copia-, *gypsy*- and LINE-Like Retrotransposon Fragments in the Mitochondrial Genome of *Arabidopsis thaliana*

Volker Knoop, Michael Unseld, Joachim Marienfeld, Petra Brandt, Sabine Sünkel, Henriette Ullrich and Axel Brennicke

Institut für Genbiologische Forschung, D-14195 Berlin, Germany and Allgemeine Botanik, Universität Ulm, D-89069 Ulm, Germany

Manuscript received June 30, 1995
Accepted for publication October 24, 1995

ABSTRACT

Several retrotransposon fragments are integrated in the mitochondrial genome of *Arabidopsis thaliana*. These insertions are derived from all three classes of nuclear retrotransposons, the Ty*1*/*copia*-, Ty*3*/*gypsy*- and non-LTR/LINE-families. Members of the Ty*3*/*gypsy* group of elements have not yet been identified in the nuclear genome of Arabidopsis. The varying degrees of similarity with nuclear elements and the dispersed locations of the sequences in the mitochondrial genome suggest numerous independent transfer-insertion events in the evolutionary history of this plant mitochondrial genome. Overall, we estimate remnants of retrotransposons to cover ≥5% of the mitochondrial genome in Arabidopsis.

INTRONS, duplicated sequences, additional genes and inserted chloroplast sequences account at least partly for the large size of the mitochondrial genome in plants (BENDICH 1993; SCHUSTER and BRENNICKE 1994). Additionally, nuclear DNA sequences have been reported in the mitochondrial genome of one plant, the evening primrose *Oenothera berteriana*, where a fragment of the cytoplasmic 18S rDNA sequence is found adjacent to a short open reading frame with similarity to reverse transcriptase (RT) sequences (SCHUSTER and BRENNICKE 1987). The identification of these two fragments in a plant mitochondrial genome has shown that nuclear sequences can be stably integrated into plant mitochondrial DNA. It is as yet unclear however, how frequently such transfers from nucleus to mitochondrion occur. Nuclear sequence data are available to date only for selected genes and their immediate vicinities thus prohibiting an exhaustive search of nuclear sequences in the mitochondrial genomes of plants. Of particular interest in this respect are sequences mobile in the nuclear genomes and the possibility that they reach the mitochondrial compartment. Retrotransposons contribute significantly to the expansion of nuclear DNA as evidenced for example by the high copy number of these elements in the genome of lily (SENTRY and SMYTH 1989; SMYTH 1991). How the amplification rate of individual elements is controlled is still unclear (GRANDBASTIEN 1992). In the 10 times smaller genome of *Arabidopsis thaliana*, the number of retrotransposon sequences appears to be much smaller than in other plants (KONIECZNY *et al.* 1991; VOYTAS *et al.* 1992). The increasing sequence information of mobile elements in the nucleus such as retrotransposons (see also the accompanying paper; WRIGHT *et al.* 1995) now allows a selective search for these sequences in mitochondrial genomes.

To evaluate the frequency of retrotransposon-like sequences in the mitochondrial genome of plants, we searched the consolidated sequence data presently available from the Arabidopsis mitochondrial genome sequencing project in our laboratory (KLEIN *et al.* 1994) for homologies to retrotransposons. Nine regions with high similarity to members of all three classes of these usually nuclear elements were identified in the mitochondrial DNA. The observed similarities between individual retrotransposon sequences in the nuclear and mitochondrial genomes suggest independent transfer events for the retrotransposon-like sequences from the nuclear to the mitochondrial compartment.

MATERIALS AND METHODS

Nucleotide sequence determination and analysis: The sequences of the mitochondrial genome of *A. thaliana* var. Columbia were determined in the course of the ongoing sequencing project of this genome. Cloning and mapping of the genome have been detailed previously (KLEIN *et al.* 1994). Determination and analysis of the sequence were done by fluorescent DNA sequencing protocols on EMBL automatic machines and will be described in detail elsewhere.

General sequence data handling was done with the UWGCG program package for VAX/VMS systems (Genetics Computer Group 1994). Sequence similarity searches used the BLASTX algorithm at the BLAST server at the National Institutes of Health in Bethesda, Maryland (ALTSCHUL *et al.* 1994). BLASTX translates the nucleic acid query sequence into six reading frames before screening the protein database.

BLAST searches were performed against the nr (nonredundant) combined protein database using default search parameters. Query sequences were extracted from 111.7-kb consolidated nucleotide sequences of the mitochondrial DNA organized in five contigs of 68.2, 21.2, 2.5, 10.8 and 9.0 kb, respectively. These contigs were split into separate queries of

Corresponding author: Axel Brennicke, Allgemeine Botanik, Universität Ulm, Albert-Einstein-Allee, D-89069 Ulm, Germany.

~3 kb to avoid computing and output limitations. A VAX/VMS shell script function to select a sequence range before submitting the query via e-mail was kindly provided for this purpose by S. KLOSKA, Berlin. The contig sequences will be part of a database entry comprising the entire mitochondrial DNA sequence of *A. thaliana* to be reported elsewhere. Regions of the query sequences with similarity to retrotransposon sequences identified in the BLASTX searches were translated into continuous peptide sequences by compensating for frame shifts, but keeping stop codon identities unaltered. Peptide sequences were rerun against the protein database using BLASTP with default settings. Only examples with a calculated probability of a random hit below 10^{-9} are reported here.

Southern blot analysis: Total cellular DNA was extracted from green leaf material of *A. thaliana* ecotypes by the CTAB method (DOYLE and DOYLE 1990). Mitochondrial DNA was purified from suspension culture cells of ecotype Columbia as described previously (KLEIN *et al.* 1994). DNA was cut with *Eco*RI before electrophoresis on a 1% agarose gel and transferred to PALL Biodyne B membranes by vacuum blotting. Hybridization was performed in 5× Denhardt's, 5× SSC, 0.1% SDS, 50% formamide and 100 μg/ml herring sperm DNA overnight, and blots were washed in 0.2× SSC, 0.1% SDS at 60° two times for 10 min. Blots were exposed for 5 days to Agfa Curix RP1 X-ray films.

RESULTS

Retrotransposon sequences in the mitochondrial genome of *A. thaliana*: As part of the routine preliminary scans of consolidated sequence information obtained in the mitochondrial sequencing project of Arabidopsis (KLEIN *et al.* 1994), we analyze reading frames larger than 100 amino acid codons by similarity screenings against the databases. When one of these reading frames revealed significant similarity to a retrotransposon sequence, we searched systematically the presently verified 111 kb of the mitochondrial genome sequence of Arabidopsis for further similarities to these elements with the BLASTX algorithm. This search identified nine retrotransposon-derived fragments (designated RT1–RT9) with lengths between 67 and 420 codons integrated in the mitochondrial genome of Arabidopsis (Figures 1 and 2).

Retrotransposon fragments are dispersed in the mitochondrial genome: The individual stretches of retrotransposon origin are located at separate regions of the mitochondrial genome (Figure 1). Fragments with similarities to the different retrotransposon classes are spread throughout the entire genome with no recognizable preference for hotspots of integration. The nine mitochondrial retrotransposon sequences are found in intergenic regions at varying distances to coding regions.

Mitochondrial sequences are derived from all three types of nuclear retrotransposons: The retrotransposon-related sequences in the mitochondrial genome of Arabidopsis reported here can be classified according to the individual similarities to nuclear representatives. In the nuclear genomes three different classes of retrotransposons have been distinguished, the Ty*1/copia*, Ty*3/gypsy* and LINE or non-LTR families (see GRANDBASTIEN 1992). The nine mitochondrial retrotransposon fragments can be clearly assigned to either of these three groups, with sequences RT1–RT3 and RT7–RT8 derived from Ty*1/copia* and RT5, RT6 and RT9 related to Ty*3/gypsy* retrotransposons (Figure 2). The single non-LTR/LINE-like sequence RT4 has also been identified independently in a scan for such sequences in total cellular DNA from Arabidopsis (see companion paper).

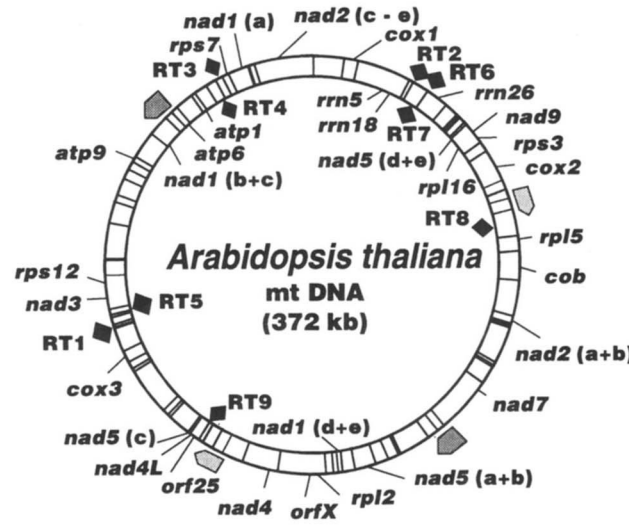

FIGURE 1.—Circular map of the mitochondrial genome of *A. thaliana* (KLEIN *et al.* 1994). Approximate positions of selected coding regions are indicated for orientation. Retrotransposon fragments RT1–RT9 (rhomboids) are distributed randomly around the genome without any connection to the recombinationally active repeated sequences (arrowheads).

Sequence similarity of the nine retrotransposon fragments to known nuclear retrotransposons is confined to the RT/RNase H domains in most instances. RT sequences are also found in intron-encoded reading frames (maturases), which could also be potential sources of the mitochondrial sequences RT1–RT9. Sequence similarities to these intron-encoded RTs are however insignificant and are confined to the signature amino acids of RTs. The mitochondrial sequences are thus clearly derived from retrotransposon structures.

Retrotransposon sequences in the mitochondrial genome show divergent degrees of conservation: Most of the retrotransposon sequences identified in nuclear genomes of various species are incomplete fragments or are distorted by reading frame shifts. These remnants are either degenerated from formerly complete sequences or originated from incomplete reverse transcription/amplification (VOYTAS *et al.* 1990, 1992; XIONG and EICKBUSH 1990). The retrotransposon sequences in the mitochondrial genome of Arabidopsis reported here are likewise incomplete and are frequently interrupted by translation termination codons or reading frame shifts. Higher degrees of degeneration of other mitochondrial retrotransposon-derived se-

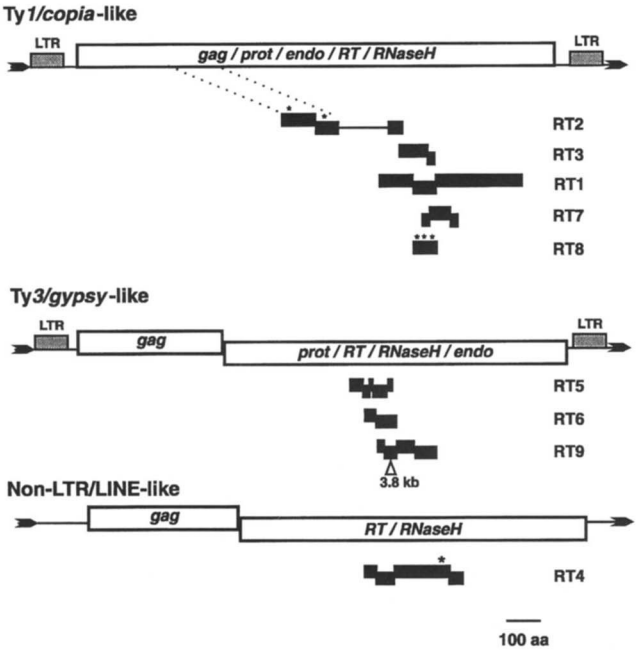

FIGURE 2.—Schematic alignment of the three classes of retrotransposons with the respective identified mitochondrial retrotransposon fragments RT1–RT9. Ty*1/copia*- and Ty*3/gypsy*-like retrotransposons carry long terminal repeats (LTR) that are absent from LINE-like sequences. The graphic presentation is schematic, not depicting the differences between individual retrotransposons in size and gene arrangement. The entire coding information of these retrotransposons can potentially be expressed as a large single polyprotein, *e.g.*, the 1440-amino acids-long reading frame of the Ty*1/copia*-like *hopscotch* element of maize (WHITE *et al.* 1994). Ty*3/gypsy*-like elements are distinguished by a different arrrangement of the endonuclease domain (endo) relative to the RT downstream of the nucleic acid-binding (gag) and protease (prot) domains. The extent of the mitochondrial insertions is shown as black bars with offset boxes indicating frameshifts. The locations of termination codons are indicated by *, and the 3.8-kb insertion of unknown origin in the RT9 element is indicated by △. Another sequence insertion of uncharacterized origin in element RT2 is represented by a thin line. . . . indicate similarity with the gag-sequence, adjacent to which a large part of the internal polyprotein sequence has been lost in RT2. Nucleotide sequences of RT1–RT8 and RT9a (upstream region) and RT9b (downstream region) are deposited in the database under accession numbers X91201–X91210.

quences may prohibit their clear identification. The mitochondrial retrotransposon sequences show divergent degrees of similarity to individual nuclear sequences suggesting different origins of the nine identified mitochondrial retrotransposon sequences. The group of Ty*3/gypsy*-like mitochondrial sequences RT5, RT6 and RT9 exemplifies this. In the protein database screens RT5 shows the highest similarity to the Ty*3* sequence of yeast (accession number S41556), while RT6 is best conserved to the *gypsy* transposon of Drosophila (acc. S26840), and RT9 displays the greatest similarity to the SKIPPY transposon of the fungus *Fusarium oxysporum* (acc. L34658). The previously identified RT reading frame in Oenothera mitochondria (SCHUSTER and BRENNICKE 1987) is also of the *gypsy* type and is very similar to RT9 and RT6 on the protein level. The Oenothera mitochondrial RT sequence shows the highest similarity to the second reading frame of the MAGGY retrotransposon from the fungus *Magnaporthe grisea* (acc. L35053) in screenings against the database.

Derived amino acid alignments are shown exemplary for each class of retrotransposon sequences, *i.e.*, for the mitochondrial element RT1 (Figure 3) of the Ty*1/copia* type, for RT5 as a Ty*3/gypsy* representative (Figure 4) and for RT4 as a non-LTR sequence (Figure 5). The Ty*1/copia*-like sequence RT1 shows significantly higher similarity to the maize *hopscotch* retrotransposon (WHITE *et al.* 1994) than to any of the Ty*1/copia* sequences as yet identified in the nuclear genome of Arabidopsis (VOYTAS *et al.* 1990). Similarities between mitochondrial and nuclear retrotransposon-like sequences are in general comparable to the conservation between nuclear retrotransposon families. These families are defined within the three classes of retrotranposon sequences by the degrees of conservation between their RT sequences (CASACUBERTA *et al.* 1995). In the nuclear as well as in the mitochondrial retrotransposon sequences, similarities are clustered in distinct regions, predominantly the RT and RNase H domains. Most of the mitochondrial sequences are identified with these functionally defined sequences (Figure 2).

A mitochondrial retrotransposon insertion predates ecotype evolution and speciation in Arabidopsis: The most striking sequence similarity is observed between the mitochondrial RT1 sequence and nuclear retrotransposons of the Ty*1/copia* class identifed in Arabidopsis and other plant species (Figures 2 and 3) (VOYTAS and AUSUBEL 1988; WHITE *et al.* 1994). Intriguingly sequence similarity is higher to the *hopscotch* retrotransposon of maize rather than to any of the known nuclear Arabidopsis sequences (Figure 3). Between the RT1 and maize *hopscotch* sequences 52% of all derived amino acids are identical and 70% are similar, while between RT1 and the most similar Arabidopsis nuclear retrotransposon (Ta1-2) only 40% of the codons are identical and 60% are similar.

To investigate whether the RT1 sequence identified in ecotype Columbia is also present in other ecotypes, total cellular DNAs from various Arabidopsis ecotypes were hybridized with an internal probe of the mitochondrial RT1 sequence (Figure 6). This experiment shows the RT1 sequence to be conserved in the mitochondrial genomes of all 12 analyzed ecotypes as evidenced by the presence of the 4.7-kb *Eco*RI fragment expected from the sequence data of ecotype Columbia. A second hybridizing *Eco*RI fragment is in accord with the internal *Eco*RI recognition site in the probe. For this fragment two alternative arrangements of adjacent downstream sequences are observed in the different ecotypes, one resulting in a 2.7-kb fragment, the other yielding an 8-kb *Eco*RI fragment. The mitochondrial

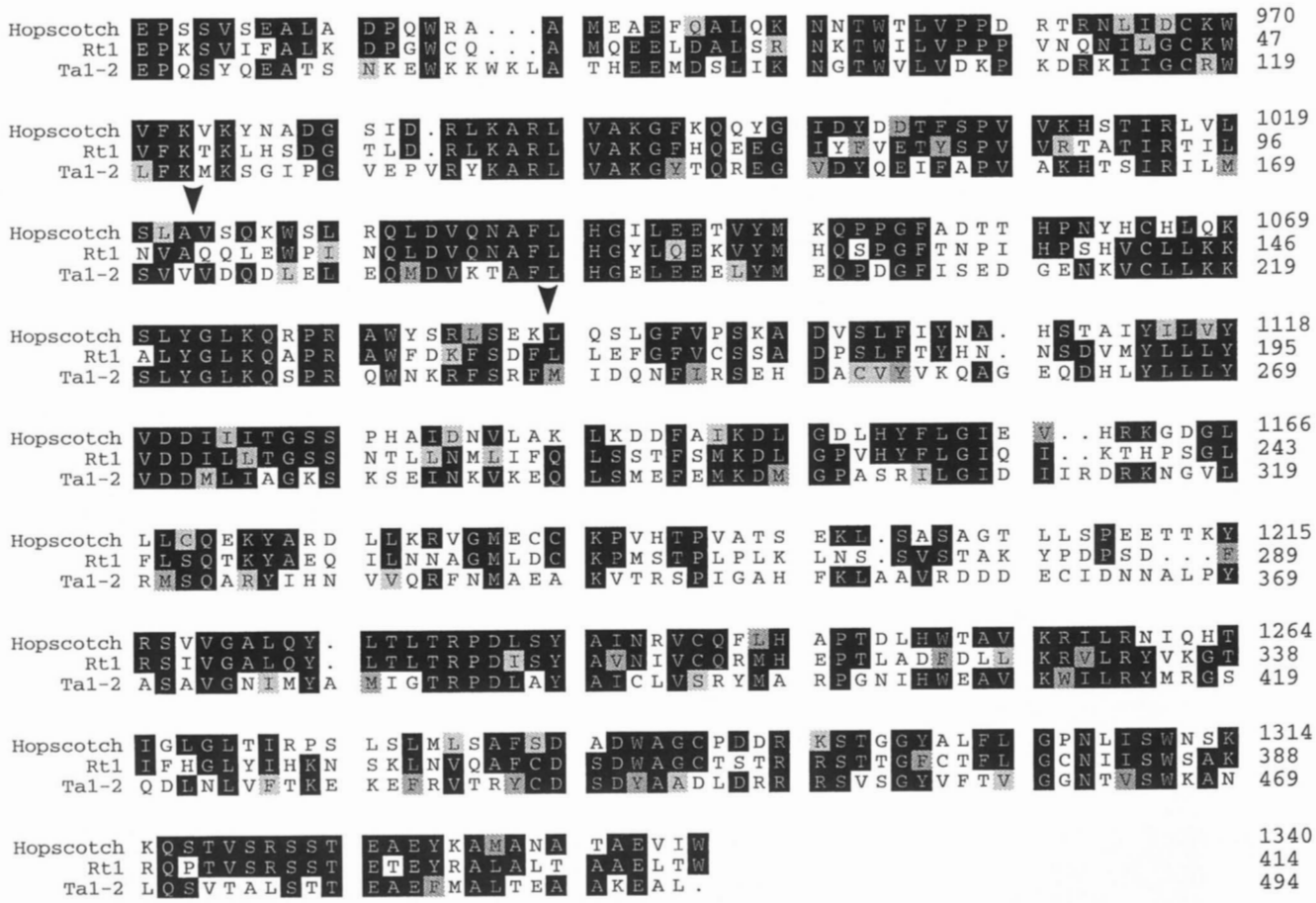

FIGURE 3.—Alignment of the *Zea mays hopscotch* retrotransposon-encoded polyprotein (1439 amino acids, accession number U12626) and hypothetical protein 3 (494 amino acids) of the Ta1-2 retrotransposon from *A. thaliana* (accession numbers S23315, X53975) with the mitochondrial retrotransposon fragment RT1 (nucleotide sequence accession number X91201). Numbering is according to database entries for *hopscotch* and Ta1-2, respectively. Two frameshifts (vertical arrowheads) are assumed to translate the nucleotide sequence of RT1. Gaps introduced for improved alignment are shown as dots. Between the Arabidopsis mitochondrial RT1 sequence and the *Z. mays hopscotch* sequence 52% of the amino acids are identical, while only 40% of the amino acids are conserved between the Arabidopsis mitochondrial and nuclear sequences RT1 and Ta1-2, respectively. Identical amino acids are indicated by black boxes, similar amino acids are shaded.

genome of ecotype Enkheim apparently contains both arrangements, which most likely are the result of mitochondrial intragenomic rearrangements. The additional weak and variable hybridization signals could indicate similarities to other mitochondrial sequences or alternatively to nuclear retrotransposon sequences. These observations are in line with the high diversity and the high rates of loss of retrotransposons in the nuclear genome of Arabidopsis (VOYTAS *et al.* 1990; KONIECZNY *et al.* 1991).

DISCUSSION

Retrotransposon sequences expand the mitochondrial genome: The mitochondrial retrotransposon fragments described here contribute significantly to the overall size of the mitochondrial genome in Arabidopsis. The nine regions with retrotransposon homology together cover 6 of the 111 kb mitochondrial DNA analyzed. Supposing a comparable frequency of retrotransposon sequences in the remaining 260 kb of the Arabidopsis mitochondrial genome, we can assume ~5% of the entire mitochondrial complexity to be derived from such retroelement sequences, considerably more than the proportion of retrotransposon sequences in the nuclear genome of Arabidopsis. In the nuclear genome of Arabidopsis retrotransposons either failed to amplify beyond a low copy number or were reduced severely during streamlining of the nuclear genome to only a few examples, which now contribute probably <1% to the nuclear sequence complexity (see companion paper).

A large portion of the Arabidopsis mitochondrial genomic sequences analyzed to date has no recognizable function and shows no significant similarity to any sequence in the present databases (M. UNSELD, unpublished observations). Much of the mitochondrial genome complexity in plants may be derived from imported nuclear sequences. Even if all of these as yet unassigned sequences were of nuclear origin, the pro-

FIGURE 4.—Amino acid-sequence alignment of the mitochondrial RT5 insertion as an example of the Ty*3/gypsy* class of retrotransposons with the open reading frame B of the Ty*3* retrotransposon from *Saccharomyces cerevisiae* (accession number Z46728), the deduced protein of 1745 amino acids from the retrotransposon F44e2-1 from *Caenorhabditis elegans* (accession number S44816) and the polyprotein of 1237 amino acids deduced from retrotransposon T412 of *Drosophila melanogaster* (accession number P10394). The accession number of the *A. thaliana* mitochondrial sequence RT5 is X91205. Four frameshifts (vertical arrowheads) are assumed to obtain the protein translation shown. Graphic details are as in Figure 3.

portion of retrotransposon sequences in the mitochondrial genome appears to be higher than in the nuclear genome. A preference for retrotransposon sequences over other nuclear sequences for import into the mitochondrial DNA could be due to actual retrotransposition events from the nuclear compartment to the mitochondrial genome. Virus-like particles as carriers of the cDNA sequence and integration into the mitochondrial genome mediated by element-encoded integrase and/or the frequent mitochondrial recombination events

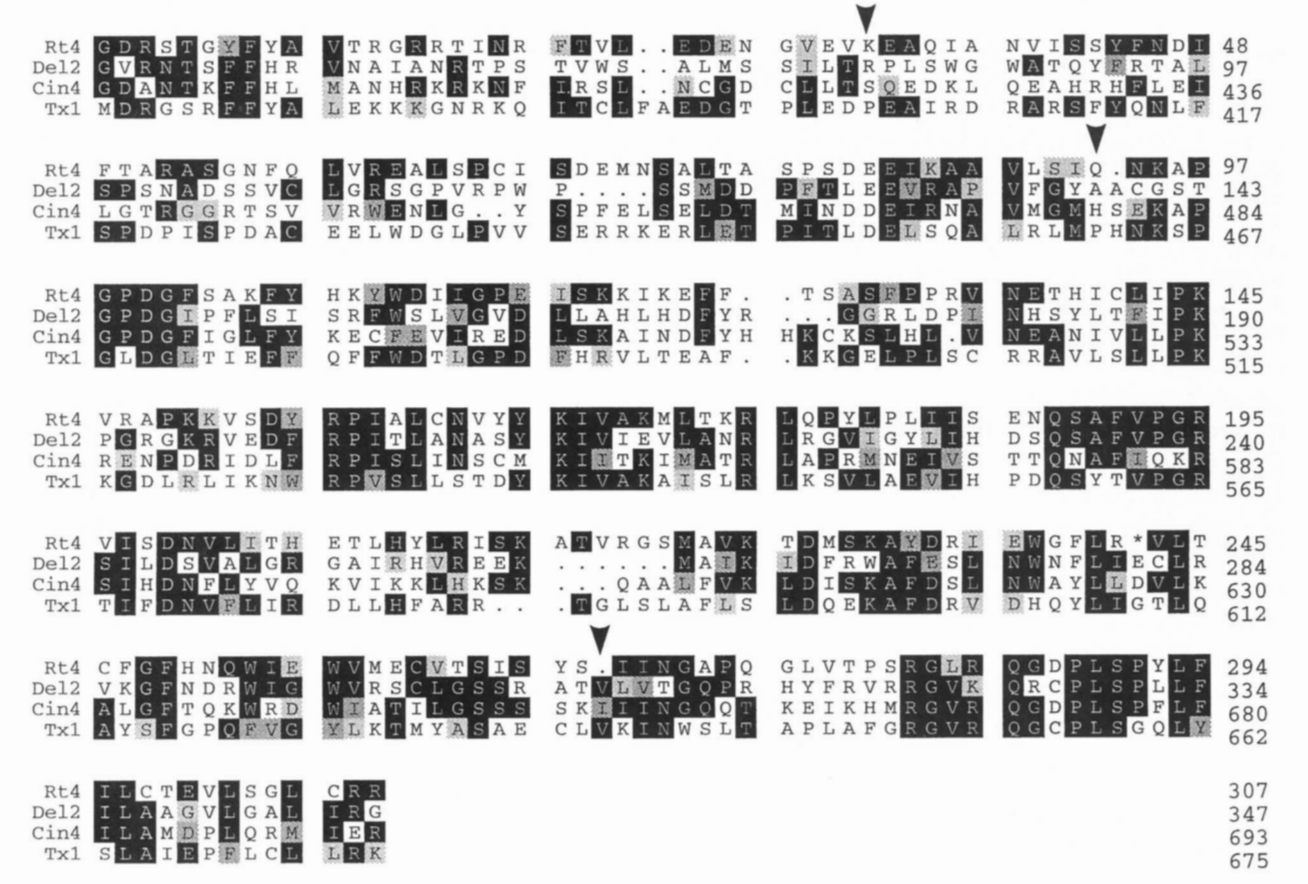

FIGURE 5.—Amino acid similarity of the mitochondrial RT4 sequence with non-LTR/LINE-like retrotransposons. The presentation is as in Figures 3 and 4. A stop codon in the RT4 frame is marked by *. Sequences can be found in the databases under X91204 for *A. thaliana* mitochondrial RT4, Y00086 for *Z. mays* cin4, Z17425 for *Lilium speciosum* del2 and P14381 for *Xenopus laevis* TX1. The 999-amino acids-long del2 polypeptide results from translation of Z17425 starting at position 1324 without an ATG startcodon.

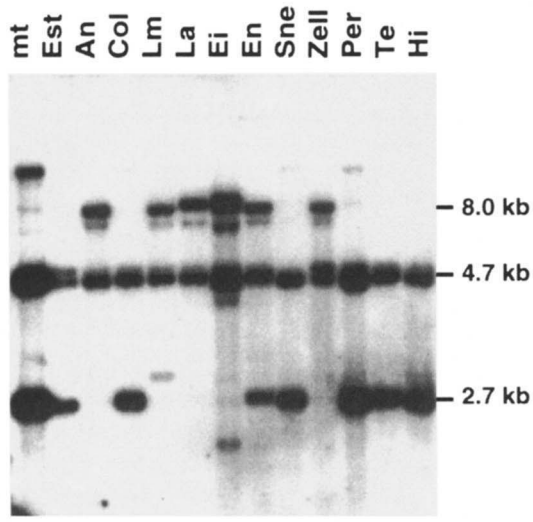

FIGURE 6.—Hybridization with the mitochondrial RT1 sequence shows the conservation of this sequence in other ecotypes of Arabidopsis. Total cellular DNA was cut with *Eco*RI and probed with an internal 300-bp *Bam*HI/*Pst*I fragment of the mitochondrial RT1 sequence. Two hybridizing fragments are observed due to the presence of an *Eco*RI site in the probe. The upstream region is conserved in the mitochondrial genomes of all ecotypes (4.7 kb fragment), while two alternative arrangements can be seen for the downstream sequence as 2.7- and 8.0-kb fragments. Ecotype Enkheim (En) contains both arrangements in the mitochondrial DNA. Weak hybridizations are due to either nuclear- or mitochondrial-encoded sequence similarities. Several fragments can be identified as mitochondrial homologies as they are also apparent in purified mitochondrial DNA (lane mt). Ecotypes are abbreviated as follows: Estland (Est), Antwerpen (An), Columbia (Col), Le Mans (Lm), Landsberg (La), Eifel (Ei), Enkheim (En), Sierra Nevada (Sne), Zermatt (ZeII), Perm (Per), Tenela (Te), Hilversum (Hi).

could promote such interorganellar transfer of retrotransposon sequences. The contribution of genuine retrotransposition is not excluded by the incomplete and interrupted structures of the retrotransposon sequences in the mitochondrial genome of Arabidopsis. Most of the presently known nuclear retrotransposon sequences likewise appear to be nonfunctional due to stop codons and frame shifts introduced by mutations (VOYTAS *et al.* 1992; WHITE *et al.* 1994).

It is conceivable that a massive influx of retrotransposon sequences into the organelle occurred in a progenitor of Arabidopsis that contained a much greater number of these sequences in the nuclear genome and that only later most of the nuclear copies were lost under the evolutionary pressure to streamline the nuclear genome. This scenario is supported by the finding that at least the RT1 retrotransposon insertion in the mitochondrial genome predates the ecotype evolution of Arabidopsis.

Multiple transfer/insertion events of retrotransposon sequences into the mitochondrial DNA: Each of the nine retrotransposon sequences reported here has most likely been independently transferred and integrated into the mitochondrial genome. Also within each of the three retrotransposon classes, different degrees of similarities between the mitochondrial RT fragments and distinct nuclear retrotransposons are observed, further emphasizing independent and unique origins in each instance. The mitochondrial insertions are thus most likely derived from different nuclear retrotransposons rather than having arisen from duplications of a unique mitochondrial insertion within this organelle.

Timing of these different transfer/integration events is at present hardly possible, since most likely not all of the current nuclear elements have yet been identified that could be related to potential progenitors. Only the LINE-like mitochondrial element RT4, which was also identified as Ta 17 by the approach used to scan the nuclear genome for retrotransposon sequences (see accompanying paper), clearly groups with the Arabidopsis LINE-like sequences rather than with those of other plant species. All other mitochondrial retrotransposon fragments show higher similarities to nuclear retrotransposon sequences from other species. These closer sequence relationships to other species, as for example between RT1 and a maize retrotransposon, can be explained by three alternative scenarios. First, a comparatively recent transfer event may have originated from an as yet unidentifed progenitor element in the nuclear DNA of Arabidopsis. Second, a recent interspecific exchange between the nuclear or mitochondrial genomes of other plant species and the mitochondrial DNA of Arabidopsis could be considered. Third, ancient transfers from the Arabidopsis nuclear genome inserted these sequences into the mitochondrial genome after which the respective progenitor sequences were lost from the nuclear genome. This latter scenario of ancient sequence transfers from nucleus to mitochondrion is to be considered the most likely one, in which transposition was initiated from retrotransposons that did not survive in the nuclear genome. The progenitors were either lost completely or degenerated beyond recognition in the nucleus of Arabidopsis during the extensive streamlining of this genome, which is suggested by the comparatively small size of this genome and the dense population of coding sequences.

The weak hybridization signals observed when probing total cellular DNA from Arabidopsis with the mitochondrial RT1 sequence, which at present shows the highest similarity to the maize *hopscotch* sequence, may be due to retrotransposon copies with low sequence similarity or to short fragments of the nuclear RT1 progenitor. Such direct probings with these mitochondrial sequences might offer an approach to identify as yet unknown retrotransposon sequences in the nuclear genome of Arabidopsis, which are more closely related to the homologous mitochondrial sequences. Due to its slow primary sequence evolution plant mitochondrial DNA will have conserved the originally transferred ret-

rotransposon sequence much better than the nuclear genome.

Ty*3/gypsy* retrotransposon sequences identified in the mitochondrial, but not the nuclear, genome of Arabidopsis: Three of the mitochondrial retrotransposon sequences, RT5, RT6 and RT9, group clearly with the Ty*3/gypsy* retrotransposons, which have not yet been found in the nuclear genome of Arabidopsis. This class of sequences, however, is present in the nuclear genomes of other plant species as evidenced by the dell retrotransposon in lily, the IFG7 sequence in pine and magellan in corn (PURUGGANAN and WESSLER 1994). The Ty*3/gypsy* class of retrotransposons may not occur in the Arabidopsis nuclear genome either due to their never having arrived in this genome or because of having been eliminated during streamlining of the nuclear DNA.

It is more likely, however, that Ty*3/gypsy* sequences are present but have just not yet been identified in the Arabidopsis nuclear genome. Combined oligonucleotide and PCR approaches, such as those used to identify representatives of the non-LTR class of retrotransposons (see companion paper), will provide a comparatively rapid answer. The mitochondrial sequences identified here may be useful to design derived consensus oligonucleotides for this experimental approach.

We thank DANIEL F. VOYTAS for communicating unpublished results and for helpful comments on the manuscript and SEBASTIAN KLOSKA for providing the shell script for more convenient BLAST searches. Research in the authors' laboratories was supported by the Bundesministerium für Forschung und Technologie in the Genzentrum Berlin and a Landesschwerpunkt from the State of Baden-Würtemberg.

LITERATURE CITED

ALTSCHUL, S. F., M. S. BOGUSKI, W. GISH and J. C. WOOTTON, 1994 Issues in searching molecular sequence databases. Nature Genet. **6:** 119–129.

BENDICH, A. J., 1993 Reaching for the ring: the study of mitochondrial genome structure. Curr. Genet. **17:** 421–425.

CASACUBERTA, J. M., S. VERNHETTES and M.-A. GRANDBASTIEN, 1995 Sequence variability within the tobacco retrotransposon Tnt1 population. EMBO J. **14:** 2670–2687.

DOYLE, J. J., and J. L. DOYLE, 1990 Isolation of plant DNA from fresh tissue. FOCUS **12:** 13–15.

Genetics Computer Group, 1994 Program manual for the GCG package version 8. Madison, WI.

GRANDBASTIEN, M.-A., 1992 Retroelements in higher plants. Trends Genet. **8:** 103–108.

KLEIN, M., U. ECKERT-OSSENKOPP, I. SCHMIEDEBERG, P. BRANDT, M. UNSELD *et al.*, 1994 Physical mapping of the mitochondrial genome of *Arabidopsis thaliana* by cosmid and YAC clones. Plant J. **6:** 447–455.

KONIECZNY, A., D. F. VOYTAS, M. P. CUMMINGS and F. M. AUSUBEL, 1991 A superfamily of retrotransposable elements in *Arabidopsis thaliana*. Genetics **127:** 801–809.

PURUGGANAN, M. D., and S. R. WESSLER, 1994 Molecular evolution of magellan, a maize Ty3/*gypsy*-like retrotransposon. Proc. Natl. Acad. Sci. USA **91:** 11674–11678.

SCHUSTER, W., and A. BRENNICKE, 1987 Plastid, nuclear and reverse transcriptase sequences in the mitochondrial genome of Oenothera: is genetic information transferred between organelles via RNA? EMBO J. **6:** 2857–2863.

SCHUSTER, W., and A. BRENNICKE, 1994 The plant mitochondrial genome—physical structure, information content, RNA editing and gene migration to the nucleus. Annu. Rev. Plant Physiol. Plant Mol. Biol. **45:** 61–78.

SENTRY, J. W., and D. R. SMYTH, 1989 An element with long terminal repeats and its variant arrangements in the genome of *Lilium henryi*. Mol. Gen. Genet. **215:** 349–354.

SMYTH, D. R., 1991 Dispersed repeats in plant genomes. Chromosoma **100:** 355–359.

VOYTAS, D. F., and F. M. AUSUBEL, 1988 A *copia*-like transposable element family in *Arabidopsis thaliana*. Nature **336:** 242–244.

VOYTAS, D. F., A. KONIECZNY, M. P. CUMMINGS and F. M. AUSUBEL, 1990 The structure, distribution and evolution of the Ta1 retrotransposable element family of *Arabidopsis thaliana*. Genetics **126:** 713–721.

VOYTAS, D. F., M. P. CUMMINGS, A. KONIECZNY, F. M. AUSUBEL and S. R. RODERMEL, 1992 *Copia*-like retrotransposons are ubiquitous among plants. Proc. Natl. Acad. Sci. USA **89:** 7124–7128.

WHITE, S. E., L. F. HABERA and S. R. WESSLER, 1994 Retrotransposons in the flanking regions of normal plant genes: a role for *copia*-like elements in the evolution of gene structure and expression. Proc. Natl. Acad. Sci. USA **91:** 11792–11796.

WRIGHT, D. A., N. KE, J. SMALLE, B. HAUGE, H. M. GOODMAN *et al.*, 1995 Multiple non-LTR retrotransposons in the genome of *Arabidopsis thaliana*. Genetics **142:** 569–578.

XIONG, Y., and T. H. EICKBUSH, 1990 Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J. **9:** 3353–3362.

Communicating editor: J. CHORY