18 bp flanked by a *hAT*-specific eight bp target site duplication (TSD) on BAC 78E10, thus revealing a full length *hAT* transposon of 2,649 bp designated as BvhAT1. On two further BAC clones, the 5′-truncated *hAT* transposons BvhAT2-1 and BvhAT2-2 were identified (Fig. 1a).

The open reading frame (ORF) of the BvhAT1 transposase, which was detected by a GeneWise comparison with the *hAT* transposase *Tag2* from *A. thaliana* (Henk et al. 1999), is 1,785 bp long and splitted by an intron of 109 bp (Fig. 1a). The BvhAT1 ORF is defective due to three frameshifts and two internal stop codons.

The structural assignment of the BvhAT1 transposase to the *hAT* transposon superfamily is demonstrated by a sequence alignment (Fig. 1b) of the corresponding 595 amino acids to the transposases *Tag2* from *A. thaliana* (Henk et al. 1999) and *Ac* from *Zea mays* (Kunze et al. 1987), revealing the presence of six conserved *hAT*-specific amino acid blocks A-F (Rubin et al. 2001).

For a subsequent in silico identification of autonomous *hAT* transposons, the 720.4 MB of *B. vulgaris* genomic sequences were searched using 1,362 bp of the BvhAT1 transposase ORF spanning the domains A-F as a FASTA query. From approximately 1,000 search hits with an e-value threshold of 1×10^{-10}exclusively retrieved from the *B. vulgaris* 630 Mbp genome assembly, a set of 80 sequences were further analyzed by GeneWise using the BvhAT1 tranposase protein as reference sequence. In addition to the 2,649 bp BvhAT1 sequence, two out of these 80 sequences represent complete *hAT* transposons of 4,324 and 4,207 bp, designated as BvhAT3 and BvhAT4, respectively. The significant differences in length of BvhAT3 and BvhAT4 to BvhAT1 are due to non-homologous extensions in the 5′-untranslated regions of 1,704 and 1,457 bp, respectively, which presumably represent different integrations of genomic fragments of unknown origin. Both BvhAT3 and BvhAT4 harbour a complete *hAT* transposase ORF, which is defective because of the presence of numerous stop codons. Moreover, BvhAT3 is flanked by TSDs of eight bp, while no TSD could be detected for BvhAT4. The remaining 78 *hAT* sequences exclusively represent individual truncations of defective *hAT* transposase ORFs, In particular, 54 out of these 78

Fig. 1 **a** Schematic representation of the complete *hAT* transposons BvhAT1 and the truncated transposons BvhAT2-1 and BvhAT2-2, respectively. The transposase ORFs (*grey shaded*) include ORF length and the relative position of the 109 bp intron (*vertical black bar*). TIRs and TSDs are represented by *grey arrows* and *black triangles*, respectively. **b** Alignment of the conserved *hAT* transposase domains A–F (Rubin et al. 2001) deduced from BvhAT1, *Tag2* from *A. thaliana* (AAD24567) and *Ac* from *Zea mays* (P08770). Amino acid similarities ≥66% are indicated by *grey shading*. Gaps within the alignment are shown by *dashes*, and the number of amino acids of each transposase is given at the ends

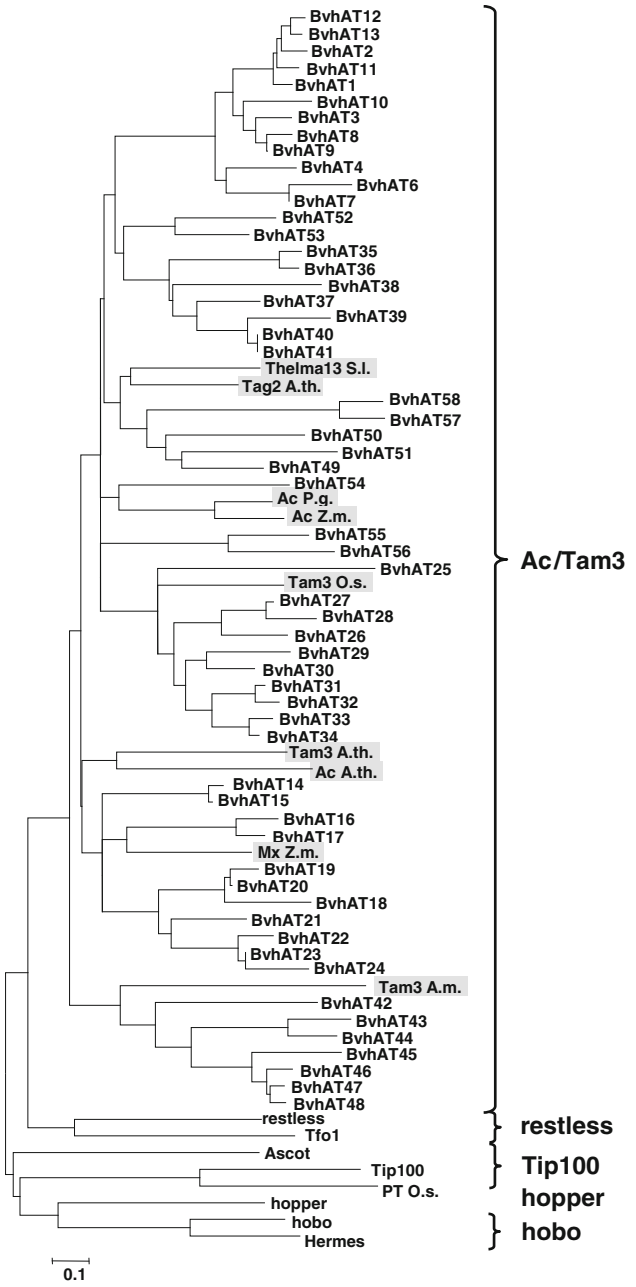

Fig. 2 Diversity of *hAT* transposase fragments harbouring the conserved amino acid domains A–F (Rubin et al. 2001). The Neighbour Joining (NJ) tree shows the relationship of 57 sugar beet BvhAT transposases to *hAT* elements (*grey shading*) of the Ac/Tam3 clade from plants such as *Arabidopsis thaliana* (*A.t.*; Ac AAC61291, Tam3 CAB41922, Tag2 AAD24567), *Antirrhinum majus* (*A.m.*; Tam3 BAA28817), *Silene latifolia* (*S.l.*; Thelma13 AAP59878), *Pennisetum glaucum* (*P.s.*; Ac 2021344A), *Oryza sativa* (*O.s.*; Tam3 AAG13541, putative transposase (PT) AAL86479), and *Zea mays* (*Z.m.*; Ac P08770, Mx AAV82322). The remaining *hAT* transposon clades restless, Tip100, hobo and hopper are represented by animal and fungi transposase sequences from restless (*Tolypocladium inflatum*; CAA93759), Tfo1 (*Fusarium oxysporum*; BAA32244), Ascot (*Ascobolus immersus*; CAA68959); hobo (*Drosophila melanogaster*; A39622), Hermes (*Musca domestica*; AAB60236), hopper (*Bactrocera dorsalis*; AAL93203) and from the plant *Ipomoea purpurea* (Tip100; BAA36225). The genetic distance of the sequences is given by the *scale bar*

sequences were in a size range between 1,098 and 1,416 bp and consist of an continuous ORF for the complete set of transposase domains A–F, as identified by sequence alignments (results not shown). On the remaining 24 sequences, continuous ORFs for a subset of the six transposase-specific A–F domains, ranging from 555 to 1,374 bp, could be identified (results not shown).

The structural diversity of *B. vulgaris hAT* transposases and their relationship to *hAT* transposons from plants, animals and fungi, was analyzed by the neighbour joining algorithm of the MEGA4.0 software. All BvhAT transposase sequences spanning the conserved domains A–F were compared to the corresponding transposase regions from *hAT* transposons of the five *hAT* families forming the *hAT* transposon superfamily (Xu and Dooner 2005). Within the resulting neighbour joining tree (Fig. 2), the subdivision of the transposase sequences into the five families of Ac/Tam3, restless, Tip100, hopper and hobo (Xu and Dooner 2005) is unambiguous. The BvhAT transposases were exclusively assigned to the Ac/Tam3 family. Apart from a subclade containing 12 members, a clustering of *B. vulgaris hAT* transposases has not been observed. Generally, grouping of the analyzed Ac/Tam3 transposons is only weakly supported by bootstrap values (results not shown).

In silico identification of non-autonomous *hATpin* transposons

Valuable sequence information might be lost by using heterologous queries in genome analyses, as presumed for the analysis of the palm (*Phoenix dactylifera*) genome (Al-Dous et al. 2011). To generate a specific query sequence for the identification of *B. vulgaris*-specific non-autonomous MITE-like *hAT* transposon, a top-down approach recruiting a single primer resembling the outer 12 bp of the 5′-TIR sequence of BvhAT1 was used in PCRs. A PCR product of 927 bp revealed 64 bp of the 5′-region and 75 bp of the 3′-region showing similarities of 74.2% and 82.4%, respectively, to the corresponding regions of BvhAT1. The remaining 788 bp of the PCR fragment showed no significant similarities to either the 5′-non-coding, 3′-non-coding, or the transposase coding regions of BvhAT1.

The 927 bp fragment was subsequently used as a FASTA query within the *B. vulgaris* nucleotide database. Search hits up to an e-value of 1×10^{-10} were further analyzed for the presence of both 5′- and 3′-TIR structures. Thus, 116 individual non-autonomous *hAT* sequences ranging from 338 to 2,260 bp with both 5′- and 3-TIRs, but

no *hAT* transposase-specific motifs were identified. Moreover, 84 out of these 116 complete non-autonomous transposons are flanked by *hAT*-specific TSDs of eight bp.

The nucleotide alignment of the three autonomous transposons BvhAT1, BvhAT3 and BvhAT4 and 116 non-autonomous elements revealed the presence of the conserved pentamers 5′-CGAGC-3′ and 5′-GCTCG-3′, ranging from 6 to 14 copies in the 5′-subterminal regions (STR), and from 4 to 20 copies in the 3′-STRs, respectively (see below, Fig. 3b). This shared CGAGC/GCTCG motifs structurally resemble the STR of a family of MITE-like *hAT* transposons identified in plant species like lettuce and rice. By the ability to form stable secondary structures mediated by subterminal CGAGC/GCTCG repetitions, these repetitive elements were designated as *hATpin* transposons (Moreno-Vázquez et al. 2005). Similarly, the 116 non-autonomous *B. vulgaris* sequences were regarded as *hATpin* MITEs named BvhATpin.

The diversity of BvhATpin transposons was displayed by a neighbour joining (NJ) analysis including the autonomous BvhAT1, BvhAT3 and BvhAT4 sequences. Based on a sequence aligment of the common 195 bp of the 5′-ends, including TIR and subterminal CGAGC/GCTCG motifs, the resulting cladogram indicates the subdivision of the sugar beet BvhAT/BvhATpin transposons into four families (Fig. 3a). The autonomous *hAT* transposons BvhAT1, BvhAT3 and BvhAT4 were grouped into the BvhATpin families II, III and I, respectively (Fig. 3a). Moreover, the cladogram suggests a derivation of BvhATpin III MITEs from the autonomous BvhAT3. Similarly, BvhATpin II elements were shown to be possibly derived from BvhAT1 (Fig. 3a). An exemplary alignment of BvhAT1 and its derivative BvhATpin II-37 reveals a high degree of nucleotide identity and an amplification of CGAGC/GCTCG motifs within 5′- and 3′-subterminal regions of BvhATpin II-37 (Fig. 3b). In addition to a specific TIR length and nucleotide composition, each BvhATpin family is characterized by a family-specific average size (Table 1). These distinct size classes of BvhATpin elements are detectable by PCR with a TIR primer, revealing three amplicons with sizes of approximately 600, 900 and 1,100 bp, respectively (results not shown).

For a detailed analysis of TIR structures, the length and the relative frequency of the 5′-TIR nucleic acid residues from 84 BvhAT and BvhATpin transposons showing at most one mismatch between 5′- and 3-TIRs were depicted by a family-specific sequence logo (Fig. 4a). An overall comparison of the of group I–IV TIR logos reveals the 5′-terminal consensus sequence TAGGG, which meets the *hAT*-specific TIR consensus (T/C)A(A/G)NG detected by Rubin et al. (2001). The 8 bp TIR of BvhATpin family IV resembles the TIR structure described for rice *hATpin* MITEs (Moreno-Vázquez et al. 2005), while the TIRs of

Fig. 3 Diversity of autonomous BvhAT transposons and non-autonomous BvhATpin MITEs. **a** Cladogram, derived from the nucleotide alignment using 195 bp of the 5′-terminal regions including TIR motifs of three BvhAT transposons (*asterisks*) and 116 BvhATpin MITEs, resulting in the four clades BvhATpin I–IV. **b** Nucleotide alignment of the 5′- and 3′-subterminal regions, respectively, from BvhAT1 and its putative derivative BvhATpin II-37. Positions of the pentamer repeats GCTCG (*block arrow right*) and the reverse complement motif CGAGC (*block arrow left*) within both sequences are given above and below the corresponding sequence. Additional pentamer motifs within the BvhATpin II-37 sequence are marked by *triangles*, and common TIR sequences are framed

the BvhATpin I–III families are significantly extended to 15, 18 and 12 bp, respectively. The presence of almost perfect TIRs of the 84 BvhAT/BvhATpin transposons indicates a conservation of the corresponding TSDs, therefore TSD structures were also analyzed by the generation of a sequence logo. As shown in Fig. 4b, the TSD consensus reveals a slight preference of target sites nucleotide composition rather than a random nucleotide sequence of BvhAT and BvhATpin integration sites.

In order to estimate the approximate number of BvhAT/BvhATpin transposons in the beet genome, the presence and abundance of truncated BvhATpin elements was also determined, using each of the 116 BvhATpin transposons as a query to search within an early draft *B. vulgaris* genome sequence. In this dataset, 304 truncated BvhATpin copies ranging in size between 21 and 942 bp exclusively terminated by a 5′-TIR motif were identified, while 204 BvhATpin fragments with sizes from 17 to 950 bp are only terminated by a 3′-TIR motif. These 513 BvhATpin MITE truncations were unambiguously assigned to each of the four BvhATpin families (Table 1). Therefore, by adding the 0.23 Mbp of complete and truncated BvhATpin sequences to the 0.1 Mbp of autonomous complete and partial BvhAT transposons, an amount of at least 0.33 Mbp equaling 0.04% of autonomous as well as non-autonomous *hAT* transposons were detected within the *B. vulgaris* genome, assuming a genome size of 758 Mbp.

Genomic organization of BvhAT and BvhATpin families in *B. vulgaris*

The genomic organization of the BvhAT family in *B. vulgaris* was analyzed by Southern hybridization (Fig. 5), using the 971 bp transposase probe amplified from BvhAT1 as described above to enable the recognition of the majority of sugar beet *hAT* transposases.

Hybridization to genomic *B. vulgaris* DNA digested with the enzymes *Alu*I, *Hae*III and *Rsa*I (Fig. 5a, lanes 1–3) showed conserved fragments superimposed on weak background hybridization over a wide molecular weight range. Probing of the BvhAT1 transposase to genomic DNA digested with the methylation-sensitive enzymes

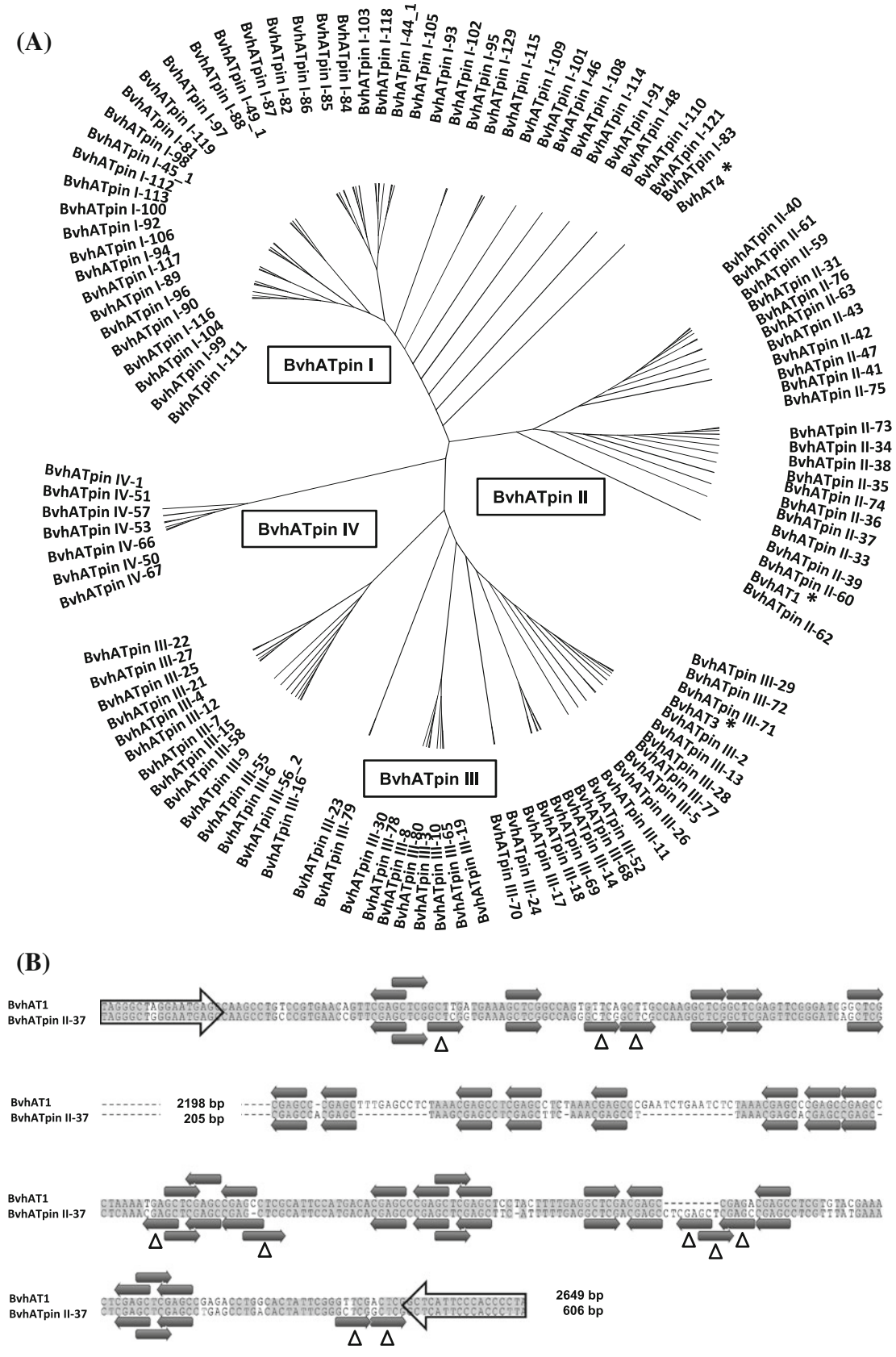
(A)
BvhATpin I
BvhATpin II
BvhATpin III
BvhATpin IV
BvhATpin I-103
BvhATpin I-118
BvhATpin I-44_1
BvhATpin I-105
BvhATpin I-93
BvhATpin I-102
BvhATpin I-95
BvhATpin I-129
BvhATpin I-115
BvhATpin I-109
BvhATpin I-101
BvhATpin I-46
BvhATpin I-108
BvhATpin I-114
BvhATpin I-91
BvhATpin I-48
BvhATpin I-110
BvhATpin I-121
BvhATpin I-83
BvhAT4 *
BvhATpin I-85
BvhATpin I-84
BvhATpin I-86
BvhATpin I-82
BvhATpin I-87
BvhATpin I-49_1
BvhATpin I-88
BvhATpin I-97
BvhATpin I-119
BvhATpin I-81
BvhATpin I-98
BvhATpin I-45_1
BvhATpin I-112
BvhATpin I-113
BvhATpin I-100
BvhATpin I-92
BvhATpin I-106
BvhATpin I-94
BvhATpin I-117
BvhATpin I-89
BvhATpin I-96
BvhATpin I-90
BvhATpin I-116
BvhATpin I-104
BvhATpin I-99
BvhATpin I-111
BvhATpin II-40
BvhATpin II-61
BvhATpin II-59
BvhATpin II-31
BvhATpin II-76
BvhATpin II-63
BvhATpin II-43
BvhATpin II-42
BvhATpin II-47
BvhATpin II-41
BvhATpin II-75
BvhATpin II-73
BvhATpin II-34
BvhATpin II-38
BvhATpin II-35
BvhATpin II-74
BvhATpin II-36
BvhATpin II-37
BvhATpin II-33
BvhATpin II-39
BvhATpin II-60
BvhAT1 *
BvhATpin II-62
BvhATpin IV-1
BvhATpin IV-51
BvhATpin IV-57
BvhATpin IV-53
BvhATpin IV-66
BvhATpin IV-50
BvhATpin IV-67
BvhATpin III-22
BvhATpin III-27
BvhATpin III-25
BvhATpin III-21
BvhATpin III-4
BvhATpin III-12
BvhATpin III-7
BvhATpin III-15
BvhATpin III-58
BvhATpin III-9
BvhATpin III-55
BvhATpin III-6
BvhATpin III-56_2
BvhATpin III-16
BvhATpin III-23
BvhATpin III-79
BvhATpin III-30
BvhATpin III-78
BvhATpin III-8
BvhATpin III-80
BvhATpin III-10
BvhATpin III-65
BvhATpin III-19
BvhATpin III-70
BvhATpin III-24
BvhATpin III-17
BvhATpin III-18
BvhATpin III-69
BvhATpin III-14
BvhATpin III-68
BvhATpin III-52
BvhATpin III-11
BvhATpin III-26
BvhATpin III-5
BvhATpin III-77
BvhATpin III-28
BvhATpin III-13
BvhATpin III-2
BvhAT3 *
BvhATpin III-71
BvhATpin III-72
BvhATpin III-29
(B)
BvhAT1
BvhATpin II-37
2198 bp
205 bp
2649 bp
606 bp

Table 1 Characterization of four *Beta vulgaris hATpin* families

Family	Number of intact *hATpin* elements	Sizes of intact *hATpin* elements (bp)		Number of truncated *hATpin* elements	
		Size range	Average size	3′-truncations	5′-truncations
BvhATpin I	45	471–881	576	88	52
BvhATpin II	22	338–1,634	792	23	–
BvhATpin III	42	733–2,260	1,125	101	90
BvhATpin IV	7	475–1,210	683	97	62

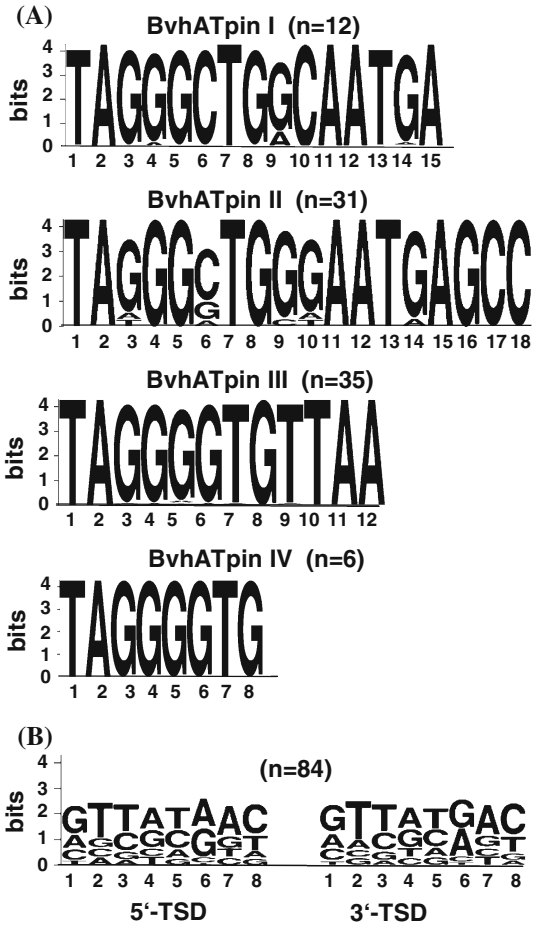

Fig. 4 **a** Sequence logos, generated with (n) sequences, giving the relative frequency of the nucleic acid residue within the 5′-TIRs of the four *B. vulgaris* hATpin clades and **b** within the 8 bp TSD motifs flanking the TIRs of all 84 members of BvhATpin clades *I–IV* analyzed in (**a**). The overall height of each nucleotide stack indicates the sequence conservation at the given nucleotide positions measured in bits

*Hpa*II and *Msp*I (Fig. 5a, lanes 7 and 8) showed strong hybridization in a high molecular weight range, indicating the presence of most BvhAT members in chromosomal regions with a high degree of CG or CNG methylation on CCGG restriction sites.

To compare the abundance of BvhAT and BvhATpin elements in sugar beet, genomic DNA was hybridized with the 606 bp probe of BvhATpin II-37, which is most likely derived from BvhAT1 (Fig. 3b). Conserved fragments are similar to the hybridization pattern of BvhAT1 (Fig. 5b, lanes 1–6) while the number of *Hpa*II and *Msp*I fragments increases. Moreover, the intensity of the hybridization smear (Fig. 5b, lanes 2–8) is slightly stronger than in the BvhAT1 hybridization and indicates a higher copy number of *hATpin* elements.

The distribution and abundance of BvhAT1 was also investigated in the Amaranthaceae family, including species of the genus *Beta*, the genus *Patellifolia*, and the two distantly related species spinach and quinoa. As outgroup species, the gymnosperm *Pinus elliotii*, and *A. thaliana* and *Zea mays* containing *Tag2* and *Ac* as members of the Ac/Tam3 transposon family, respectively, were chosen.

The genus *Beta* is divided into the sections *Beta*, *Corollinae*, and *Nanae*. The section *Beta* includes the species *B. vulgaris* with the subspecies *vulgaris*, representing all cultivars of sugar beet, fodder beet, garden beet and leaf beet, and the two wild subspecies *Beta vulgaris* subsp. *maritima* and *Beta vulgaris* subsp. *adanensis*. Further wild beet species within this section are *Beta macrocarpa* and *Beta patula*. Hybridization of the 971 bp BvhAT1 transposase coding fragment to *Hae*III-digested genomic DNA of these *Beta* species (Fig. 5c, I: lanes 1–8) reveals a species-specific hybridization pattern, with a higher number of conserved fragments in *B. vulgaris* subspecies (Fig. 5c, lanes 1–6), compared to *B. macrocarpa* and *B. patula* (Fig. 5c, lanes 7 and 8). Hybridization to genomic DNA of species from the sections *Corollinae* and *Nanae*, respectively, shows a section-specific pattern of conserved fragments (Fig. 5c, II: lanes 9 and 10; III: lane 11) within the species tested, with an almost similar strength of hybridization and amount of fragments compared to the wild species *B. macrocarpa* and *B. patula* from the section *Beta* (Fig. 5c, lanes 7 and 8). The two species from the genus *Patellifolia* (Fig. 5c, P: lanes 12 and 13) display a low number of weakly conserved fragments. Hybridization to genomic DNA of outgroup species from the Amaranthaceae family as well as *A. thaliana*, *Z. mays* and *P. elliotii* results only in an extremely weak hybridization smear (Fig. 5c, O: lanes 14–18).

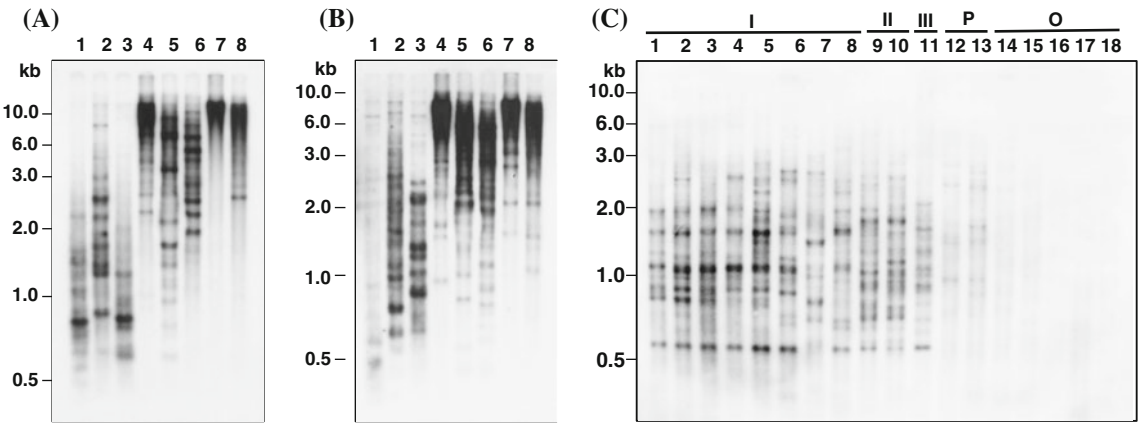

Fig. 5 Comparative Southern hybridization of the autonomous *hAT* transposons BvhAT1 (**a**) and the non-autonomous BvhATpin II-37 (**b**) with genomic DNA of *B. vulgaris* ssp. *vulgaris* "KWS 2320", digested with *Alu*I (*1*), *Hae*III (*2*), *Rsa*I (*3*), *Bam*HI (*4*), *Eco*RI (*5*), *Hin*dIII (*6*), *Hpa*II (*7*) and *Msp*I (*8*). **c** Species distribution of BvhAT1 shown by Southern hybridization of *Hae*III-restricted genomic DNA. The species tested were cultivars of *B. vulgaris* ssp. *vulgaris* of the section *Beta* (*I*): sugar beet "KWS2320" (*1*), fodder beet "Brigadier" (*2*), garden beet "Rote Bete" (*3*) and chard "Lukullus" (*4*), and wild beet species of the section *Beta* (I): *B. vulgaris* ssp. *maritima* (*5*), *B. vulgaris* ssp. *adanensis* (*6*), *B. macrocarpa* (*7*) and *B. patula* (*8*); species of the section *Corollinae* (*II*): *B. corolliflora* (*9*) and *B. macrorhiza* (*10*); species of the section *Nanae* (*III*): *B. nana* (*11*); species of the genus *Patellifolia* (P): *P. procumbens* (*12*), *P. patellaris* (*13*); outgroup species (O): *Spinacia oleracea* (*14*), *Chenopodium quinoa* (*lane 15*), *Arabidopsis thaliana* (*lane 16*), *Zea mays* (*lane 17*) and *Pinus elliotii* (*lane 18*)

Distribution of BvhAT and BvhATpin elements along *B. vulgaris* chromosomes

The chromosomal distribution of autonomous BvhAT and non-autonomous BvhATpin transposons was analyzed on *B. vulgaris* mitotic metaphase chromosomes and interphase nuclei by fluorescent in situ hybridization, using the 971 bp BvhAT1 transposase probe and the 606 bp BvhATpin II-37 sequence, respectively, previously applied for Southern hybridizations (Fig. 5a, b).

The transposase probe revealed a dispersed arrangement of fluorescent signals along all chromosome arms (Fig. 6a), with hybridization sites indicated as doublets of signals. These signals were mostly observed in intercalary and subterminal regions of the euchromatin, while signals were depleted in heterochromatic centromeric regions. This is particularly apparent on interphase chromosomes, where the BvhAT transposases are predominantly localized in euchromatic chromosomal regions, which are less stained with DAPI (Fig. 6b).

After hybridization with the BvhATpin II-37 probe, signals of BvhATpin transposons are predominantly visible in euchromatic localizations on all metaphase chromosomes. In accordance with the results of the Southern

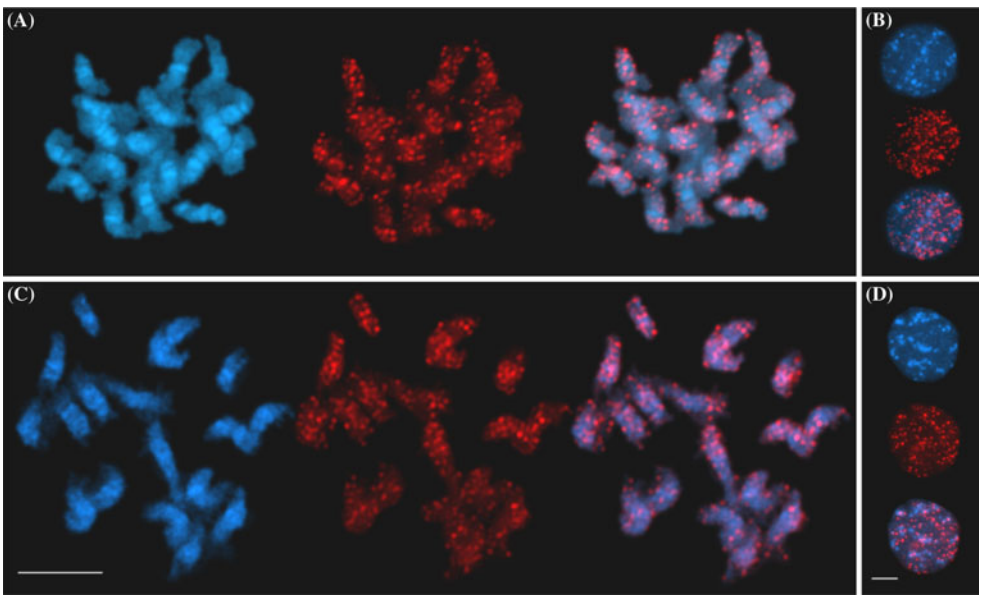

Fig. 6 Fluorescent in situ hybridization of BvhAT1 and BvhATpin II-37 probes to metaphase chromosomes (**a**, **c**) and interphases nuclei (**b**, **d**, respectively) from *B. vulgaris* ssp. *vulgaris* "KWS 2320". In each panel, the DAPI stained DNA (*blue fluorescence*) shows the morphology of the chromosomes. The BvhAT1 as well as the BvhATpin II-37 hybridization signals are visible as red fluorescent signals (middle and overlay). The *scale bars* in **c** and **d** correspond to 5 μm

hybridization, the strength of signals indicates that the number of the corresponding BvhATpin members is increased compared to the BvhAT1 transposase (Fig. 6c, d).

Discussion

Sequencing and subsequent comparison of whole genomes from different plant species significantly accelerates the analysis of the abundance, diversity and evolutionary history of transposon superfamilies, as exampled in *Arabidopsis* and *Brassica* (Zhang and Wessler 2004), *Lotus japonicus* (Holligan et al. 2006), grapevine (Benjak et al. 2008), soybean (Du et al. 2010), and sunflower (Cavallini et al. 2010). Nevertheless, the genomic fractions of *hAT* transposons within these species show significant size differences. In genomes of *Arabidopsis* and *Brassica*, relatively small *hAT* fractions were detected compared to the number of members in other transposon families (Zhang and Wessler 2004). In contrast, 3.64 Mb comprising of 1459 *hAT* elements representing 0.66% of the genome were identified in grapevine, thus forming the largest DNA transposon superfamily within this species (Benjak et al. 2008).

By a combination of molecular analyses and whole-genome in silico approaches, we have assigned at least 0.1 Mbp comprising of 81 either complete or truncated autonomous BvhAT transposons to 0.01% of the *B. vulgaris* genome. Therefore, the *hAT* portion within the *B. vulgaris* genome is most likely smaller than the percentage of *hAT* transposons estimated for other plant genomes. In particular, the number of only three complete *B. vugaris*-specific *hAT* transposons significantly contrasts the amount of 65 intact *hAT* transposons in soybean with its genome size of 975 Mbp (Du et al. 2010), or 118 complete *hAT* elements within 32 Mbp of the 472 Mbp genome of *Lotus japonicus* (Holligan et al. 2006). These contrasting values indicate *hAT* activity, which has been demonstrated by analyses of *hAT* mobility in diverse plant genomes (Altinkut et al. 2006, Fujino et al. 2009). In maize, transcriptional activity of *hAT* transposons as well as the *hAT* mediated excision and spreading of a *Ds* element from a single donor site has been shown (Vicient 2010; Vollbrecht et al. 2010). The transcriptional activity of the intact *hAT Thelma13* from *S. latifolia*, which shows the highest degree of homology to BvhAT transposases, was revealed by RT-PCR (Pritham et al. 2003). In contrast, despite their structural completeness the three *B. vulgaris* hAT transposons are functionally defective due to several ORF frameshifts. As a consequence, sugar beet most likely does not contain any *hAT* transposase activity that might mediate a mobilization of BvhAT and non-autonomous BvhAT derivatives in the genome.

Members of the *hAT* superfamily in plants are predominantly grouped into the *Ac/Tam3* family. Few exceptions like *Tip100* from *Ipomoea purpurea*, or *Tag1* and *Bg* from *Arabidopsis* and maize, respectively, form distinct *hAT* families (Xu and Dooner 2005). All BvhAT transposases analyzed belong to the *Ac/Tam3* clade. Nevertheless, a species-specific grouping of *B. vulgaris* sequences and separation from prominent transposases of other plant species like *Ac* and *Tam3* could not be observed in a cladogram. This high degree of amino acid conservation within *hAT* transposases from *Beta* and different monocotyledonous and dicotyledonous plant species indicates that the *Ac/Tam3* transposon family is ancient and conserved throughout the plant kingdom.

Class II "cut and paste" DNA transposon superfamilies like *Mariner*, PIF/*Harbinger* and *hAT* have been shown to harbour non-autonomous derivatives, commonly designated as miniature inverted-repeat transposable elements (MITEs), in diverse plant species (Feschotte et al. 2003; Zhang et al. 2004; Macas et al. 2005; Moreno-Vázquez et al. 2005; Menzel et al. 2006). Due to the lack of an internal transposase gene, MITE mobilization most likely depends on the transposase provided from related autonomous elements (Feschotte et al. 2002). Within the *hAT* transposon superfamily, several non-cross reacting combinations of autonomous members and non-autonomous derivatives have been described, like Ac/Ds and Mx/rMx from maize (Kunze and Weil 2002; Xu and Dooner 2005) as well as Dart/nDart, Tok/dTok and DaiZ/nDaiZ from rice (Fujino et al. 2005; Moon et al. 2006; Huang et al. 2009). Another MITE-like *hAT* family, designated *hATpin*, was in silico identified in diverse plant species. Common feature of *hATpin* elements, in particular from rice, are a TIRs composed of TAGGG(C/G)TG nucleotide residues, and subterminal regions (STRs) consisting of numerous CGAGC/GCTCG motifs mediating the formation of stable secondary structures (Moreno-Vázquez et al. 2005). Similarly, the presence of at least three repetitions of the GCTCG/CGAGC pentamer in 5′- and 3′-STRs structurally characterize the 119 complete *B. vulgaris*-specific *hAT* elements as *hATpin* transposons.

In rice, 242 *hATpin* transposons, four of which represent autonomous elements, were grouped into three main clades based on variations of STR motifs and exchange of a single nucleotide residue within the TAGGG(C/G)TG TIR (Moreno-Vázquez et al. 2005). In contrast, the diversity analysis of three BvhAT and 116 *hATpin* transposons containing TIRs as well as common subterminal regions of the 5′-end resulted in the formation of four clades, with the TIRs of all clades revealing the typical consensus (T/C)A(A/G)NG of *Ac/Tam3* transposons (Rubin et al. 2001).

Nevertheless, only clade IV harbouring only 7 out of 116 *B. vulgaris hATpin* members show TIRs, which are structurally similar to the *hATpin* elements of rice. Interestingly, the TIRs of the BvhATpin transposons in clades I–III are extendend in length in a clade-specific manner. Based on this structural TIR differences it is tempting to speculate about the putative presence of at least four families of BvhAT transposons in the *B. vulgaris* genome, because the mobilization of non-autonomous transposons requires the recognition of their TIRs by the transposase of a related autonomous element (Feschotte et al. 2002). However, the in silico analyses of *B. vulgaris* whole genome data did not provide a physical link of the numerous truncated nucleotide sequences to the fragments of BvhAT transposase ORFs. Although the close relationship of BvhAT and BvhATpin transposons is revealed in the clades I, II and III, and several BvhATpin elements most likely originate from BvhAT1 and BvhAT3, no indications for a clade IV-specific BvhAT transposase has been obtained. As the analysis of the *B. vulgaris* genome sequencing data is constantly improving, a future identification of additional complete BvhAT transposons, in particular of clade IV BvhAT elements, might be achieved. Alternatively, the degradation of clade IV transposons in evolutionary time scales might have resulted in a loss of the corresponding nucleotide sequences.

The number of three complete *hAT* transposons as well as 78 transposase fragments, respectively, identified in silico in the currently available sugar beet nucleotide sequences contradicts the calculation of at least 10 hAT transposases based on the hybridization to a high density BAC filter representing 1.5× coverage of the *B. vulgaris* genome. Therefore, more than one *hAT* transposase sequence per BAC can be assumed, supporting the idea of *hAT* transposition into linked sites as observed for Ac/Ds transposons (Kunze and Weil 2002). Similarly, signals in FISH on *B. vulgaris* metaphases, indicating a dispersed distribution of BvhAT and BvhATpin transposons on all chromosomes, might represent physically linked integrations.

FISH also demonstrates the insertion of BvhAT and BvhATpin transposons predominantly in euchromatic regions, which are characterized by a relatively high gene density. This insertions preference for genic regions seems to be common for numerous MITE-like transposon families in different plant species (Moreno-Vázquez et al. 2005; Menzel et al. 2006; Vollbrecht et al. 2010; Takagi et al. 2010). By their integration in promoter regions MITEs might provide novel regulatory sequences possibly resulting in an alteration of gene expression. Moreover, MITEs are able to incorporate and subsequently spread genomic sequences (Menzel et al. 2006). In internal regions of BvhATpin transposons, we generally have observed genomic sequences of different length resulting in distinct clade-specific size classes. These insertions are not related to transposase ORFs and therefore have been inserted from unknown genomic regions. Thus, by the integration near or even into gene coding sequences, BvhATpin transposons might contribute to changes in gene expression and hence genome evolution.

The integration of *hAT* transposons into genomic regions does not occur randomly, but is most likely influenced by structural features of the 8 bp target site. The analysis of the target site of 28 hAT elements, including derivatives, in 12 *Drosophila* species revealed a target site specificity requiring the nucleotide residues T and A at positions 2 and 7, respectively (de Freitas et al. 2010). This strong specificity of nucleotide positions could not be observed in the sequence logo generated from the TSD of 84 BvhAT and BvhATpin transposons. Nevertheless, the weak dominance of nucleotide residues within each position of BvhAT/BvhATpin target sites presumably indicates a target site preference, which is not defined directly by the nucleotide sequence, but by special features encrypted by the sequence composition. Such situation was identified in the analysis of 1,741 insertion sites of maize *Ds* elements, indicating a target site-specific increase in GC-content inducing a hydrogen bonding signature possibly mediating the preferential recognition by the transposase (Vollbrecht et al. 2010). Similarly, target sites of BvhAT and BvhATpin transposons might be characterized by such properties rather than the nucleotide sequence.

Compared to the low copy number of the autonomous founder elements, MITE derivatives of certain DNA transposon superfamilies are highly amplified within plant genomes, as shown for *Mariner*-derived *stowaway* MITEs (Feschotte et al. 2003, Menzel et al. 2006). In contrast, *hAT* derivatives in plant genomes are present in significantly lower amount, as exampled for the nDart, dTok and nDaiz families in the rice cultivar 'Nipponbare' comprising of 13, 25 and 16 members, respectively (Fujino et al. 2005; Moon et al. 2006; Huang et al. 2009). MITE-like *hATpin* elements of the rice genome are present in only 238 copies (Moreno-Vázquez et al. 2005). Similarly, only a weak amplification of BvhATpin elements could be observed in Southern and FISH analyses. Moreover, the identification of BvhATpin fragments and their assignment to the four BvhATpin families has revealed a considerably higher number of truncations than complete elements. It is tempting to speculate that the high degree of degradation and fragmentation observed for the small genomic fraction of BvhAT and BvhATpin sequences might be an indication for the supression of the *hAT* transposon superfamily during the evolution of *Beta* genomes. Nevertheless, future comparative whole genome analyses of other DNA transposon superfamilies will reveal whether this repression is selective for *hAT*

transposons, or a general control mechanism of *Beta* transposon activity.

Acknowledgments We thank Ines Walter for excellent technical assistance. This work has been funded by the "Verbundprojekt GABI BeetSeq: Erstellung einer Referenzsequenz für das Genom der Zuckerrübe (*Beta vulgaris*)", FKZ 0315069A and 0315069B (to H.H. and B.W.).

References

Al-Dous EK, George B, Al-Mahmoud ME, Al-Jaber MY, Wang H, Salameh YM, Al-Azwani EK, Chaluvadi S, Pontaroli AC, Debarry J, Arondel V, Ohlrogge J, Saie IJ, Suliman-Elmeer KM, Bennetzen JL, Kruegger RR, Malek JA (2011) De novo genome sequencing and comparative genomics of date palm (Phoenix dactylifera). Nat Biotechnol 29:521–527

Altinkut A, Kotseruba V, Kirzhner VM, Nevo E, Raskina O, Belyayev A (2006) Ac-like transposons in populations of wild diploid *Triticeae* species: comparative analysis of chromosomal distribution. Chromosom Res 14:307–317

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

Benjak A, Forneck A, Casacuberta JM (2008) Genome-wide analysis of the "cut-and-paste" transposons of grapevine. PLoS One 3:e3107

Benjak A, Boué S, Forneck A, Casacuberta JM (2009) Recent amplification and impact of MITEs on the genome of grapevine (Vitis vinifera L.). Genome Biol Evol 20:75–84

Bennetzen JL (2005) Transposable elements, gene creation and genome rearrangement in flowering plants. Curr Opin Genet Dev 15:621–627

Birney E, Clamp M, Durbin R (2004) GeneWise and Genomewise. Genome Res 14:988–995

Cavallini A, Natali L, Zuccolo A, Giordani T, Jurman I, Ferrillo V, Vitacolonna N, Sarri V, Cattonaro F, Ceccarelli M, Cionini PG, Morgante M (2010) Analysis of transposons and repeat composition of the sunflower (Helianthus annuus L.) genome. Theor Appl Genet 120:491–508

Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. Genome Res 14:1188–1190

de Freitas OrtizM, Lorenzatto KR, Corrêa BR, Loreto EL (2010) *hAT* transposable elements and their derivatives: an analysis in the 12 *Drosophila* genomes. Genetica 138:649–655

De Keukeleire P, De Schepper S, Gielis J, Gerats T (2004) A PCR-based assay to detect hAT-like transposon sequences in plants. Chromosom Res 12:117–123

Desel C (2002) Chromosomale Lokalisierung von repetitiven und unikalen DNA-Sequenzen durch Fluoreszenz- in situ- Hybridisierung in der Genomanalyse bei Beta-Arten. Dissertation Christian-Albrechts-University Kiel, Germany

Drummond AJ, Ashton B, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, Wilson A (2009) Geneious v4.7. Available from http://www.geneious.com/

Du J, Grant D, Tian Z, Nelson RT, Zhu L, Shoemaker RC, Ma J (2010) SoyTEdb: a comprehensive database of transposable elements in the soybean genome. BMC Genomics 11:113

Fedoroff N, Wessler S, Shure M (1983) Isolation of the transposable maize controlling elements *Ac* and *Ds*. Cell 35:235–242

Feschotte C, Pritham EJ (2007) DNA transposons and the evolution of eukaryotic genomes. Annu Rev Genet 41:331–368

Feschotte C, Jiang N, Wessler SR (2002) Plant transposable elements: where genetics meets genomics. Nat Rev Genet 3:329–341

Feschotte C, Swamy L, Wessler SR (2003) Genome-wide analysis of mariner-like transposable elements in rice reveals complex relationships with stowaway miniature inverted repeat transposable elements (MITEs). Genetics 163:747–758

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

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

Fujino K, Sekiguchi H, Kiguchi T (2005) Identification of an active transposon in intact rice plants. Mol Genet Genomics 273:150–157

Fujino K, Matsuda Y, Sekiguchi H (2009) Transcriptional activity of rice autonomous transposable element Dart. J Plant Physiol 166:1537–1543

Gindullis F, Dechyeva D, Schmidt T (2001) Construction and characterization of a BAC library for the molecular dissection of a single wild beet centromere and sugar beet (*Beta vulgaris*) genome analysis. Genome 44:846–855

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Hehl R, Nacken W, Krause A, Saedler H, Sommer H (1991) Structural analysis of Tam3, a transposable element from *Antirrhinum majus*, reveals homologies to the Ac element from maize. Plant Mol Biol 16:369–371

Heitkam T, Schmidt T (2009) BNR—a LINE family from *Beta vulgaris*- contains a RRM domain in open reading frame 1 and defines a L1 sub-clade present in diverse plant genomes. Plant J 59:872–882

Henk AD, Warren RF, Innes RW (1999) A new Ac-like transposon of *Arabidopsis* is associated with a deletion of the RPS5 disease resistance gene. Genetics 151:1581–1589

Holligan D, Zhang X, Jiang N, Pritham EJ, Wessler SR (2006) The transposable element landscape of the model legume Lotus japonicus. Genetics 174:2215–2228

Huang J, Zhang K, Shen Y, Huang Z, Li M, Tang D, Gu M, Cheng Z (2009) Identification of a high frequency transposon induced by tissue culture, nDaiZ, a member of the hAT family in rice. Genomics 93:274–281

Jacobs G, Dechyeva D, Menzel G, Dombrowski C, Schmidt T (2004) Molecular characterization of *Vulmar1*, a complete *mariner* transposon of sugar beet and diversity of *mariner*- and *En/Spm*-like sequences in the genus *Beta*. Genome 47:1–10

Kunze R, Weil CF (2002) The hAT and CACTA superfamily of plant transposons. In: Craig NL (ed) Mobile DNA. ASM Press, Washington, DC, pp 565–610

Kunze R, Stochaj U, Laufs J, Starlinger P (1987) Transcription of transposable element Activator (Ac) of *Zea mays* L. EMBO J 6:1555–1563

Lange C, Holtgräwe D, Schulz B, Weisshaar B, Himmelbauer H (2008) Construction and characterization of a sugar beet (*Beta vulgaris*) fosmid library. Genome 51:948–951

Macas J, Koblízková A, Neumann P (2005) Characterization of Stowaway MITEs in pea (*Pisum sativum* L.) and identification of their potential master elements. Genome 48:831–839

McClintock B (1947) Cytogenetic studies of maize and neurospora. Carnegie Inst. Washington Year Book, pp 146–152

McGrath JM, Shaw RS, de los Reyes BG, Weiland JJ (2004) Construction of a sugar beet BAC library from a hybrid with diverse traits. Plant Mol Biol Rep 22:23–28

Menzel G, Dechyeva D, Keller H, Lange C, Himmelbauer H, Schmidt T (2006) Mobilization and evolutionary history of miniature inverted-repeat transposable elements (MITEs) in *Beta vulgaris* L. Chromosom Res 14:831–844

Menzel G, Dechyeva D, Wenke T, Holtgräwe D, Weisshaar B, Schmidt T (2008) Diversity of a complex centromeric satellite and molecular characterization of dispersed sequence families in sugar beet (*Beta vulgaris*). Ann Bot 102:521–530

Moon S, Jung KH, Lee DE, Jiang WZ, Koh HJ, Heu MH, Lee DS, Suh HS, An G (2006) Identification of active transposon dTok, a member of the hAT family, in rice. Plant Cell Physiol 47:1473–1483

Moreno-Vázquez S, Ning J, Meyers BC (2005) *hATpin*, a family of MITE-like *hAT* mobile elements conserved in diverse plant species that forms highly stable secondary structures. Plant Mol Biol 58:869–886

Pritham EJ, Zhang YH, Feschotte C, Kesseli RV (2003) An Ac-like transposable element family with transcriptionally active Y-linked copies in the white campion, Silene latifolia. Genetics 165:799–807

Rubin E, Lithwick G, Levy AA (2001) Structure and evolution of the hAT transposon superfamily. Genetics 158:949–957

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA 81:8014–8018

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, New York

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

Schwarzacher T, Heslop-Herrison JS (1991) In situ hybridization to plant telomeres using synthetic oligomeres. Genome 34:317–323

Schwarzacher T, Heslop-Herrison JS (2000) Practical In situ hybridization. BIOS Scientific Publishers, Oxford, p 60

Takagi K, Maekawa M, Tsugane K, Iida S (2010) Transposition and target preferences of an active nonautonomous DNA transposon nDart1 and its relatives belonging to the hAT superfamily in rice. Mol Genet Genomics 284:343–355

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Vicient CM (2010) Transcriptional activity of transposable elements in maize. BMC Genomics 11:601

Vicient CM, Suoniemi A, Anamthawat-Jónsson K, Tanskanen J, Beharav A, Nevo E, Schulmann AH (1999) Retrotransposon *BARE-1* and its role in genome evolution in the genus *Hordeum*. Plant Cell 11:1769–1784

Vollbrecht E, Duvick J, Schares JP, Ahern KR, Deewatthanawong P, Xu L, Conrad LJ, Kikuchi K, Kubinec TA, Hall BD, Weeks R, Unger-Wallace E, Muszynski M, Brendel VP, Brutnell TP (2010) Genome-wide distribution of transposed dissociation elements in maize. Plant Cell 22:1667–1685

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, Schulman AH (2007) A unified classification system for eukaryotic transposable elements. Nat Rev Genet 8:973–982

Xu Z, Dooner HK (2005) *Mx-rMx*, a family of interacting transposons in the growing *hAT* superfamily of maize. Plant Cell 17:375–388

Yang G, Weil CF, Wessler SR (2006) A rice Tc1/mariner-like element transposes in yeast. Plant Cell 18:2469–2478

Zhang X, Wessler SR (2004) Genome-wide comparative analysis of the transposable elements in the related species *Arabidopsis thaliana* and *Brassica oleracea*. Proc Natl Acad Sci USA 101:5589–5594

Zhang X, Jiang N, Feschotte C, Wessler SR (2004) PIF- and Pong-like transposable elements: distribution, evolution and relationship with Tourist-like miniature inverted-repeat transposable elements. Genetics 166:971–986