TRUNCATED NON-NUCLEAR TRANSPOSABLE ELEMENTS IN GRAPEVINE: A MINI REVIEW

A.V. Milovanov[1], J.Tello[2], U.C.M. Anhalt[2], A. Forneck[2]

[1]Kuban State Agrarian University, Department of Viticulture, Krasnodar, Russia
[2]University of Natural Resources and Life Sciences, Vienna, Department of Crop Sciences, Division of Viticulture and Pomology, Tulln, Austria

In this mini-review we present insight to the non-nuclear transposable elements and *in silico* analysis of miniature inverted transposable elements (MITEs) in the grapevine mitochondrial genome. Here we report the identification of 17 truncated sequences in grapevine (*Vitis vinifera* L.) mitochondrial genome which expectedly belongs to the four ancient transposon families (hAT, Tc1Mariner, Mutator and PIF/Harbinger). Some sequences with a high rate of homology in chloroplast and nuclear genomes were also identified. Thus, it suggests the intercellular gene transfer between these three organelles. These partial sequences showed a high level of similitude with full MITE sequences, and they were found in their inner region, supporting their MITE origin. Further analysis revealed these sequences in other life kingdoms (including eubacteria and archaea), which indicates their ancient origin. Further research showed that 13 out of the 17 sequences are conserved domains of the genes where they are located, suggesting their contribution to gene evolution. Therefore, we suppose that more studies of nature, origin and functional meaning of these sequences and their fusion with genes are necessary. In the light of our observations it will be useful for further studies of *V. vinifera* genome organizing and systematics, as well as for other species.

Vitis vinifera L., mitochondrial genome, mobile elements, MITEs, intracellular gene transfer, ancient transposon families

sciendo

doi: 10.2478/sab-2019-0030
Received for publication on March 15, 2018
Accepted for publication on September 9, 2019

INTRODUCTION

Transposable elements (TEs), also well-known as 'selfish genes', are ubiquitous parts of DNA that can move into new location of chromosomes and create its copies. Originally, they were discovered by Barbara McClintock more than half a century ago as structures that are responsible for the pigmentation mutations in maize (McClintock, 1951). A further study revealed more of their various roles in genetics, functions and host genome evolution for different species. One of the most demonstrative examples was displayed on the grapevine cultivars Ruby Okuyama and Flame Muscat (Kobayashi, 2004). The berry colour of these cultivars was changed by movement of *Gret1* element and the alleles of *VvmybA1* gene were discovered. Then, comparative analysis of grass genomes demonstrated different roles of TEs in various genetic content (Morgante et al., 2007). Another interesting example was shown for rice; its genome contains chimeric transposable elements called Pack-MULEs (Jiang et al., 2004). Surprisingly, these TEs contain fragments derived from different cellular genes and multiple chromosomal loci that are fused and form open read frames. Moreover, some of them are active and expressed in chimeric transcripts. Other Pack-MULE resembling element was found in *Glycine max* L. and it was called Tgm-Express1 (Zabala et al., 2005). This TE contains several truncated cellular genes and, significantly, causes mutation in Wp locus, which consequently leads to pink phenotype.

There exist extensive reports on TEs in different species (Kleckner, 1981; Daboussi, Capy, 2003; Wicker et al., 2007; Bennetzen, Wang, 2014; Sotero-Caio et al., 2017). The TEs are commonly divided into two major groups according to their transposition mechanism, which can be RNA (class I-TEs) or a DNA intermediate (class II-TEs). Class I-TEs (also known as retrotransposons) is divided into two subgroups: LTR-retrotransposons (which have long terminal repeats and are subsequently divided into autonomous and non-autonomous TEs) and non-LTR. Autonomous LTR-retrotransposons encode a capsid-like protein and reverse transcriptase (*gag*

and *pol* genes), whereas non-autonomous do not. Non-LTR elements are divided into two subgroups: LINEs (long interspersed nuclear elements) and SINEs (short interspersed nuclear elements). They encode ORF1, *gag*-like protein and reverse transcriptase. Lastly, summarizing the Class 2 of transposons, they have terminal inverted repeats, encode a transposase and have several target-site duplications which create together a different length for each transposon family (Feschotte et al., 2002).

Since the discovering of TEs, huge data have accumulated leading to the formation of various databases (Du et al., 2010; Chen et al., 2013; Bao et al., 2015). In this regard, several works were dedicated to the classification of TEs and explained how they are organized in families as well as their mutual relationships (Wicker et al., 2007; Kapitonov, Jurka, 2008). Also, this data allowed conducting a comparative analysis of structure and revealed a function relevance of TEs. Ultimately, the most important findings were: (1) loss of TEs during the evolution of species; (2) different distribution and content of TEs in various species; (3) horizontal transfer of TEs; (4) amplification and removal of TEs in plants are rapid processes; (5) gene creation and diversity in plants caused by transposon's activity (for more detailed information see Bennetzen, 2005; Velasco et al., 2007; Carrier at al., 2012; Bennetzen, Wang, 2014; Di Genova et al., 2014; Chalopin et al., 2015; Chuong et al., 2016).

In the present article, we briefly review currently available information on nuclear, chloroplast and mitochondrial TEs in grapevine (*Vitis vinifera* L.). Additionally, we present some initial results dealing with the *in silico* identification of MITE sequences (a type of class-II TEs) in the grapevine mitochondrial genome. Then, we searched for their presence in chloroplast and nuclear genomes, for suggesting likely transfer-insertion events in the evolutionary process of this important crop.

TEs in grapevine nuclear genome

TEs are ubiquitous in the plant kingdom, being suggested to be the most common genetic elements contributing to plant nuclear genome organization and evolution (Feschotte et al., 2002). TEs are the single largest component of the genome of most plant species, like seen in Arabidopsis (The Arabidopsis Genome Initiative, 2000), rice (International Rice Genome Sequencing Project, 2005), and, more recently, coffee (Denoeud et al., 2014) and poplar genomes (Pinosio et al., 2016). Regarding grapevine, the sequencing of the highly homozygous PN40024 genotype revealed that 41.4% of its genome is composed of TEs (Jaillon et al., 2007) and its comparison with Sultanina genome revealed that TEs corresponded to approximately 80% of the total amount of its repetitive sequences (Di Genova et al., 2014). This fraction might have contributed to the generation of new genotypes and (to some extent) new phenotypes during grapevine evolution (Carrier et al., 2012). Indeed, the insertion of different TEs in the promoter of different grapevine genes (like *Gret1*, *Hatvine1-rrm* and *Mila-flb* TEs in *VvmybA1*, *VvTFL1A*, and *VvPI* genes, respectively) has been reported to be the main factor causing somatic variations that affect relevant quality traits like berry colour, flowering date, crop yield, size and compactness of grape clusters (Kobayashi, 2004; Fernandez et al., 2010, 2013, 2014).

As mentioned previously, TEs can be divided into two main classes according to their transposition intermediate: RNA (class I TEs, or retrotransposons) or DNA (class II TEs, or DNA transposons). Different subclasses, orders, superfamilies and families have been further described for these two classes (Wicker et al., 2007), and diverse in-depth genome-wide analyses have provided a detailed overview of some TE groups in grapevine (Benjak et al., 2008, 2009; Moisy et al., 2008; Wang et al., 2013). One of the most abundant class II TEs are the Miniature Inverted Transposable Elements (MITEs) family, which groups a set of TEs characterized by their small size (usually, less than 600 bp) and the presence of terminal inverted repeats (Feschotte et al., 2002). They are frequently found in plants within or close to genes, suggesting their great potential to modify gene expression (and occasionally phenotype) upon mobilization (Guermonprez et al., 2012; Vitte et al., 2014). Representatives of four 'cut-and-paste' grapevine transposon superfamilies have been previously described, and potential MITE sequences related to the CACTA, hAT and PIF superfamilies were reported by Benjak et al. (2008). Following this work, some subfamilies were identified, and their relative abundance and close association to genes indicate their potential impact on gene evolution (Benjak et al., 2009).

TEs in grapevine chloroplast genome

One of the first steps in detecting TEs in chloroplast genomes made Thompson et al. (1981) and Ohyama et al. (1983). Later, the research teams of Sugita and Kolodner indicated inverted repeats as possible candidates for TEs of plant's chloroplast DNA (Kolodner, Tewari, 1979; Sugita et al., 1984; Milligan et al., 1989). Thus, only transposon-mediated fossils without functional meaning were found by Zhou et al. (1988), Reboud, Zeyl (1994) and Fortune et al. (2008) before 1995, when two DNA elements called 'Wendi' were explored by Fan et al. (1995). A year later, in 1996, the first monocot's chloroplast TE was proposed by Prina (1996). But, unfortunately, there were still no efforts to detect TEs in the Vitaceae chloroplast genome.

Table 1. Truncated MITE sequences identified in *Vitis vinifera* L. mitochondrial genome

MITE	P-MITE Database ID	MITE family	Chloroplast genome[1]	Nuclear genome[1,2]	Gene ID[3]
MT-01	SQ123090308	hAT	–	+ (3, 16)	VIT_03s0110g00200 (intron); VIT_16s0022g00160(exon)
MT-02	SQ161007872	Mariner	–	+ (1, Un)	intergenic region/s
MT-03	SQ181097200	Mariner	+	+ (18)	intergenic region/s
MT-04	SQ182098594	Mutator	+	+ (12)	VIT_12s0055g00520 (exon); VIT_12s0055g00530 (exon)
MT-05	SQ183115307	hAT	+	+ (1, 9)	VIT_01s0011g00480 (intron); VIT_09s0002g08330 (exon); VIT_09s0002g08340 (exon)
MT-06	SQ202041669	Mutator	+	+ (10, 12)	intergenic region/s
MT-07	SQ205090478	PIF/Harbinger	+	+ (7, 10, 16)	VIT_07s0129g00790 (exon); VIT_16s0013g00330 (exon)
MT-08	SQ205100655	PIF/Harbinger	–	+ (3, 12)	VIT_03s0110g00190 (exon/intron); VIT_12s0028g02450 (exon)
MT-09	SQ265152084	PIF/Harbinger	+	+ (11, 13)	VIT_11s0052g01680 (exon); VIT_13s0019g02630 (exon)
MT-10	SQ302142638	Mutator	+	+ (6, 8)	VIT_06s0004g07410 (exon)
MT-11	SQ372011074	Mutator	+	+ (Un)	intergenic region/s
MT-12	SQ385086057	PIF/Harbinger	+	+ (11)	intergenic region/s
MT-13	SQ393043634	hAT	+	+ (5)	VIT_05s0029g01200 (intron)
MT-14	SQ395060092	PIF/Harbinger	–	+ (Un)	intergenic region/s
MT-15	SQ442012989	Mutator	–	–	n.i.
MT-16	SQ452094622	Mutator	+	+ (7)	VIT_07s0031g03000 (exon)
MT-17	SQ465051794	PIF/Harbinger	–	–	n.i.

[1]+/– = presence and/or absence, [2]number of chromosome is in brackets; [3]data obtained from EnsemblPlants and VTCdb: ViTis Co-expression Database; n.i. = not identified, Un = Unplaced contig

According to Jansen et al. (2006) the grapevine chloroplast genome extents 161 kbp, it includes several inverted repeats of 26 kbp; a similar rearrangement to that was found in other angiosperms and up to 20 different inverted repeats were found. In the light of our future findings, notably one 39-bp repeat was found three times, whereas one time in intergenic spacer. To our knowledge, no studies have detailedly identified TEs in the grapevine chloroplast genome so far, whereas in other species some of them have already been dealt with (Bendich, 2013).

TEs in grapevine mitochondrial genome

Insights to the TEs identification in mitochondrion were done for different species. For example, one of the first TE fragments was found in the mitochondrial genome of *Arabidopsis thaliana* L. and recognized as retrotransposon's remnants (Knoop et al., 1996). Then 21 *gypsy*-like retrotransposon sequences were identified in the sugar beet (Kubo et al., 2000) and 19 TE fragments in the rice mitochondrion (Notsu et al., 2002). Overall, according to Goremykin et al. (2012) sequences of TEs, reported for mitochondrial genomes, vary from 0.1 to 6.4% (also reviewed by Kubo, Mikami, 2007).

V. vinifera mitochondrial genome was described in detail by Goremykin et al. (2008). It was characterized as a 780 kbp long structure that contains 38 genes and 30 fragments from the chloroplast genome, suggesting a rampant gene transfer between these two organelles. On the other hand, several evidences indicate that intracellular gene transfer (IGT) between the three DNA-containing organelles might have occurred during the evolution of flowering plants, more specifically during the determination of the different cell organelles (Nugent, Palmer, 1991; Adams et al., 1999; Huang, 2005; Rousseau-Gueutin et al., 2013). In this regard, the high similarity between individual TE sequences found in the nuclear and mitochondrial genomes of *A. thaliana* has been indicated as an example of TE transfer between these two organelles (Knoop et al., 1996).

Considering that *V. vinifera* has one of the largest mitochondrial genomes described so far, and that its large size is mainly due to the expansion of the

spacer regions (around 90% of mitochondrial genome) (Goremykin et al., 2008), grapevine is a nice target for the in-depth analysis of mitochondrial TEs (mtTEs) and for the analysis of IGT between nuclear, chloroplast and mitochondrial genomes.

Can MITE-related sequences be identified in the grapevine mitochondrial genome?

To prove the presence of MITE sequences in the grapevine mitochondrial genome, the MITE-related sequences of the P-MITE database (Chen et al., 2013) were used as templates of a comparison analysis. *V. vinifera* mitochondrial genome sequence (GenBank code: NC_012119.1) was BLASTed in the P-MITE database (Chen et al., 2013) to detect homologous MITE sequences with: (I) P-MITE BLAST e-value < 0.001; (II) identity of candidate sequence and full MITE sequence > 80%; (III) location of candidate sequences (Supplementary Online Material 1); and (IV) location of candidate sequences in predicted gene sequences. We selected 17 candidate mitochondrial genome regions (Table 1 and Supplementary Online Materials 1 and 2) associated with four MITE families (Table 1). These 17 selected regions were BLASTed in NCBI (Altschul et al., 1990), NCBI-CD (Marchler-Bauer, Bryant, 2004), Chloroplast Genome Database (Cui, 2006), EnsemblPlants (Kersey et al., 2016), Grape Genome Browser (Jaillon et al., 2007) to identify homologue sequences. Identified sequences were aligned using Clustal Omega (Goujon et al., 2010) to establish homology. The average length of 17 sequences was 192 bp, although a high degree of variation was seen, detecting homolog fragments from 337 to just 58 bp long (MT-09 and MT-10, respectively). The length of the P-MITE sequence actually covered by the fragment detected in the grapevine mitochondrial genome was also highly variable, covering from 9.1% (for MT-09) to 60.5% (for MT-14) of the original sequence. We did not find any relationship between sequence length and value of obtained coverage. Sequences (MT-01 to MT-17) were then compared in detail with their corresponding homolog from the P-MITE database (Table 1). Interestingly, such mitochondrial sequences were identified as truncated (incomplete) instead of full, since in all cases the truncated sequence was aligned in the inner region of the full MITE sequence (Supplementary Online Material 1), and in none of the cases we found the 5' and 3' regions of the original sequence. Analogously, in a work aimed to evaluate the similarity between retrotransposon sequences in the nuclear and mitochondrial genomes of *A. thaliana*, Knoop et al. (1996) suggested that the presence of incomplete retrotransposon sequences can be a result of degeneration of formerly complete sequences, or incomplete reverse transcription/amplification processes. Previous works propose that these MITE families (hAT, Tc1Mariner, Mutator and PIF/Harbinger) are some of the most ancient ones, and probably they originated before the separation of life kingdoms (Rubin et al., 2001; Kapitonov, Jurka, 2004; Wicker et al., 2007). Therefore, we observed the presence of most of the sequences (MT-1, MT-2, MT-3, MT-4, MT-5, MT-7, MT-8, MT-9, MT-11, MT-12, MT-13 and MT-16) in ancient bacterial genomes and two of them (MT-2 and MT-12) in archaeal genomes. According to these results and the fact that some of them were located in the coding region of some nuclear and, consequently, mitochondrial genes (Table 1), we suggest that the detected TEs might have played an important role during the evolution process, being recruited as structural parts of hosting genes. This fact has already been indicated for different species (Robertson, 2002; Cordaux et al., 2006; Feschotte, 2008), but, to the best of our knowledge, it has not been previously reported for grapevine organelles genomes.

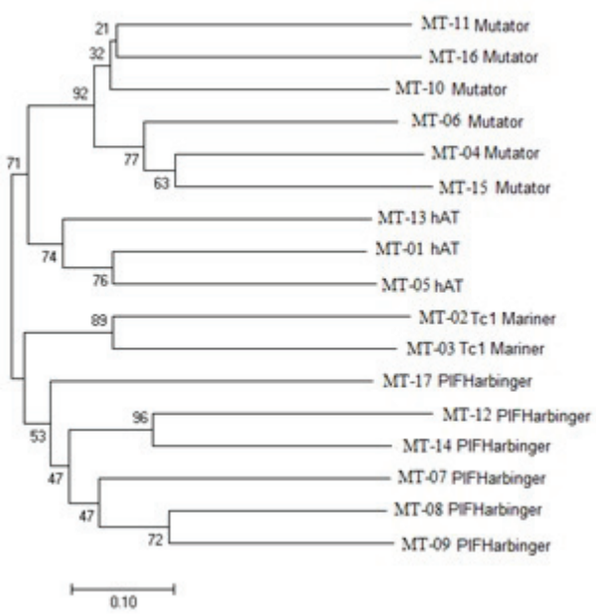

Fig. 1. Neighbour-joining phylogenetic tree of truncated MITE sequences (MT-01 to MT-17). Numbers at the nodes indicate bootstrap values (2000 replications)

The selected sequences were subjected to phylogenetic analysis (Fig. 1) with MEGA7 software (Kumar et al., 2016). Results indicated the presence of two main clusters, each one with two clear sub-clusters. The first main cluster is composed of the sequences of the Mutator and hAT families, whereas the second one is composed by those of the Mariner and PIF/Harbinger families, which is in general agreement with previous works (Wicker et al., 2007; Yuan, Wessler, 2011; Santos et al., 2016). Nonetheless, and according to Wicker et al. (2007), Mutator family should be placed between hAT and PIF/Harbinger

families, but in our cluster analysis it was on one of the edges of the tree. It has been indicated that different incongruences may appear during the construction of phylogenetic trees, attributed to artifacts during tree reconstruction procedures, an insufficient number of characters and/or gene paralogy (Goremykin et al., 2012). Further analyses revealed that truncated MITEs of Mutator family were distributed in a low number of living kingdoms (species), whereas other MITE families were widely distributed. Particularly MT-02 and MT-12 were found in archaeal genomes. If we compare the distribution of hAT and Mutator families, we can see that 3 of 6 Mutator sequences were found in plants genomes only (other sequences of Mutator were found in bacterial and plants genomes only), whereas hAT family is presented in every kingdom except for archaea.

BLAST search with other kingdoms revealed that sequences MT-02 and MT-03 (Mariner) were found in organisms like *Candidatus Thalassoarchaea* L. and *Thiomicrospira crunogena* L. and some endosymbionts (*Trabutina mannipara* L. and *Buchnera aphidicola* L.) (data not shown). Similarly, some sequences identified as hAT members were found in different bacterial genomes: MT-01 in *Candidatus Pelagibacter* L. (one of the smallest self-replicating one-cell organisms (Giovannoni, 2005) and in *Wolbachia* sp., and MT-05 in the genome of *Acaryochloris marina* L. Additionally, MT-4, MT-11 and MT-13 were recognized as parts of different transformation vectors. Together these findings add supplementary evidence for a likely ancient nature of these truncated transposons.

Lastly, we searched for conserved domains in 17 sequences using the NCBI CD-search tool (Supplementary Online Material 3). As a result, 13 sequences (MT-01, MT-03, MT-04, MT-05, MT-07, MT-08, MT-09, MT-10, MT-11, MT-13, MT-14, MT-16 and MT-17) were identified as conserved domains. All of these sequences (except for MT-14) were recognized as conserved domains of homologue genes from mitochondrial, chloroplast and nuclear genomes in grapevine and other species' genomes. Interestingly, MT-14 (previously identified as a Harbinger family member) was identified as a conserved domain of a gene coding for a Ribonuclease H-like superfamily. As previously indicated, MT-14 belongs to the Harbinger family and transposases are related to retroviral integrases (Rice, Baker, 2001; Majorek et al., 2014), which can be considered as another evidence of ancient nature of these sequences.

Are homologous MITEs in grapevine mitochondrial, nuclear and chloroplast genomes a case of IGT?

A significant relationship between lateral gene transfer (LGT) and gene translocation has been suggested because translocated genes have a higher probability to be truncated (Hao, Golding, 2009). To evaluate if the identified sequences are the result of intracellular gene transfer (IGT) between organelles, they were used as templates to detect homologous sequences in the nuclear and chloroplast grapevine genomes (Table 1). Results indicate that 11 and 15 out of 17 mitochondrial sequences were found in the grapevine chloroplast and nuclear genomes, respectively. In this regard, we found homologous sequences for MT-03, MT-04, MT-05, MT-06, MT-07, MT-09, MT-10, MT-11, MT-12, MT-13 and MT-16 in the three grapevine genomes, which might suggest gene transfer between these three organelles.

Major events of eukaryotic evolution were caused by their endosymbiosis with mitochondrion and plastids, and during this evolution process, horizontal gene transfer was a common natural phenomenon (Campbell, 2000; Koonin et al., 2001; Zhaxybayeva, 2006; Keeling, Palmer, 2008). According to our findings, some of grapevine truncated TE sequences identified here are present in various bacteria and cyanobacteria genomes, too (data not shown), and some of them were also found in three grapevine DNA-containing organelles (Table 1). It was suggested that ancient events of horizontal gene transfer happened between single-cell organisms and in the early stages of evolution (Brown, 2003), and this was reported also for grapevine (Goremykin et al., 2008; Bock, 2010). BLAST of TE sequences with grapevine mitochondrial genome revealed that some of them were located in coding regions (Table 1). To identify IGT between chloroplast and mitochondrial genomes we took GeneBank codes of chloroplast genes to find homologous sequences in grapevine mitochondrial genome. Most of the chloroplast sequences showed a high similarity with mitochondrial genome. In the case of MT-9, MT-10 and MT-12, only partial sequences were observed, but with a high homology level. This fact may suggest gene truncation by IGT processes, as shown previously (Hao, Golding, 2009), reinforcing the evidence of IGT in grapevine.

Location of sequences identified in the grapevine nuclear genome

Lastly, we searched for the physical location of these truncated sequences in the annotated nuclear genes, using the structural and functional data of the grapevine genome (12x) (Grimplet et al., 2012), based on the genes identified in PN40024 (Jaillon et al., 2007). According to this analysis, nine truncated sequences were found in annotated genes, whereas six of them in intergenic regions.

MT-1, MT-4, MT-5, MT-7, MT-8, MT-9, MT-10, MT-13 and MT-16 were found in annotated genes of grapevine nuclear genome and in homologous genes of mitochondrial and chloroplast genomes (except for MT-13, found in non-homologous genes) of different species, which adds additional evidence for

IGT process. Some TEs were identified in the exonic region of the gene they code for, potentially modifying the protein primary structure (Table 1). On the other hand, only one of these sequences (MT-13) was exclusively found in a gene intron. MT-13 (identified as a conserved domain; Supplementary Online Material 3) was identified as an intron of a gene coding for a methionyl-tRNA synthetase, which contains at least three conserved domains (C u s a c k et al., 1991). The location of these nine truncated TEs, their likely effect on protein structure and/or gene expression and their identification as conserved domains, highlight their possible role in gene structure and evolution. Interestingly, two sequences (MT-1 and MT-8) were observed in different regions of two copies of the same gene, suggesting a different role in both copies. As an example, MT-8 was found in one exon of the cytochrome oxidase gene in chromosome 12 (VIT_12s0028g02450) and as an exon and an intron in the cytochrome oxidase gene on chromosome 3 (VIT_03s0110g00190). The analysis of these sequences with the cytochrome oxidase gene from *Tetrastigma diepenhorstii* L. revealed high homology (above 90%), but only with the truncated gene sequence from mitochondrial genome. On the other hand, six truncated TEs (MT-02, MT-03, MT-06, MT-11, MT-12 and MT-14) were found in intergenic regions.

CONCLUSION

Most of the literature focuses on the analysis of grapevine nuclear TEs, while the studies on chloroplast and mitochondrial transposable elements are inextensive. Here, we identified 17 sequences in the grapevine (*V. vinifera*) mitochondrial genome as truncated (incomplete) MITEs, which were adequately grouped in four clusters according to their structure. Some of them were also found in the chloroplast and nuclear genomes, suggesting the existence of intracellular gene transfer. In addition, some of these sequences were identified as highly conserved domains in the coding and non-coding regions of different genes, suggesting their likely role in gene structure and/or expression. Some of these truncated MITEs were also identified in the genomes of species of diverse kingdoms (including some algae, archaea and bacteria), suggesting that these TE sequences might have appeared before the separation of life kingdoms. Our results indicate that sections from ancient transposons could be recruited to host genomes in early evolution stages and today we can observe them as highly conserved functional gene sequences.

Future works might aid in the understanding of (1) the mutual evolution of TEs and host genomes; (2) the intracellular (as well as horizontal) gene transfer, and (3) the prediction of possibly chimeric or/and fused genes. In this regard, grapevine is presented as an interesting model for this type of analysis.

REFERENCES

Adams KL, Song K, Roessler PG, Nugent JM, Doyle JL, Doyle JJ, Palmer JD (1999): Intracellular gene transfer in action: Dual transcription and multiple silencings of nuclear and mitochondrial cox2 genes in legumes. Proceedings of the National Academy of Sciences of the United States of America, 96, 13863–13868. doi: 10.1073/pnas.96.24.13863.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990): Basic local alignment search tool. Journal of Molecular Biology, 215, 403–410. doi: 10.1016/s0022-2836(05)80360-2.

Bao W, Kojima KK, Kohany O (2015): Repbase Update, a database of repetitive elements in eukaryotic genomes. Mobile DNA, 6, Article No. 11. doi: 10.1186/s13100-015-0041-9.

Bendich AJ (2013): DNA abandonment and the mechanisms of uniparental inheritance of mitochondria and chloroplasts. Chromosome Research, 21, 287–296. doi: 10.1007/s10577-013-9349-9.

Benjak A, Forneck A, Casacuberta JM (2008): Genome-wide analysis of the "cut-and-paste" transposons of grapevine. PLoS ONE, 3, e3107. doi: 10.1371/journal.pone.0003107.

Benjak A, Boue S, Forneck A, Casacuberta JM (2009): Recent amplification and impact of MITEs on the genome of grapevine (*Vitis vinifera* L.). Genome Biology and Evolution, 1, 75–84. doi: 10.1093/gbe/evp009.

Bennetzen JL (2005): Transposable elements, gene creation and genome rearrangement in flowering plants. Current Opinion in Genetics and Development, 15, 621–627. doi: 10.1016/j.gde.2005.09.010.

Bennetzen JL, Wang H (2014): The contributions of transposable elements to the structure, function, and evolution of plant genomes. Annual Review of Plant Biology, 65, 505–530. doi: 10.1146/annurev-arplant-050213-035811.

Bock R (2010): The give-and-take of DNA: Horizontal gene transfer in plants. Trends in Plant Science, 15, 11–22. doi: 10.1016/j.tplants.2009.10.001.

Brown JR (2003): Ancient horizontal gene transfer. Nature Reviews Genetics, 4, 121–132. doi: 10.1038/nrg1000.

Campbell AM (2000): Lateral gene transfer in prokaryotes. Theoretical Population Biology, 57, 71–77. doi: 10.1006/tpbi.2000.1454.

Carrier G, Le Cunff L, Dereeper A, Legrand D, Sabot F, Bouchez O, Audeguin L, Boursiquot J-M, This P (2012): Transposable elements are a major cause of somatic polymorphism in *Vitis vinifera* L. PLoS ONE, 7, e32973. doi: 10.1371/journal.pone.0032973.

Chalopin D, Naville M, Plard F, Galiana D, Volff J-N (2015): Comparative analysis of transposable elements highlights mobilome diversity and evolution in vertebrates. Genome Biology and Evolution, 7, 567–580. doi: 10.1093/gbe/evv005.

Chen J, Hu Q, Zhang Y, Lu C, Kuang H (2013): P-MITE: A database for plant miniature inverted-repeat transposable elements. Nucleic Acids Research, 42, D1176–D1181. doi: 10.1093/nar/gkt1000.

Chuong EB, Elde NC, Feschotte C (2016): Regulatory activities of transposable elements: From conflicts to benefits. Nature Reviews Genetics, 18, 71–86. doi: 10.1038/nrg.2016.139.

Cordaux R, Udit S, Batzer MA, Feschotte C (2006): Birth of a chimeric primate gene by capture of the transposase gene from a mobile element. Proceedings of the National Academy of Sciences of the United States of America, 103, 8101–8106. doi: 10.1073/pnas.0601161103.

Cui L (2006): ChloroplastDB: The Chloroplast Genome Database. Nucleic Acids Research, 34, D692–D696. doi: 10.1093/nar/gkj055.

Cusack S, Hartlein M, Leberman R (1991): Sequence, structural and evolutionary relationships between class 2 aminoacyl-tRNA synthetases. Nucleic Acids Research, 19, 3489–3498. doi: 10.1093/nar/19.13.3489.

Daboussi M-J, Capy P (2003): Transposable elements in filamentous fungi. Annual Review of Microbiology, 57, 275–299. doi: 10.1146/annurev.micro.57.030502.091029.

Denoeud F, Carretero-Paulet L, Dereeper A, Droc G, Guyot R, Pietrella M, Zheng C, Alberti A, Anthony F, Aprea G, Aury J-M, Bento P, Bernard M, Bocs S, Campa C, Cenci A, Combes M-C, Crouzillat D, Da Silva C, Daddiego L, De Bellis F, Dussert S, Garsmeur O, Gayraud T, Guignon V, Jahn K, Jamilloux V, Joet T, Labadie K, Lan T, Leclercq J, Lepelley M, Leroy T, Li L-T, Librado P, Lopez L, Munoz A, Noel B, Pallavicini A, Perrotta G, Poncet V, Pot D, Priyono, Rigoreau M, Rouard M, Rozas J, Tranchant-Dubreuil C, VanBuren R, Zhang Q, Andrade AC, Argout X, Bertrand B, de Kochko A, Graziosi G, Henry RJ, Jayarama, Ming R, Nagai C, Rounsley S, Sankoff D, Giuliano G, Albert VA, Wincker P, Lashermes P (2014): The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. Science, 345, 1181–1184. doi: 10.1126/science.1255274.

Di Genova A, Almeida A, Munoz-Espinoza C, Vizoso P, Travisany D, Moraga C, Pinto M, Hinrichsen P, Orellana A, Maass A (2014): Whole genome comparison between table and wine grapes reveals a comprehensive catalog of structural variants. BMC Plant Biology, 14, 7. doi: 10.1186/1471-2229-14-7.

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. doi: 10.1186/1471-2164-11-113.

Fan W-H, Woelfle MA, Mosig G (1995): Two copies of a DNA element,'Wendy', in the chloroplast chromosome of *Chlamydomonas reinhardtii* between rearranged gene clusters. Plant Molecular Biology, 29, 63–80. doi: 10.1007/bf00019119.

Fernandez L, Torregrosa L, Segura V, Bouquet A, Martinez-Zapater JM (2010): Transposon-induced gene activation as a mechanism generating cluster shape somatic variation in grapevine. The Plant Journal, 61, 545–557. doi: 10.1111/j.1365-313x.2009.04090.x.

Fernandez L, Chaib J, Martinez-Zapater J-M, Thomas MR, Torregrosa L (2013): Mis-expression of a PISTILLATA-like MADS box gene prevents fruit development in grapevine. The Plant Journal, 73, 918–928. doi: 10.1111/tpj.12083.

Fernandez L, Le Cunff L, Tello J, Lacombe T, Boursiquot JM, Fournier-Level A, Bravo G, Lalet S, Torregrosa L, This P, Martinez-Zapater JM (2014): Haplotype diversity of VvT-FL1A gene and association with cluster traits in grapevine (*V. vinifera*). BMC Plant Biology, 14, Article No. 209. doi: 10.1186/s12870-014-0209-3.

Feschotte C (2008): Transposable elements and the evolution of regulatory networks. Nature Reviews Genetics, 9, 397–405. doi: 10.1038/nrg2337.

Feschotte C, Jiang N, Wessler SR (2002): Plant transposable elements: Where genetics meets genomics. Nature Reviews Genetics, 3, 329–341. doi: 10.1038/nrg793.

Fortune PM, Roulin A, Panaud O (2008): Horizontal transfer of transposable elements in plants. Communicative and Integrative Biology, 1, 74–77. doi: 10.4161/cib.1.1.6328.

Giovannoni SJ (2005): Genome streamlining in a cosmopolitan oceanic bacterium. Science, 309, 1242–1245. doi: 10.1126/science.1114057.

Goremykin VV, Salamini F, Velasco R, Viola R (2008): Mitochondrial DNA of *Vitis vinifera* and the issue of rampant horizontal gene transfer. Molecular Biology and Evolution, 26, 99–110. doi: 10.1093/molbev/msn226.

Goremykin VV, Lockhart PJ, Viola R, Velasco R (2012): The mitochondrial genome of *Malus domestica* and the import-driven hypothesis of mitochondrial genome expansion in seed plants. The Plant Journal, 71, 615–626. doi: 10.1111/j.1365-313x.2012.05014.x.

Goujon M, McWilliam H, Li W, Weizhong L, Valentin F, Squizzato S, Paern J, Lopez R (2010): A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Research, 38, W695–W699. doi: 10.1093/nar/gkq313.

Grimplet J, Van Hemert J, Carbonell-Bejerano P, Diaz-Riquelme J, Dickerson J, Fennell A, Pezzotti M, Martinez-Zapater JM (2012): Comparative analysis of grapevine whole-genome gene predictions, functional annotation, categorization and integration of the predicted gene sequences. BMC Research Notes, 5, 213. doi: 10.1186/1756-0500-5-213.

Guermonprez H, Henaff E, Cifuentes M, Casacuberta JM (2012): MITEs, miniature elements with a major role in plant genome evolution. In: Grandbastien MA, Casacuberta J (eds): Plant transposable elements. Topics in Current Genetics, 24, 113–124. Springer, Berlin, Heidelberg.

Hao W, Golding GB (2009): Does gene translocation accelerate the evolution of laterally transferred genes? Genetics, 182, 1365–1375. doi: 10.1534/genetics.109.104216.

Huang CY (2005): Mutational decay and age of chloroplast and mitochondrial genomes transferred recently to angiosperm

nuclear chromosomes. Plant Physiology, 138, 1723–1733. doi: 10.1104/pp.105.060327.

International Rice Genome Sequencing Project (2005): The map-based sequence of the rice genome. Nature, 436, 793–800. doi: 10.1038/nature03895.

Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Wincker P (2007): The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature, 449, 463–467. doi: 10.1038/nature06148.

Jansen RK, Kaittanis C, Saski C, Lee SB, Tomkins J, Alverson AJ, Daniell H (2006): Phylogenetic analyses of *Vitis* (Vitaceae) based on complete chloroplast genome sequences: Effects of taxon sampling and phylogenetic methods on resolving relationships among rosids. BMC Evolutionary Biology, 6, Article No. 32. doi: 10.1186/1471-2148-6-32.

Jiang N, Bao Z, Zhang X, Eddy SR, Wessler, SR (2004): Pack-MULE transposable elements mediate gene evolution in plants. Nature, 431, 569. doi: 10.1038/nature02953

Kapitonov VV, Jurka J (2004): Harbinger transposons and an ancient HARBI1 gene derived from a transposase. DNA and Cell Biology, 23, 311–324. doi: 10.1089/104454904323090949.

Kapitonov VV, Jurka J (2008): A universal classification of eukaryotic transposable elements implemented in Repbase. Nature Reviews Genetics, 9, 411–412. doi: 10.1038/nrg2165-c1.

Keeling PJ, Palmer JD (2008): Horizontal gene transfer in eukaryotic evolution. Nature Reviews Genetics, 9, 605–618. doi: 10.1038/nrg2386.

Kersey PJ, Allen JE, Armean I, Boddu S, Bolt BJ, Carvalho-Silva D, Christensen M, Davis P, Falin LJ, Grabmueller C, Humphrey J, Kerhornou A, Khobova J, Aranganathan NK, Langridge N, Lowy E, McDowall MD, Maheswari U, Nuhn M, Ong CK, Overduin B, Paulini M, Pedro H, Perry E, Spudich G, Tapanari E, Walts B, Williams G, Tello-Ruiz M, Stein J, Wei S, Ware D, Bolser DM, Howe KL, Kulesha E, Lawson D, Maslen G, Staines DM (2016): Ensembl Genomes 2016: More genomes, more complexity. Nucleic Acids Research, 44, D574–D580. doi: 10.1093/nar/gkv1209.

Kleckner N (1981): Transposable elements in prokaryotes. Annual Review of Genetics, 15, 341–404. doi: 10.1146/annurev.ge.15.120181.002013.

Knoop V, Unseld M, Marienfeld J, Brandt P, Sunkel S, Ullrich H, Brennicke A (1996): Copia-, gypsy- and line-like retrotransposon fragments in the mitochondrial genome of *Arabidopsis thaliana*. Genetics, 142, 579–585.

Kobayashi S (2004): Retrotransposon-induced mutations in grape skin color. Science, 304, 982–982. doi: 10.1126/science.1095011.

Kolodner R, Tewari KK (1979): Inverted repeats in chloroplast DNA from higher plants. Proceedings of the National Academy of Sciences of the United States of America, 76, 41–45. doi: 10.1073/pnas.76.1.41.

Koonin EV, Makarova KS, Aravind L (2001): Horizontal gene transfer in prokaryotes: Quantification and classification. Annual Review of Microbiology, 55, 709–742. doi: 10.1146/annurev.micro.55.1.709.

Kubo T, Mikami T (2007): Organization and variation of angiosperm mitochondrial genome. Physiologia Plantarum, 129, 6–13. doi: 10.1111/j.1399-3054.2006.00768.x.

Kubo T, Nishizawa S, Sugawara A, Itchoda N, Estiati A, Mikami T (2000): The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for tRNACys (GCA). Nucleic Acids Research, 28, 2571–2576.

Kumar S, Stecher G, Tamura K (2016): MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. Molecular Biology and Evolution, 33, 1870–1874. doi: 10.1093/molbev/msw054.

Majorek KA, Dunin-Horkawicz S, Steczkiewicz K, Muszewska A, Nowotny M, Ginalski K, Bujnicki JM (2014): The RNase H-like superfamily: New members, comparative structural analysis and evolutionary classification. Nucleic Acids Research, 42, 4160–4179. doi: 10.1093/nar/gkt1414.

Marchler-Bauer A, Bryant SH (2004): CD-Search: protein domain annotations on the fly. Nucleic Acids Research, 32, W327–W331. doi: 10.1093/nar/gkh454.

McClintock B (1951): Chromosome organization and genic expression. Cold Spring Harbor Laboratory Press, 16, 13–47.

Milligan BG, Hampton JN, Palmer JD (1989): Dispersed repeats and structural reorganization in subclover chloroplast DNA. Molecular Biology and Evolution, 6, 355–368. doi: 10.1093/oxfordjournals.molbev.a040558.

Moisy C, Garrison K, Meredith CP, Pelsy F (2008): Characterization of ten novel Ty1 copia-like retrotransposon families of the grapevine genome. BMC Genomics, 9, 469. doi: 10.1186/1471-2164-9-469.

Morgante M, De Paoli E, Radovic S (2007): Transposable elements and the plant pan-genomes. Current opinion in plant biology, 10, 149-155. doi: 10.1016/j.pbi.2007.02.001

Notsu Y, Masood S, Nishikawa T, Kubo N, Akiduki G, Nakazono M, Hirai A, Kadowaki K (2002): The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: Frequent DNA sequence acquisition and loss during the evolution of flowering plants. Molecular Genetics and Genomics, 268, 434–445. doi: 10.1007/s00438-002-0767-1.

Nugent JM, Palmer JD (1991): RNA-mediated transfer of the gene coxII from the mitochondrion to the nucleus during flowering plant evolution. Cell, 66, 473–481. doi: 10.1016/0092-8674(81)90011-8.

Ohyama K, Yamano Y, Fukuzawa H, Komano T, Yamagishi H, Fujimoto S, Sugiura M (1983): Physical mappings of chloroplast DNA from liverwort *Marchantia polymorpha* L. cell suspension cultures. MGG Molecular and General Genetics, 189, 1–9. doi: 10.1007/bf00326047.

Pinosio S, Giacomello S, Faivre-Rampant P, Taylor G, Jorge V, Le Paslier MC, Zaina G, Bastien C, Cattonaro F, Mar-

roni F, Morgante M (2016): Characterization of the poplar pan-genome by genome-wide identification of structural variation. Molecular Biology and Evolution, 33, 2706–2719. doi: 10.1093/molbev/msw161.

Prina AR (1996): Mutator-induced cytoplasmic mutants in barley: Genetic evidence of activation of a putative chloroplast transposon. Journal of Heredity, 87, 385–389. doi: 10.1093/oxfordjournals.jhered.a023020.

Reboud X, Zeyl C (1994): Organelle inheritance in plants. Heredity, 72, 132–140. doi: 10.1038/hdy.1994.19.

Rice PA, Baker TA (2001): Comparative architecture of transposase and integrase complexes. Nature Structural and Molecular Biology, 8, 302–307. doi: 10.1038/86166.

Robertson HM (2002): Evolution of DNA transposons in eukaryotes. In: Craig NL, Craigie R, Gellert M, Lambowitz AM (eds): Mobile DNA II. ASM Press, Washington, DC, 1093–1110. doi: 10.1128/9781555817954.

Rousseau-Gueutin M, Huang X, Higginson E, Ayliffe M, Day A, Timmis JN (2013): Potential functional replacement of the plastidic acetyl-CoA carboxylase subunit (accD) gene by recent transfers to the nucleus in some angiosperm lineages. Plant Physiology, 161, 1918–1929. doi: 10.1104/pp.113.214528.

Rubin E, Lithwick G, Levy AA (2001): Structure and evolution of the hAT transposon superfamily. Genetics, 158, 949–957. doi: 10.1038/srep27101.

Santos BZ, Brazil P, Vens C, Cerri R (2016): Decision trees for hierarchical classification of transposable elements. In: Proc. 25th Belgian-Dutch Machine Learning Conference (Benelearn), Kortrijk, Belgium, 3 pp..

Sotero-Caio CG, Platt RN II, Suh A, Ray DA (2017): Evolution and diversity of transposable elements in vertebrate genomes. Genome Biology and Evolution, 9, 161–177. doi: 10.1093/gbe/evw264.

Sugita M, Kato A, Shimada H, Sugiura M (1984): Sequence analysis of the junctions between a large inverted repeat and single-copy regions in tobacco chloroplast DNA. MGG Molecular Genetics and Genomics, 194, 200–205. doi: 10.1007/bf00383517.

The Arabidopsis Genome Initiative (2000): Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature, 408, 796–815. doi: 10.1038/35048692.

Thompson JA, Hansmann P, Knoth R, Link G, Falk H (1981): Electron microscopical localisation of the 23S and 16S rRNA genes within an inverted repeat for two chromoplast DNAs. Current Genetics, 4, 25–28. doi: 10.1007/bf00376782.

Velasco R, Zharkikh A, Troggio M, Cartwright DA, Cestaro A, Pruss D, Pindo M, Fitzgerald LM, Vezzulli S, Reid J, Malacarne G, Iliev D, Coppola G, Wardell B, Micheletti D, Macalma T, Facci M, Mitchell JT, Perazzolli M, Eldredge G, Gatto P, Oyzerski R, Moretto M, Gutin N, Stefanini M, Chen Y, Segala C, Davenport C, Dematte L, Mraz A, Battilana J, Stormo K, Costa F, Tao Q, Si-Ammour A, Harkins T, Lackey A, Perbost C, Taillon B, Stella A, Solovyev V, Fawcett JA, Sterck L, Vandepoele K, Grando SM, Toppo S, Moser C, Lanchbury J, Bogden R, Skolnick M, Sgaramella V, Bhatnagar SK, Fontana P, Gutin A, Van de Peer Y, Salamini F, Viola R (2007): A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. PLoS ONE, 2, e1326. doi: 10.1371/journal.pone.0001326.

Vitte C, Fustier M-A, Alix K, Tenaillon MI (2014): The bright side of transposons in crop evolution. Briefings in Functional Genomics, 13, 276–295. doi: 10.1093/bfgp/elu002.

Wang N, Xiang Y, Fang L, Wang Y, Xin H, Li S (2013): Patterns of gene duplication and their contribution to expansion of gene families in grapevine. Plant Molecular Biology Reporter, 31, 852–861. doi: 10.1007/s11105-013-0556-5.

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. Nature Reviews Genetics, 8, 973–982. doi: 10.1038/nrg2165.

Yuan Y-W, Wessler SR (2011): The catalytic domain of all eukaryotic cut-and-paste transposase superfamilies. Proceedings of the National Academy of Sciences of the United States of America, 108, 7884–7889. doi: 10.1073/pnas.1104208108.

Zabala G, Vodkin LO (2005): The wp mutation of Glycine max carries a gene-fragment-rich transposon of the CACTA superfamily. The Plant Cell, 17, 2619-2632. doi: 10.1105/tpc.105.033506

Zhaxybayeva O (2006): Phylogenetic analyses of cyanobacterial genomes: Quantification of horizontal gene transfer events. Genome Research, 16, 1099–1108. doi: 10.1101/gr.5322306.

Zhou DX, Massenet O, Quigley F, Marion MJ, Moneger F, Huber P, Mache R (1988): Characterization of a large inversion in the spinach chloroplast genome relative to *Marchantia*: A possible transposon-mediated origin. Current Genetics, 13, 433–439. doi: 10.1007/bf00365665.

Corresponding Author:

Alexander M i l o v a n o v, Dr., Kuban State Agrarian University, Faculty of Horticulture and Viticulture, Department of Viticulture, Kalinina 13, 350044, Krasnodar, Russian Federation, phone: +79615084523, email: alexander.milovanov1991@gmail.com