The role of LTR retrotransposons in plant genetic engineering : How to control their transposition in the genome

Ramakrishnan, Muthusamy

2023-01-15

Ramakrishnan , M , Papolu , P K , Mullasseri , S , Zhou , M , Sharma , A , Ahmad , Z , Satheesh , V , Kalendar , R & Wei , Q 2023 , ' The role of LTR retrotransposons in plant genetic engineering : How to control their transposition in the genome ' , Plant Cell Reports , þÿvol. 42 , no. 1 , pp. 3 15 . https://doi.org/10.1007/s00299-022-02945-z

http://hdl.handle.net/10138/355872
https://doi.org/10.1007/s00299-022-02945-z

Plant Cell Reports
https://doi.org/10.1007/s00299-022-02945-z

REVIEW

The role of LTR retrotransposons in plant genetic engineering: how to control their transposition in the genome

Muthusamy Ramakrishnan[1] · Pradeep K. Papolu[2] · Sileesh Mullasseri[3] · Mingbing Zhou[2,4] · Anket Sharma[2,5] · Zishan Ahmad[1] · Viswanathan Satheesh[6] · Ruslan Kalendar[7,8] · Qiang Wei[1]

Received: 27 May 2022 / Accepted: 23 October 2022

Abstract

Key message **We briefly discuss that the similarity of LTR retrotransposons to retroviruses is a great opportunity for the development of a genetic engineering tool that exploits intragenic elements in the plant genome for plant genetic improvement.**

Abstract Long terminal repeat (LTR) retrotransposons are very similar to retroviruses but do not have the property of being infectious. While spreading between its host cells, a retrovirus inserts a DNA copy of its genome into the cells. The ability of retroviruses to cause infection with genome integration allows genes to be delivered to cells and tissues. Retrovirus vectors are, however, only specific to animals and insects, and, thus, are not relevant to plant genetic engineering. However, the similarity of LTR retrotransposons to retroviruses is an opportunity to explore the former as a tool for genetic engineering. Although recent long-read sequencing technologies have advanced the knowledge about transposable elements (TEs), the integration of TEs is still unable either to control them or to direct them to specific genomic locations. The use of existing intragenic elements to achieve the desired genome composition is better than using artificial constructs like vectors, but it is not yet clear how to control the process. Moreover, most LTR retrotransposons are inactive and unable to produce complete proteins. They are also highly mutable. In addition, it is impossible to find a full active copy of a LTR retrotransposon out of thousands of its own copies. Theoretically, if these elements were directly controlled and turned on or off using certain epigenetic mechanisms (inducing by stress or infection), LTR retrotransposons could be a great opportunity to develop a genetic engineering tool using intragenic elements in the plant genome. In this review, the recent developments in uncovering the nature of LTR retrotransposons and the possibility of using these intragenic elements as a tool for plant genetic engineering are briefly discussed.

Keywords Transposable elements · Retrotransposons · Plants · Genetic engineering · Retroviruses · Targeted integration

Communicated by Wusheng Liu.

✉ Ruslan Kalendar
ruslan.kalendar@helsinki.fi

✉ Qiang Wei
weiqiang@njfu.edu.cn

[1] Co-Innovation Center for Sustainable Forestry in Southern China, Bamboo Research Institute, Key Laboratory of National Forestry and Grassland Administration on Subtropical Forest Biodiversity Conservation, College of Biology and the Environment, Nanjing Forestry University, Nanjing 210037, Jiangsu, China

[2] State Key Laboratory of Subtropical Silviculture, Institute of Bamboo Research, Zhejiang A&F University, Lin'an, Hangzhou 311300, Zhejiang, China

[3] Department of Zoology, St. Albert's College (Autonomous), Kochi, 682018, Kerala, India

[4] Zhejiang Provincial Collaborative Innovation Center for Bamboo Resources and High-Efficiency Utilization, Zhejiang A&F University, Lin'an, Hangzhou 311300, Zhejiang, China

[5] Department of Plant Science and Landscape Architecture, University of Maryland, College Park, USA

[6] Shanghai Center for Plant Stress Biology, CAS Center for Excellence in Molecular Plant Sciences, Chinese Academy of Sciences, Shanghai 200032, China

[7] Helsinki Institute of Life Science HiLIFE, University of Helsinki, Biocenter 3, Viikinkaari 1, F1-00014 Helsinki, Finland

[8] Institute of Plant Biology and Biotechnology (IPBB), Timiryazev Street 45, 050040 Almaty, Kazakhstan

Published online: 19 November 2022

Introduction: LTR retrotransposons

Transposable elements (TEs) are mobile genetic elements which represent a significant portion of eukaryotic genomes. Based on their mechanism of transposition, TEs are classified into DNA transposons (Class II) and retrotransposons (Class I). Retrotransposons are divided into long terminal repeat retrotransposons (LTR retrotransposons), non-LTR retrotransposons, and DIRS (Dictyostelium intermediate repeat sequence) (Bourque et al. 2018). LTR retrotransposons, the most abundant group of TEs in the plant genome, use a "copy-and-paste" mechanism via an RNA intermediate for their transposition (Fig. 1). LTR retrotransposons are generally classified into three superfamilies, Ty1/*copia*, Ty3/*Gypsy*, and endogenous retroviruses (ERVs). Ty1/*Copia* and Ty3/*Gypsy* are dispersed across the genomes of higher plants. Members of these superfamilies are capable of increasing in copy numbers, and are often activated by

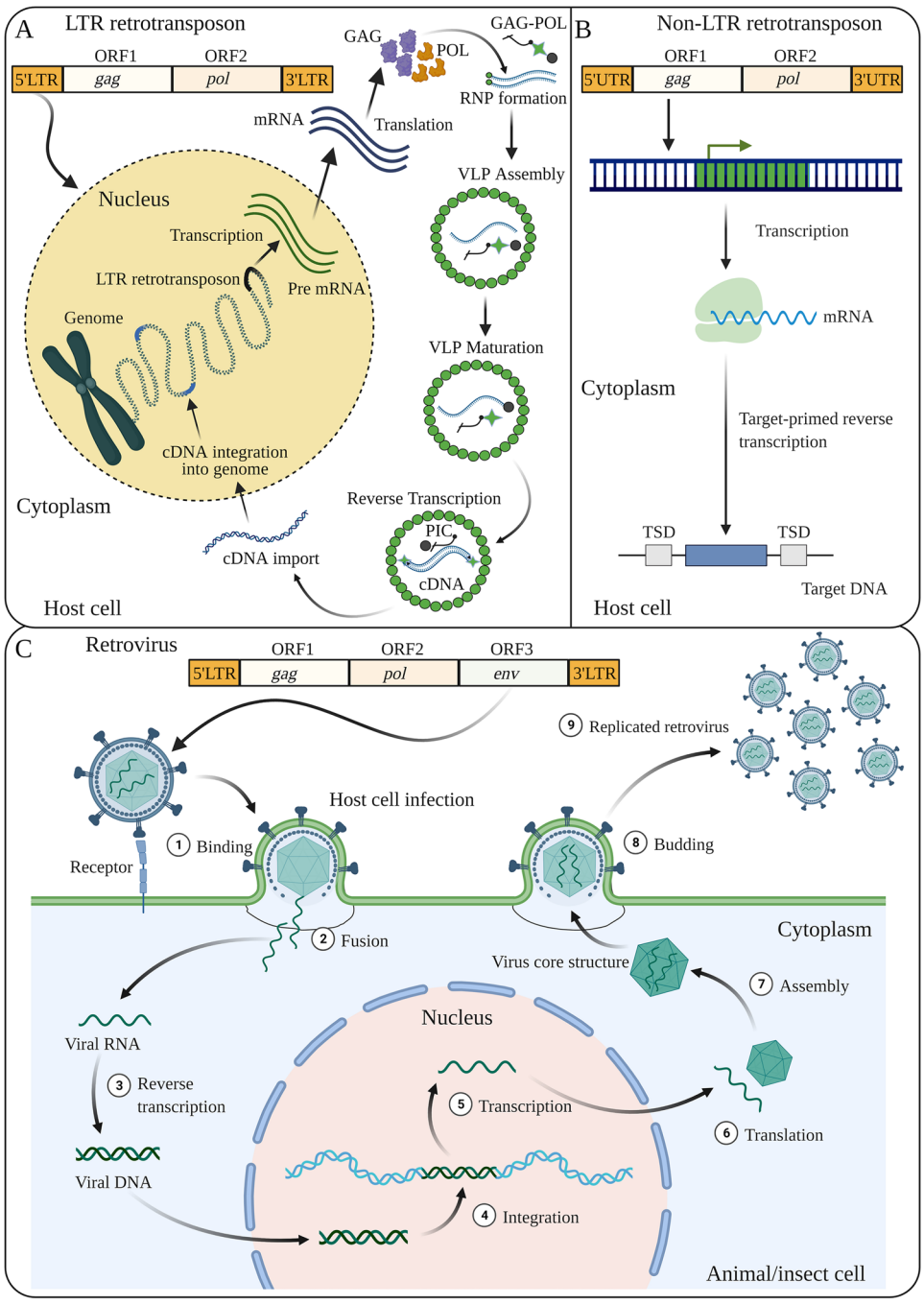

Fig. 1 The mechanism of transposition of long terminal repeat (*LTR*) retrotransposons, *non-LTR* retrotransposons, and retroviruses. **A** After the transcription of a LTR retrotransposon, the mRNA encodes GAG and POL proteins to produce virus-like particles (*VLPs*) with reverse transcriptase, which synthesises the cDNA; then, the cDNA is imported into the nucleus and integrated into the genome, which is called replicate retrotransposition; *RNP* and *PIC* the ribonucleoprotein particle and pre-integration complex, respectively. **B** Non-LTR retrotransposons use target site-primed reverse transcription and usually terminate in a poly(A) sequence; *TSD* target site duplication. **C** The transposition of retroviruses and LTR retrotransposons is relatively similar, but retroviruses have an *envelope* (*env*) gene that infects animal/insect cells; also, retroviruses have an additional open reading frame (*ORF*) in their genome. Created with BioRender.com

various biotic and abiotic stresses due to retrotransposition bursts (Havecker et al. 2004).

Retrotransposons outnumber genes in plant genomes, comprising the bulk of the genome, and they are largely inactive during development (Alzohairy et al. 2014). However, they replicate through cycles of transcription, reverse transcription, and integrate new copies, without deleting original copies from the genome during replication (Quesneville 2020; Ramakrishnan et al. 2022). The replication of LTR retrotransposons is equivalent to the intracellular phase of the retroviral life cycle (Li et al. 2022). LTR retrotransposons cause easily detectable genetic changes in the genome (Bourque et al. 2018). The replication of retroviruses and retrotransposons depends on the selection of a favorable chromosomal site for the integration of their genomic DNA (Sultana et al. 2017). Therefore, LTR retrotransposon insertions and their mechanisms of targeted integration could have significant applications in genome engineering. This review provides a brief account of the current understanding of these elements and their roles in crop plants, and explores how LTR retrotransposons can be used as genetic engineering tools for plant breeding and agriculture.

LTR retrotransposons and retroviruses

LTR retrotransposons and retroviruses are somewhat similar (Li et al. 2022). Like retroviruses, LTR retrotransposons replicate through a cycle of transcription of integrated copies as if they were cellular genes, followed by translation of their encoded products and reverse transcription of RNA into cDNA (Fig. 1). These proteins are present in two main open reading frames (ORFs) that specify GAG, the structural protein forming the nucleocapsid, and the POL polyprotein, which is processed by its own aspartic proteinase (AP) domain. The ORF also contains reverse transcriptase (RT) and RNAse H (RH) to perform reverse transcription, and an integrase (IN) to insert the new copy into the genome (Quesneville 2020; Ramakrishnan et al. 2022).

However, retroviruses have an *envelope* (*env*) gene that is used to infect animal/insect cells (Sultana et al. 2017) (Fig. 1). Further, retroelements with an extra ORF in the same position as the *env* gene have been found in retrovirus genomes (Leblanc et al. 2000; Pelisson et al. 2002). In host cells, 5′ LTR is known to control the expression of retrovirus genes responsible for producing infectious particles. In certain retroviruses, 3′ LTR is oriented in the inverse direction from that of the transcription controlled by the 5′ LTR (Barbeau and Mesnard 2011).

In contrast to retroviruses, LTR retrotransposons, with a few exceptions (Ty3/*Gypsy* superfamily), do not contain the *env* gene encoding the envelope protein necessary for retrovirus integration (Quesneville 2020; Vicient and Casacuberta 2020). Because the envelope protein is absent in most plant LTR retrotransposons, the infection does not occur as it does with retroviruses. The retrotransposon copies do not leave the host cell. Instead, they migrate out of the nucleus and integrate the newly synthesized cDNA into another locus of the same genome (Havecker et al. 2004), resulting in the accumulation of multiple copies of a particular retrotransposon. In contrast, for retroviruses, a complete cycle consists of the retroviruses infecting a cell and migrating from that cell to the next (Havecker et al. 2004; Sultana et al. 2017). Hence, the copy number of retroviruses is not high. In addition to the envelope protein, animal cells are characterized by the presence of appropriate membrane proteins to which the envelope protein binds during retrovirus integration (Grandi and Tramontano 2018). Therefore, the presence of membrane proteins on the animal cell wall, and the presence of an envelope protein on the surface of the virus, allow retroviruses to easily integrate into animal cells.

Having the property of being infectious, retroviruses have several advantages as vectors for gene delivery, such as receptor-mediated uptake of a membrane-coated viral particle into target cells, reverse transcription of a plus-stranded RNA genome into double-stranded DNA, and cytoplasmic assembly of particles with the full-length retroviral mRNA as the mobile form of genetic information (Baum et al. 2006). Despite their infectious properties, retroviral sequences have been chosen to produce beneficial immune functions through immune epigenetic regulation in mammals (Buttler and Chuong 2022). Although retroviral vector-mediated gene transfer systems have been a good choice for animals, this method is not suitable for plants, as plants lack the appropriate membrane/receptor proteins needed to bind the retroviral envelope protein (Grandi and Tramontano 2018).

In plants, activating or silencing LTR retrotransposons could produce favorable epigenetic modifications. For epigenetic modification, however, using LTR retrotransposons, as candidates, presents many advantages (such as genome stability, gene imprinting, introduction of new gene functions, genetic variability, stress tolerance, etc.) and disadvantages (such as mutagenic effects, genetic rearrangements, genomic stress, loss of gene function, high copy numbers, etc.) (Ramakrishnan et al. 2021; Zhang et al. 2018). For example, hypomethylation of retrotransposon, related to rice *Karma*, reduced the yield in African oil palm (*Elaeis guineensis*), while hypermethylation of the retrotransposon, near the *Karma* splice improved the normal fruit set (Ong-Abdullah et al. 2015). It is up to future research on the artificial activation of LTR retrotransposons, and the mechanisms involved in their mobility and silencing, to provide a better understanding of their involvement in plant genome evolution and genetic diversity.

Role of LTR retrotransposons in crop plants

Some LTR retrotransposons associated with molecular functions have been identified in crop plants, and can be used as tools in plant genetic engineering (reviewed by Galindo-Gonzalez et al. 2017; Orozco-Arias et al. 2019). More candidate LTR retrotransposon tools for genetic engineering will be identified over time with advances in technologies such as high-throughput long-read sequencing. This review focuses on LTR retrotransposons, mostly characterized through tissue culture approaches (Table 1), which can be used as genetic engineering tools. Tobacco *Tnt1* LTR sequences have been characterized in several higher plants by tissue culture approaches. The transposition of *Tnt1* in tobacco mesophyll protoplasts showed its potential as a genetic engineering tool in plants (Grandbastien et al. 1989). Moreover, two *Tnt1* elements in transgenic tobacco were expressed in leaf-derived protoplast, but not in leaf tissues, indicating that the transcription features of *Tnt1* could provide a molecular basis for somaclonal variation and tissue culture-induced mutagenesis (Pouteau et al. 1991). Moreover, fungal extracts can efficiently activate *Tnt1* transposition and increase the number of new copies of *Tnt1* with high sequence similarities to subpopulations; therefore, *Tnt1* transposition might play a significant role in activating the host's genetic plasticity in response to environmental stress (Melayah et al. 2001). For instance, in *Medicago truncatula*, a *Tnt1* element was activated during protoplast culture, generating the highest copy number insertions per plant, and the copy numbers were stable during the life cycle (d'Erfurth et al. 2003). Therefore, *Tnt1* can be used as a powerful genetic engineering tool in leguminous plants. In addition, the expression of the *Tnt1* promoter in heterologous species of transgenic tomato and *Arabidopsis* are capable of inducing a plant defense response, and $CuCl_2$ and salicylic acid treatment can be used to drive transgenes that confer resistance to plant parasites and pathogens (Mhiri et al. 1997).

The copy numbers of *Tos17*, an endogenous Ty1/*copia*-like LTR retrotransposon, in transgenic rice have increased with prolonged tissue culture duration, and tissue culture-induced mutations of *Tos17* could be an advantage for an insertional mutagenesis system, and for the functional analysis of genes (Hirochika et al. 1996). Furthermore, the sequence analysis of *Tos17* insertion mutant lines through tissue culture showed that *Tos17* prefers to integrate into genic rather than intergenic regions (Miyao et al. 2003). This indicates that the utility of *Tos17* insertions could lie in the rapidly evolving gene classes. On the other hand, the targeted mutagenesis of *Tos17*, using the genome editing tool, clustered regularly interspaced short palindromic repeats (CRISPR)-associated protein 9 (Cas9) (CRISPR/Cas9), in rice revealed that regenerated plants derived from callus culture and homozygous plants show a lack of *Tos17* in the next generation (Saika et al. 2019), which might be a useful tool to elucidate the functional role of TE transposition in genome evolution. Similarly, the activity of mutated *Tos17* on chromosomes 7 (*Tos17* $^{Chr.7}$) in rice through tissue culture showed 873-bp DNA deletion in the coding region of the *pol* gene by CRISPR/Cas9-mediated gene editing with single guide RNAs (sgRNAs), and, further, the generation of *Tos17 D873* indicates that *Tos17* requires *gag*, *integrase*, and *pol* domains for its transposition (Luo et al. 2020). This demonstrates that the generation of the *Tos17 D873* allele might be useful in the establishment of transgenic rice plants for gene function and epigenetics. Moreover, *Tos17* $^{Chr.7}$ has been extensively used for insertional mutagenesis as a tool for the functional analysis of rice genes.

The *Tto1* element in tobacco is an active plant retrotransposon (Hirochika 1993), and the Tto1 promoter in tobacco is responsible for a significant level of expression in transgenic plants, demonstrating that a 13-bp *cis*-regulatory element is involved in response to tissue culture, wounding fungal elicitors, and methyl-jasmonate (Takeda et al. 1999). Moreover, *Tto1* could be induced in tobacco leaves by wounding and methyl-jasmonate stimuli. The decrease in DNA methylation and gene silencing machinery in the suppression of *Tto1* and *Tar17* induced by tissue culture has been shown to become hypomethylated, and transcriptionally active in *ddm1* (for decrease in DNA methylation) mutants, suggesting that gene silencing and DNA methylation are effective strategies in LTR retrotransposon suppression (Hirochika et al. 2000). This also provides an opportunity to control retrotransposons for gene silencing, and can be used as a tool for genetic engineering in plants.

In chickpea, the *CARE1* LTR retrotransposon, from the *Gypsy* superfamily, showed that 5′ LTR was inactive in a heterologous plant under normal and tissue culture conditions, indicating that *CARE1 cis*-elements might be hindered in the recruitment of transcription factors to the promoter (Rajput and Upadhyaya 2009). The promoter region of the *h*AT superfamily LTR retrotransposon in *Saintpaulia* species showed that tissue culture-derived progeny elicit retrotransposon excision, which, in turn, alters the expression levels of *flavonoid, 3′, 5′-hydroxylase* (*F3′F5′H*) and flower color in *Saintpaulia* (Sato et al. 2011). Similarly, when a transgene driven by the promoter of the *Tnt1* element is stably integrated into tobacco, the transgene is silenced and its DNA methylation is increased. However, endogenous *Tnt1* elements remain partially methylated and incorporate histone variants upon induction, suggesting that the *Tnt1* promoter is the target of transcriptional gene silencing in tobacco (Hernández-Pinzón et al. 2012). The promoter-GUS fusion of the LTR retrotransposon, *ATCOPIA93*, in

Table 1 LTR retrotransposons used in plant genetic engineering and the outcomes

LTRs	Species	Superfamilies	Mode of plant transformation	Results	References
Tnt1	Tobacco	Ty1/*Copia*	Cell cultures derived from tobacco mesophyll protoplasts	A complete mobile retrotransposon characterized in higher plants	Grandbastien et al. (1989)
Tnt1	Tobacco	Ty1/*Copia*	Leaf-derived protoplasts	The first example of tissue culture-induced mutagenesis in plants and molecular basis for some somaclonal variation events	Pouteau et al. (1991)
Tto1–Tto2	Tobacco	Tyl/*Copia*	Leaf and callus-derived protoplast culture	Retrotransposons of tobacco are activated during tissue culture	Hirochika (1993)
Tnt1	Tobacco	Tyl/*Copia*	Leaf protoplast culture	Interacts with protoplast-specific nuclear factors	Casacuberta and Grandbastien (1993)
ΔNaeAc	Flax	Ty1/*Copia*	*Agrobacterium*-mediated regeneration method	Showed enhanced transcription and transposition in transgenic plants	Finnegan et al. (1993)
Tos6–Tos20	Rice	Ty1/*Copia*	Protoplast culture	Tissue culture-induced activation of *Tos17* may be a useful tool for insertional mutagenesis and functional analysis of genes	Hirochika et al. (1996)
ATGP3, ATCOPIA13, ATCOPIA21, ATCOPIA93	*Arabidopsis*	Ty3-*Gypsy* and Ty1/*Copia*	Parental inter-strain cross	Structuring the genome through molecular genetic approaches	Kakutani et al. (1996)
BARE-1	Barley	Ty1/*Copia*	Leaf protoplast	*BARE-1* may retain the potential for propagation in the barley genome	Suoniemi et al. (1996)
Tto1	Tobacco	Ty1/*Copia*	*Agrobacterium*-mediated regeneration method	The transcription of *Tto1* is activated by mechanical wounding	Takeda et al. (1998)
Tto1	Tobacco	Ty1/*Copia*	*Agrobacterium*-mediated regeneration method	Functions as a *cis*-regulatory element, which confers responsiveness to tissue culture, wounding, methyl jasmonate, and fungal elicitors	Takeda et al. (1999)
Tto1	*Arabidopsis*	Ty1/*Copia*	*Agrobacterium*-mediated regeneration method	The inactivation of retrotransposons and the silencing of repeated genes have mechanisms in common	Hirochika et al. (2000)
Tnt1	*Arabidopsis*	Ty1/*Copia*	Co-cultivation of root explants with *Agrobacterium*	The potential of *Tnt1* as a tool for gene tagging	Courtial et al. (2001)
Tnt1	Tobacco	Ty1/*Copia*	Callus-derived shoot regeneration protoplast	Plays a crucial role in generating host genetic plasticity in response to environmental stresses	Melayah et al. (2001)
Dasheng	Rice	Ty3/*Gypsy*	Doubled haploid mapping population	The highest-copy-number LTR elements and one of the most recent elements to be amplified in the rice genome	Jiang et al. (2002)
Tnt1	*Medicago truncatula*	Ty1/*Copia*	Functioning of an efficient transposable element in leguminous plants	*Agrobacterium*-mediated transformation of leaf explants	d'Erfurth et al. (2003)

Table 1 (continued)

LTRs	Species	Superfamilies	Mode of plant transformation	Results	References
Tos17	Rice	Ty1/*Copia*	Callus-derived protoplast culture	*Tos17* insertions are distributed throughout the rice genome, and *Tos17* prefers genic regions for integration	Miyao et al. (2003)
TLC1	Tobacco	Ty1/*Copia*	Leaf protoplasts	Ethylene-dependent signaling is the main signaling pathway involved in the regulation of the expression of the *TLC1.1* element from *Lycopersicon chilense*	Tapia et al. (2005)
LORE1	*Lotus japonicus*	Ty3/*Gypsy*	*Agrobacterium*-mediated regeneration method	A simple insertion mutagenesis based on endogenous *LORE1* elements can be established for *Lotus*	Madsen et al. (2005)
MULEs	Foxtail millet	Tyl/*Copia*	Horizontal gene transfer	Example of horizontal transfer of nuclear-encoded genes between higher plants	Diao et al. (2005)
Ogre	*M. truncatula*	Ty3/Gypsy	*Agrobacterium*-mediated regeneration method	Possess functional introns, non-coding regions that are spliced out from the element transcripts	Steinbauerová et al. (2008)
RIRE1	Rice	Tyl/*Copia*	Horizontal gene transfer	Constitutes a new case of horizontal transfer in plants	Roulin et al. (2008)
Lullaby	Rice	Ty1/*Copia*	Seed-embryo callus transformation	An interesting candidate for a *cis* element that could account for the transcriptional activation in calli	Picault et al. (2009)
Rider	Tomato	Tyl/*Copia*	Horizontal gene transfer	Active retrotransposon in the tomato genome	Cheng et al. (2009)
CARE1	Tobacco	Ty3/*Gypsy*	*Agrobacterium*-mediated regeneration method	*CARE1* from chickpea and the presence of *CARE1* family representatives in various chickpea accessions	Rajput and Upadhyaya (2009)
LORE1	*Lotus japonicus*	Ty3/*Gypsy*	*Agrobacterium*-mediated regeneration method	*LORE1* indicates that this chromovirus has considerable potential for generating genetic and epigenetic diversity in the host plant population	Fukai et al. (2010)
F3'5'H sequences	Saintpaulia	Tyl/*Copia*	Explant tissue regeneration method	Retrotransposons are underlying factors responsible for somaclonal variations under in vitro conditions	Sato et al. (2011)
Tal1	*Arabidopsis*	Tyl/*Copia*	*Agrobacterium*-mediated floral dip method	Dynamic controls for the evolution of transposon-rich heterochromatic regions	Tsukahara et al. (2012)
Tnt1	Tobacco	Tyl/*Copia*	*Agrobacterium*-mediated regeneration method	The promoter of *Tnt1* is a target of silencing in tobacco	Hernández-Pinzón et al. (2012)
Tcs1	Tobacco	Tyl/*Copia*	*Agrobacterium*-mediated regeneration method	Transposition and recombination of retroelements are likely important sources of variation in *Citrus*	Butelli et al. (2012)

Table 1 (continued)

LTRs	Species	Superfamilies	Mode of plant transformation	Results	References
LIRE	*Lilium*	Ty1/*Copia*	Wild-type Liliaceae plants	Might have contributed to the expansion of the genome in the genus *Lilium*	Lee et al. (2013)
Evadé (EVD)	*Arabidopsis*	Tyl/*Copia*	*Agrobacterium*-mediated floral dip method	TE burst causes widespread genome diversification and de novo silencing of retrotransposons	Marí-Ordóñez et al. (2013)
OAR1	*Arabidopsis*	Ty3/*Gypsy*	*Agrobacterium*-mediated floral dip method	Potential role of *OAR1* element in plant tolerance to osmotic and alkaline stresses	Zhao et al. (2014)
BAGY2	Barley	Ty3/*Gypsy*	Seed-embryo callus culture transformation	Retrotransposition events have occurred during calli development and shoot regeneration	Yilmaz et al. (2014)
MIKKI	Rice and *Arabidopsis*	*BAJIE*	Seed-embryo callus transformation; *Agrobacterium*-mediated floral dip method	*MIKKI* acts as a target mimic of *osa-miR171*	Cho and Paszkowski (2017)
ONSEN	*Arabidopsis* and *Japanese radish*	Tyl/*Copia*	Callus induction tissue culture	Heat-stressed tissues and their self-fertilized progeny, revealing the possibility of molecular breeding without genetic modification	Masuta et al. (2017)
EARE-1	*Excoecaria agallocha*	Ty1/*Copia*	Horizontal gene transfer	The significant role of post-transcriptional host control in the life cycles of transposable elements	Huang et al. (2017)
ATCOPIA93	*Arabidopsis*	Ty1/*Copia*	*Agrobacterium*-mediated floral dip method	Responsiveness of *ATCOPIA93* to biotic stress and the co-option of its LTR for plant immunity	Zervudacki et al. (2018)
HUO	Rice	Ty1/*Copia*	Reciprocal crossing	Important for genome stability during crop domestication and breeding	Peng et al. (2019)
CLEVR	*Drosophila*	Ty3-*Gypsy*	Transgenes and mutants were back-crossed to wild-type	Universally conserved features of retroviruses, and should be widely applicable to other LTR retrotransposons	Chang et al. (2019)
Tos17	Rice	Ty1/*Copia*	Calli transformed with CRISPR/Cas9 vectors	Providing rapid breeding technology as an alternative means to re-activate the expression of agronomically important genes that have been inactivated by TE insertion	Saika et al. (2019)
ZmRE-1	Maize	Tyl/*Copia*	Crossed with inbred line	Potential to improve the grain yield per unit area through increasing the planting density	Li et al. (2020)
Tos17 $^{Chr.7}$	Rice	Ty1/*Copia*	CRISPR/Cas9, *Agrobacterium*-mediated regeneration method	Transgenic rice plants for gene function study and genetic engineering	Luo et al. (2020)
MAGO1/2	Maize	Ty3/*Gypsy*	*Agrobacterium*-mediated regeneration method	Mutagenic activity of transposons in male germ cells	Lee et al. (2021)

transgenic *Arabidopsis* plants, has been shown to behave like an immune-responsive gene during pathogen defense, which might establish the connection between the responsiveness and retrotransposons to biotic stress (Zervudacki et al. 2018). Hence, LTR retrotransposons can be targeted for biotic stress like immune responses in plants.

A maize retrotransposon, *ΔNaeAc*, introduced into a flax callus by *Agrobacterium*-mediated transformation showed enhanced transcription and transposition in the callus, which could emphasize *Ac* element behavior in different plant species (Finnegan et al. 1993). More recently, the insertion of the maize LTR retrotransposon, *ZmRE-1*, in the fifth exon of *Brachytic2* (*Br2*) allele, was identified in dwarf mutants *dwarf2014* (*d2014*) at exactly the same site, which indicates that the transposition of *ZmRE-1* might be correlated with the change in height, and has the potential to improve grain yield and planting density (Li et al. 2020). A *BARE-1* LTR retrotransposon in barley's genome possessing functional TATA boxes is required to drive the transient expression of reporter genes transformed in barley protoplasts, suggesting that *BARE-1* could be the potential retrotransposon for propagation in barley (Suoniemi et al. 1996; Kalendar et al. 2000; Vicient et al. 2001). Therefore, the *BARE-1* LTR retrotransposon can be used for epigenetic modification, targeting barley propagation. A *TLC1* family of LTR retrotransposon from *Lycopersicon chilense* has the *cis*-regulatory elements required for ethylene in stress-induced gene expression in transgenic plants and protoplasts, which could play a major role in the transcriptional activation of the *TLC1* element (Tapia et al. 2005).

The *Lotus japonicus* LTR retrotransposon, *LORE1*, has been located in gene-rich to centromeric heterochromatin regions, and the new insertion mutagenesis of *LORE1* demonstrated transposition into genes of highly repetitive sequences in centromeres and telomeres, indicating that *LORE1* might be an effective LTR retrotransposon in the intact plant during *in vitro* tissue culture (Madsen et al. 2005). In another report, the chromovirus *LORE1* family of LTR retrotransposon was epigenetically silenced in transgenic plants established by *Agrobacterium*-mediated transformation. The new insertion sites of *LORE1* copies were frequently found in genic regions and showed no strong insertional preferences, and the distinct features of *LORE1* have significant potential for generating genetic and epigenetic processes on evolution in host plants (Fukai et al. 2010). The genome of rice lines derived from seed embryo callus culture showed new copies of a transcriptionally activated LTR retrotransposon, *Lullaby,* which is an interesting candidate for a *cis*-acting element that could account for transcriptional activation in rice calli (Picault et al. 2009).

The osmotic and alkaline tolerance LTR retrotransposon, *OAR1*, in *Arabidopsis*, demonstrated that the transgenic plants exhibit enhanced photochemical efficiency, membrane integrity, and biomarker gene expression during osmotic and alkaline stresses (Zhao et al. 2014). Furthermore, the genetic variation in calli culture and the regeneration of the *BAGY2* LTR retrotransposon (Leigh et al. 2003; Hosid et al. 2012) in the barley genome has shown enhanced copy number variations of the internal domains during tissue culture (Yilmaz et al. 2014). Interestingly, a root-specific LTR retrotransposon, *Mikki*, in rice and in *Arabidopsis,* is highly transcribed in roots during tissue culture, and the spliced transcripts constitute a target mimic for *miR171*, indicating that retrotransposon-derived transcripts could act as a decoy for *miR171* degradation, resulting in root-specific accumulation of *SCARECROW-Like* mRNAs in various rice species (Cho and Paszkowski 2017). The *ONSEN* retrotransposon in *Arabidopsis* and Japanese radish (*Raphanus sativus*) indicates that *ONSEN* could be transposed in heat-stressed callus tissue and subsequently regenerated tissues, which, in turn, implies that the heat shock transcription factor and RNA-directed DNA methylation (RdDM)-related genes might regulate the transcription and transposition of *ONSEN* under heat stress (Masuta et al. 2017).

The multiple copies of the *HUO* LTR retrotransposon in rice, generated by reciprocal crossing, might trigger genomic instability by altering small RNA biogenesis and genome-wide DNA methylation, resulting in decreased disease resistance and yield, as evidenced by *HUO* LTR retrotransposon elimination during rice domestication and breeding (Peng et al. 2019). Cellular labeling of endogenous retrovirus replication (*CLEVR*) reveals the replication of TY3/*Gypsy* LTR retrotransposons in both cell culture and individual neurons, while the *Gypsy-CLEVR* replication rate is enhanced when the siRNA pathway is genetically disrupted, indicating that the *CLEVR* strategy might apply to other retrotransposons in diverse plant species (Chang et al. 2019). It also suggests that control over retrotransposons has significant implications in epigenetic modification and can be used for genetic engineering. More recently, the *EARE-1* LTR retrotransposon in the genome of *Excoecaria agallocha* showed elevated expression in organs examined under stress, suggesting that both the horizontal transfer of retrotransposons and the post-transcriptional gene silencing of the host might play significant roles in the life cycle of *EARE-1* (Huang et al. 2017).

Advantages of using LTR retrotransposons in plant genetic engineering

Contributing to several molecular functions in the genome (Fig. 2), LTR retrotransposons are primarily in-house genetic elements of the genome. These elements move independently (function autonomously) around the genome without external vectors. To modify genomic functions, the activation or deactivation of LTRs can be achieved

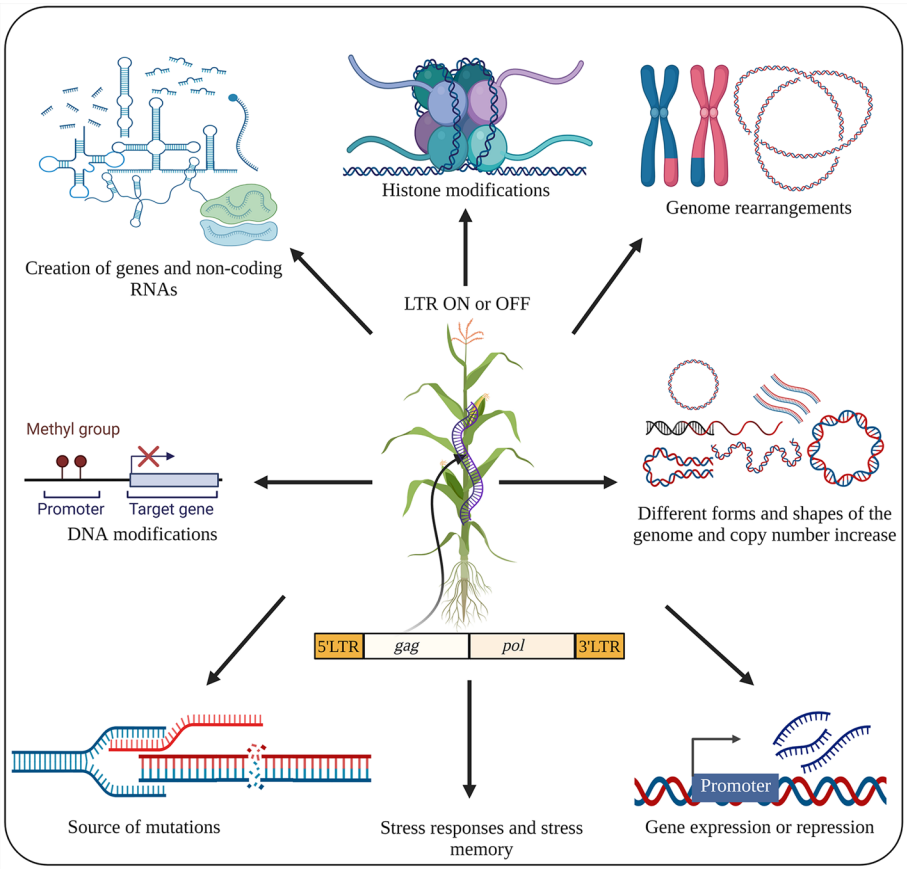

Fig. 2 Applications of LTR retrotransposons in plant genetic engineering. Created with BioRender.com

through physical stresses or chemical treatment (enzymes). For instance, RNA editing enzymes, such as adenosine deaminse acting on double-stranded RNAs (ADARs), play a dynamic role in regulating transcriptome and proteome diversity (Piontkivska et al. 2021). It has been found that ADAR editing and the ERI-6/7/MOV10 RNAi pathway (the endogenous RNA interference factor ERI-6/7, a homolog of MOV10 helicase) silence LTR retrotransposons associated with human autoimmune disorders and neurodegenerative diseases (Fischer and Ruvkun 2020). Identifying such pathways that regulate LTRs in plants may lead to the development of disease-free plants.

The activity of the retrotransposon, *Tos17* $^{\text{Chr.7}}$, in rice varieties, through 5-azacytidine (5-azaC) treatment, resulted in methylation and activation, and *Tos17* copies are not only methylated to some extent in all varieties but a gradual increase in DNA methylation was also observed with the growth of the plant (Cheng et al. 2006). More focused research in the chemical activators of enzymes that activate LTRs for desired genetic change may contribute to plant breeding, and to the successful development of crops with the desired traits. The copy numbers of LTRs associated with specific genetic changes and desired traits could also be increased (Kumar and Hirochika 2001). Moreover, endogenous active copies of retrotransposons can be isolated using simple reverse transcription PCR (Hirochika 1993). These insights into the possible advantages of using LTRs are just the tip of the iceberg: their use as genetic engineering tools awaits further advances in research.

Hurdles in using LTR retrotransposons in plant genetic engineering

Although LTR retrotransposons are predicted to be good tools for plant genetic engineering, in practice, there are many hurdles when it comes to their application: lack of control over copy numbers and differences in the expression of the copies. High copy numbers cause either beneficial or harmful mutations, regulate gene functions, and maintain genome stability (Belyayev et al. 2010). Copy numbers of retrotransposons range from hundreds to hundreds of thousands. For instance, the Ty1-*copia* element, *BARE-1*, in barley, ranges from 20,000 to 200,000 copies (Vicient et al. 2000; Shelke and Das 2015), and these copies are mainly located in heterochromatic regions including centromeres (Feng et al. 2002). To balance copy numbers, plants have evolved a complex regulatory network of epigenetic mechanisms to silence TE activity (Sinzelle et al. 2009; Lisch 2013). Thus, most copies are dead or epigenetically silenced,

and high levels of DNA methylation at cytosine nucleotides have been associated with TE silencing (Cavrak et al. 2014). The level of epigenetic modification is also associated with the silencing of retrotransposons, but how the genome balances the copy numbers remains unclear.

According to the stages of plant growth and development, many copies are often activated and deactivated by epigenetic modification. For example, a null mutation of *Arabidopsis* maintenance METHYLTRANSFERASE 1 (MET1) activates *EVADÉ* (*EVD*), a retrotransposon of the *ATCOPIA93* family capable of amplification during the sexual propagation of the mutant plant (Mirouze et al. 2009). Another major challenge is the difference in the copy number expression of the same family, and these differences also vary among species and according to plant developmental stages; moreover, they are highly heterogeneous, and polymorphic at insertion sites (Liu et al. 2022). Although recent high-throughput sequencing technologies have advanced the prediction of copy number differences, copy number expression, differences, and movement are still not accurately measurable because copy numbers are not directly controlled, while several genes and non-coding RNAs are derived from these copies. Because of these challenges, it is difficult to control active copies to make targeted changes in genomic architecture. However, the activity of the host RNA polymerase II plays a significant role in the repression of retrotransposons, because the mobility of elements mostly depends on Pol II (Hermant and Torres-Padilla 2021). Moreover, the synergistic inhibition of DNA methylation and Pol II activity can create a strong stress-dependent mobilization of the heat-responsive *ONSEN* in *Arabidopsis* (Thieme et al. 2017). Likewise, dead copies can also be activated using epigenetic modification and targeted genome editing.

Apart from this, to overcome the hurdles of LTR application in plant genetic engineering, advancement of technologies with emerging research tools, such as big data, machine learning (ML), and artificial intelligence-based deep learning methods, are necessary to predict and explore methods to control LTRs. Integration of precise genome editing technology with big data analysis, such as the development of big data tools like TEtools (Lerat et al. 2017), may produce new insights into control over retrotransposons. Moreover, ML algorithms are useful in automatically learning the parameters needed to fit a model to a specific problem (Shastry and Sanjay 2020), called supervised learning (Zou et al. 2018). More recently, a machine-learning technique has been developed to classify LTR retrotransposons in plant genomes (Orozco-Arias et al. 2021). Further such developments in computing and a new generation of long-read sequencing technologies can contribute better methods for understanding retrotransposon movement, silencing, activation, copy number changes, and expression, which may lead in future to achieving control over LTR retrotransposons.

Future perspective

1. LTR retrotransposons play a significant role in stress management. Understanding the mechanisms involved in the activation of specific LTR retrotransposons during stress can be a game changer in plant breeding techniques to develop climate-resilient crops.
2. LTR retrotransposons are in-house genetic sequences, so using them instead of foreign plasmid or viral vectors might be less likely to induce silencing in the modified locus or loci of the target genome.
3. Control of LTR retrotransposons through epigenetic modification (i.e., turning gene expression on or off) can be a promising future research focus for developing LTR retrotransposons as a genetic engineering tool.
4. Retrotransposons are known to depend on the transcriptional activity of the host RNA polymerase II for their mobility. Nonetheless, we do not know what the direct involvement of the host RNA polymerase II is in suppressing the activity of retrotransposons.
5. Recent high-throughput long-read sequencing technologies have advanced the prediction of copy number difference. However, technology must advance further in order to accurately measure the copy number expression and differences in their movement.
6. Novel methods and understandings in big data, machine learning, and deep learning-based computation of high-throughput sequencing data might help us regulate LTR retrotransposons in the future.
7. Certain chemicals have proven to be useful in activating or silencing LTR retrotransposons. So, the development of such chemical combinations can be utilized as a tool for making favorable changes in the genetic architecture of plants.
8. Identifying physical stress-mediated activation and silencing of LTR retrotransposons, and their favorable effect on stress mitigation in plants, along with heritable epigenetic memory, can help to develop stress-resistant crop plants.
9. In certain cases, activating dead copies of LTR retrotransposons in plants may induce favorable changes in gene expression and/or genome organization. Such outcomes may offer good opportunities for plant epigenetics and plant genetic engineering research.

Conclusion

Retrotransposons are retrovirus-like intragenic elements, capable of making many changes in genome organization. LTR retrotransposons form a major portion of the genomic

DNA in eukaryotes. This is an opportunity to develop LTR retrotransposons as a tool for integrating desirable genetic sequences into the target genome. However, these LTR retrotransposons are mostly inactive and cannot be controlled for targeted genomic insertion. Recently, researchers have explored the role of LTR retrotransposons in crop plants, leading to insights into specific roles of LTR retrotransposons and their molecular mechanisms. Studies on epigenetic modifications capable of controlling LTR retrotransposons for specific gene expressions can provide techniques for using LTR retrotransposons as tools for plant genetic engineering. Future research will be able to use advancements in molecular technologies. Computational methods like big data and machine learning may help develop regulating procedures for LTR retrotransposons.

Acknowledgements The authors wish to thank the University of Helsinki Language Centre, Finland for outstanding editing and proofreading of the manuscript. We sincerely thank the anonymous reviewers for their valuable time and insightful comments, which greatly improved the quality of the manuscript. We apologize to those whose original work(s) could not be included in this review owing to space limitations.

Author contribution statement MR, PKP and SM planned, designed and wrote the review. MR, SM, QW, RK and MZ outlined and edited the review. MR, PKP, RK and QW made the table and the images. MR, RK, QW, SM, AS, ZA, VS, and MZ edited and revised the review.

Funding Preparation of this review was supported by a grant from the Jiangxi "Shuangqian" Programme (S2019DQKJ2030), the Qing Lan Project of Jiangsu Higher Education Institutions, the Natural Science Foundation for Distinguished Young Scholars of Nanjing Forestry University (JC2019004), the Project for Groundbreaking Achievements of Nanjing Forestry University (202211), a project funded by the Priority Academic Programme Development of Jiangsu Higher Education Institutions, and also the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (OR11465424). The authors are also grateful for the support of Metasequoia Faculty Research Start-up Funding (grant number 163100028) at the Bamboo Research Institute, Nanjing Forestry University for the first author MR.

Declarations

Conflict of interest The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Alzohairy AM, Sabir JSM, Gyulai G, Younis RAA, Jansen RK, Bahieldin A (2014) Environmental stress activation of plant long-terminal repeat retrotransposons. Funct Plant Biol 41:557–567

Barbeau B, Mesnard J-M (2011) Making sense out of antisense transcription in human T-cell lymphotropic viruses (HTLVs). Viruses 3:456–468

Baum C, Schambach A, Bohne J, Galla M (2006) Retrovirus vectors: toward the Plentivirus? Mol Ther 13:1050–1063

Belyayev A, Kalendar R, Brodsky L, Nevo E, Schulman AH, Raskina O (2010) Transposable elements in a marginal plant population: temporal fluctuations provide new insights into genome evolution of wild diploid wheat. Mob DNA 1:6

Bourque G, Burns KH, Gehring M, Gorbunova V, Seluanov A, Hammell M, Imbeault M, Izsvák Z, Levin HL, Macfarlan TS, Mager DL, Feschotte C (2018) Ten things you should know about transposable elements. Genome Biol 19:199

Butelli E, Licciardello C, Zhang Y, Liu J, Mackay S, Bailey P, Reforgiato-Recupero G, Martin C (2012) Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. Plant Cell 24:1242–1255

Buttler CA, Chuong EB (2022) Emerging roles for endogenous retroviruses in immune epigenetic regulation. Immunol Rev 305:165–178

Casacuberta JM, Grandbastien M-A (1993) Characterisation of LTR sequences involved in the protoplast specific expression of the tobacco *Tnt1* retrotransposon. Nucleic Acids Res 21:2087–2093

Cavrak VV, Lettner N, Jamge S, Kosarewicz A, Bayer LM, Mittelsten Scheid O (2014) How a retrotransposon exploits the plant's heat stress response for its activation. PLoS Genet 10:e1004115

Chang Y-H, Keegan RM, Prazak L, Dubnau J (2019) Cellular labeling of endogenous retrovirus replication (*CLEVR*) reveals de novo insertions of the *Gypsy* retrotransposable element in cell culture and in both neurons and glial cells of aging fruit flies. PLoS Biol 17:e3000278

Cheng C, Daigen M, Hirochika H (2006) Epigenetic regulation of the rice retrotransposon *Tos17*. Mol Genet Genom 276:378–390

Cheng X, Zhang D, Cheng Z, Keller B, Ling HQ (2009) A new family of Ty1-*copia*-like retrotransposons originated in the tomato genome by a recent horizontal transfer event. Genetics 181:1183–1193

Cho J, Paszkowski J (2017) Regulation of rice root development by a retrotransposon acting as a microRNA sponge. Elife 6:e30038

Courtial B, Feuerbach F, Eberhard S, Rohmer L, Chiapello H, Camilleri C, Lucas H (2001) *Tnt1* transposition events are induced by in vitro transformation of *Arabidopsis thaliana*, and transposed copies integrate into genes. Mol Genet Genomics 265:32–42

D'erfurth I, Cosson V, Eschstruth A, Lucas H, Kondorosi A, Ratet P (2003) Efficient transposition of the *Tnt1* tobacco retrotransposon in the model legume *Medicago truncatula*. Plant J 34:95–106

Diao X, Freeling M, Lisch D (2005) Horizontal transfer of a plant transposon. PLoS Biol 4:e5

Feng Q, Zhang Y, Hao P, Wang S, Fu G, Huang Y et al (2002) Sequence and analysis of rice chromosome 4. Nature 420:316–320

Finnegan EJ, Lawrence GJ, Dennis ES, Ellis JG (1993) Behaviour of modified *Ac* elements in flax callus and regenerated plants. Plant Mol Biol 22:625–633

Fischer SEJ, Ruvkun G (2020) *Caenorhabditis elegans* ADAR editing and the ERI-6/7/MOV10 RNAi pathway silence endogenous viral elements and LTR retrotransposons. Proc Natl Acad Sci USA 117:5987–5996

Fukai E, Umehara Y, Sato S, Endo M, Kouchi H, Hayashi M, Stougaard J, Hirochika H (2010) Derepression of the plant chromovirus *LORE1* induces germline transposition in regenerated plants. PLoS Genet 6:e1000868

Galindo-Gonzalez L, Mhiri C, Deyholos MK, Grandbastien MA (2017) LTR-retrotransposons in plants: Engines of evolution. Gene 626:14–25

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

Grandi N, Tramontano E (2018) HERV envelope proteins: Physiological role and pathogenic potential in cancer and autoimmunity. Front Microbiol 9:462

Havecker ER, Gao X, Voytas DF (2004) The diversity of LTR retrotransposons. Genome Biol 5:225

Hermant C, Torres-Padilla M-E (2021) TFs for TEs: the transcription factor repertoire of mammalian transposable elements. Genes Dev 35:22–39

Hernández-Pinzón I, Cifuentes M, Hénaff E, Santiago N, Espinás ML, Casacuberta JM (2012) The *Tnt1* retrotransposon escapes silencing in tobacco, its natural host. PLoS ONE 7:e33816

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

Hirochika H, Sugimoto K, Otsuki Y, Tsugawa H, Kanda M (1996) Retrotransposons of rice involved in mutations induced by tissue culture. Proc Natl Acad Sci USA 93:7783–7788

Hirochika H, Okamoto H, Kakutani T (2000) Silencing of retrotransposons in *Arabidopsis* and reactivation by the *ddm1* mutation. Plant Cell 12:357–368

Hosid E, Brodsky L, Kalendar R, Raskina O, Belyayev A (2012) Diversity of long terminal repeat retrotransposon genome distribution in natural populations of the wild diploid wheat *Aegilops speltoides*. Genetics 190:263–274

Huang J, Wang Y, Liu W, Xu S, Fan Q, Jian S, Tang T (2017) *EARE-1*, a transcriptionally active Ty1/*Copia*-like retrotransposon has colonized the genome of *Excoecaria agallocha* through horizontal transfer. Front Plant Sci 8:45

Jiang N, Bao Z, Temnykh S, Cheng Z, Jiang J, Wing RA, Mccouch SR, Wessler SR (2002) Dasheng: a recently amplified nonautonomous long terminal repeat element that is a major component of pericentromeric regions in rice. Genetics 161:1293–1305

Kakutani T, Jeddeloh JA, Flowers SK, Munakata K, Richards EJ (1996) Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. Proc Natl Acad Sci USA 93:12406–12411

Kalendar R, Tanskanen J, Immonen S, Nevo E, Schulman Alan H (2000) Genome evolution of wild barley (*Hordeum spontaneum*) by *BARE-1* retrotransposon dynamics in response to sharp microclimatic divergence. Proc Natl Acad Sci USA 97:6603–6607

Kumar A, Hirochika H (2001) Applications of retrotransposons as genetic tools in plant biology. Trends Plant Sci 6:127–134

Leblanc P, Desset S, Giorgi F, Taddei AR, Fausto AM, Mazzini M, Dastugue B, Vaury C (2000) Life cycle of an endogenous retrovirus, ZAM, in *Drosophila melanogaster*. J Virol 74:10658–10669

Lee S-I, Park K-C, Son J-H, Hwang Y-J, Lim K-B, Song Y-S, Kim J-H, Kim N-S (2013) Isolation and characterization of novel Ty1-*copia*-like retrotransposons from lily. Genome 56:495–503

Lee Y-S, Maple R, Dürr J, Dawson A, Tamim S, Del Genio C, Papareddy R, Luo A, Lamb JC, Amantia S (2021) A transposon surveillance mechanism that safeguards plant male fertility during stress. Nat Plants 7:34–41

Leigh F, Kalendar R, Lea V, Lee D, Donini P, Schulman AH (2003) Comparison of the utility of barley retrotransposon families for genetic analysis by molecular marker techniques. Mol Genet Genom 269:464–474

Lerat E, Fablet M, Modolo L, Lopez-Maestre H, Vieira C (2017) TEtools facilitates big data expression analysis of transposable elements and reveals an antagonism between their activity and that of piRNA genes. Nucleic Acids Res 45:e17

Li C, Tang J, Hu Z, Wang J, Cao M (2020) A novel maize dwarf mutant generated by Ty1-*copia* LTR-retrotransposon insertion in *Brachytic2* after spaceflight. Plant Cell Rep 39:1–16

Li F, Lee M, Esnault C, Wendover K, Guo Y, Atkins P, Zaratiegui M, Levin HL (2022) Identification of an integrase-independent pathway of retrotransposition. Sci Adv 8:eabm9390

Lisch D (2013) How important are transposons for plant evolution? Nat Rev Genet 14:49–61

Liu J-X, Liu J, Yang T-Y, Liu L-M, Qiu L-H, Gao Y-J, Duan W-X, Lei J-C, Liu H-J, Zhang R-H, He W-Z, Li S, Li Y-R, Wei R-C, Xiong F-Q (2022) Isolation and analysis of reverse transcriptase of Ty1-*copia*-like retrotransposons in sugarcane. Sugar Tech. https://doi.org/10.1007/s12355-021-01063-6

Luo Y, Tian D, Teo JCY, Ong KH, Yin Z (2020) Inactivation of retrotransposon *Tos17*^{Chr} 7 in rice cultivar Nipponbare through CRISPR/Cas9-mediated gene editing. Plant Biotechnol 37:69–75

Madsen LH, Fukai E, Radutoiu S, Yost CK, Sandal N, Schauser L, Stougaard J (2005) *LORE1*, an active low-copy-number Ty3-*Gypsy* retrotransposon family in the model legume *Lotus japonicus*. Plant J 44:372–381

Marí-Ordóñez A, Marchais A, Etcheverry M, Martin A, Colot V, Voinnet O (2013) Reconstructing de novo silencing of an active plant retrotransposon. Nat Genet 45:1029–1039

Masuta Y, Nozawa K, Takagi H, Yaegashi H, Tanaka K, Ito T, Saito H, Kobayashi H, Matsunaga W, Masuda S (2017) Inducible transposition of a heat-activated retrotransposon in tissue culture. Plant Cell Physiol 58:375–384

Melayah D, Bonnivard E, Chalhoub B, Audeon C, Grandbastien MA (2001) The mobility of the tobacco *Tnt1* retrotransposon correlates with its transcriptional activation by fungal factors. Plant J 28:159–168

Mhiri C, Morel J-B, Vernhettes S, Casacuberta JM, Lucas H, Grandbastien M-A (1997) The promoter of the tobacco *Tnt1* retrotransposon is induced by wounding and by abiotic stress. Plant Mol Biol 33:257–266

Mirouze M, Reinders J, Bucher E, Nishimura T, Schneeberger K, Ossowski S, Cao J, Weigel D, Paszkowski J, Mathieu O (2009) Selective epigenetic control of retrotransposition in *Arabidopsis*. Nature 461:427–430

Miyao A, Tanaka K, Murata K, Sawaki H, Takeda S, Abe K, Shinozuka Y, Onosato K, Hirochika H (2003) Target site specificity of the *Tos17* retrotransposon shows a preference for insertion within genes and against insertion in retrotransposon-rich regions of the genome. Plant Cell 15:1771–1780

Ong-Abdullah M, Ordway JM, Jiang N, Ooi SE, Kok SY, Sarpan N, Azimi N, Hashim AT, Ishak Z, Rosli SK, Malike FA, Bakar NA, Marjuni M, Abdullah N, Yaakub Z, Amiruddin MD, Nookiah R, Singh R, Low ET, Chan KL, Azizi N, Smith SW, Bacher B, Budiman MA, Van Brunt A, Wischmeyer C, Beil M, Hogan M, Lakey N, Lim CC, Arulandoo X, Wong CK, Choo CN, Wong WC, Kwan YY, Alwee SS, Sambanthamurthi R, Martienssen RA (2015) Loss of *Karma* transposon methylation underlies the mantled somaclonal variant of oil palm. Nature 525:533–537

Orozco-Arias S, Isaza G, Guyot R (2019) Retrotransposons in plant genomes: structure, identification, and classification through bioinformatics and machine learning. Int J Mol Sci 20:3837

Orozco-Arias S, Candamil-Cortés MS, Jaimes PA, Piña JS, Tabares-Soto R, Guyot R, Isaza G (2021) *K*-mer-based machine learning method to classify LTR-retrotransposons in plant genomes. PeerJ 9:e11456–e11456

Pelisson A, Mejlumian L, Robert V, Terzian C, Bucheton A (2002) *Drosophila* germline invasion by the endogenous retrovirus *Gypsy*: involvement of the viral env gene. Insect Biochem Mol Biol 32:1249–1256

Peng Y, Zhang Y, Gui Y, An D, Liu J, Xu X, Li Q, Wang J, Wang W, Shi C (2019) Elimination of a retrotransposon for quenching genome instability in modern rice. Mol Plant 12:1395–1407

Picault N, Chaparro C, Piegu B, Stenger W, Formey D, Llauro C, Descombin J, Sabot F, Lasserre E, Meynard D (2009) Identification of an active LTR retrotransposon in rice. Plant J 58:754–765

Piontkivska H, Wales-Mcgrath B, Miyamoto M, Wayne ML (2021) ADAR editing in viruses: An evolutionary force to reckon with. Genome Biol Evol 13:evab240

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

Quesneville H (2020) Twenty years of transposable element analysis in the *Arabidopsis thaliana* genome. Mob DNA 11:28

Rajput MK, Upadhyaya KC (2009) *CARE1*, a Ty3-*Gypsy* like LTR-retrotransposon in the food legume chickpea (*Cicer arietinum* L.). Genetica 136:429–437

Ramakrishnan M, Satish L, Kalendar R, Narayanan M, Kandasamy S, Sharma A, Emamverdian A, Wei Q, Zhou M (2021) The dynamism of transposon methylation for plant development and stress adaptation. Int J Mol Sci 22:11387

Ramakrishnan M, Satish L, Sharma A, Kurungara Vinod K, Emamverdian A, Zhou M, Wei Q (2022) Transposable elements in plants: recent advancements, tools and prospects. Plant Mol Biol Rep. https://doi.org/10.1007/s11105-022-01342-w

Roulin A, Piegu B, Wing RA, Panaud O (2008) Evidence of multiple horizontal transfers of the long terminal repeat retrotransposon *RIRE1* within the genus *Oryza*. Plant J 53:950–959

Saika H, Mori A, Endo M, Toki S (2019) Targeted deletion of rice retrotransposon *Tos17* via CRISPR/Cas9. Plant Cell Rep 38:455–458

Sato M, Kawabe T, Hosokawa M, Tatsuzawa F, Doi M (2011) Tissue culture-induced flower-color changes in *Saintpaulia* caused by excision of the transposon inserted in the *flavonoid 3′, 5′ hydroxylase* (*F3′5′H*) promoter. Plant Cell Rep 30:929–939

Shastry KA, Sanjay HA (2020) Machine learning for bioinformatics. In: Srinivasa K, Siddesh G, Manisekhar S (eds) Statistical modelling and machine learning principles for bioinformatics techniques, tools, and applications Algorithms for intelligent systems. Springer Singapore, pp 25–39

Shelke RG, Das AB (2015) Analysis of genetic diversity in 21 genotypes of Indian banana using RAPDs and IRAPs markers. Proc Natl Acad Sci India Sect B Biol Sci 85:1027–1038

Sinzelle L, Izsvák Z, Ivics Z (2009) Molecular domestication of transposable elements: From detrimental parasites to useful host genes. Cell Mol Life Sci 66:1073–1093

Steinbauerová V, Neumann P, Macas J (2008) Experimental evidence for splicing of intron-containing transcripts of plant LTR retrotransposon *Ogre*. Mol Genet Genom 280:427–436

Sultana T, Zamborlini A, Cristofari G, Lesage P (2017) Integration site selection by retroviruses and transposable elements in eukaryotes. Nat Rev Genet 18:292–308

Suoniemi A, Narvanto A, Schulman AH (1996) The *BARE-1* retrotransposon is transcribed in barley from an LTR promoter active in transient assays. Plant Mol Biol 31:295–306

Takeda S, Sugimoto K, Otsuki H, Hirochika H (1998) Transcriptional activation of the tobacco retrotransposon *Tto1* by wounding and methyl jasmonate. Plant Mol Biol 36:365–376

Takeda S, Sugimoto K, Otsuki H, Hirochika H (1999) A 13-bp *cis*-regulatory element in the LTR promoter of the tobacco retrotransposon *Tto1* is involved in responsiveness to tissue culture, wounding, methyl jasmonate and fungal elicitors. Plant J 18:383–393

Tapia G, Verdugo I, Yañez M, Ahumada I, Theoduloz C, Cordero C, Poblete F, González E, Ruiz-Lara S (2005) Involvement of ethylene in stress-induced expression of the *TLC1. 1* retrotransposon from *Lycopersicon chilense* Dun. Plant Physiol 138:2075–2086

Thieme M, Lanciano S, Balzergue S, Daccord N, Mirouze M, Bucher E (2017) Inhibition of RNA polymerase II allows controlled mobilisation of retrotransposons for plant breeding. Genome Biol 18:134–134

Tsukahara S, Kawabe A, Kobayashi A, Ito T, Aizu T, Shin-I T, Toyoda A, Fujiyama A, Tarutani Y, Kakutani T (2012) Centromere-targeted de novo integrations of an LTR retrotransposon of *Arabidopsis lyrata*. Genes Dev 26:705–713

Vicient CM, Casacuberta JM (2020) Additional ORFs in plant LTR-retrotransposons. Front Plant Sci 11:555

Vicient CM, Kalendar R, Anamthawat-Jónsson K, Suoniemi A, Schulman AH (2000) Structure, functionality, and evolution of the *BARE-1* retrotransposon of barley. In: McDonald JF (ed) Transposable elements and genome evolution. Georgia genetics review 1, vol 1. Springer Netherlands, pp 53–63

Vicient CM, JääSkeläInen MJ, Kalendar R, Schulman AH (2001) Active retrotransposons are a common feature of grass genomes. Plant Physiol 125:1283–1292

Yilmaz S, Marakli S, Gozukirmizi N (2014) *BAGY2* retrotransposon analyses in barley calli cultures and regenerated plantlets. Biochem Genet 52:233–244

Zervudacki J, Yu A, Amesefe D, Wang J, Drouaud J, Navarro L, Deleris A (2018) Transcriptional control and exploitation of an immune-responsive family of plant retrotransposons. EMBO J 37:e98482

Zhang H, Lang Z, Zhu JK (2018) Dynamics and function of DNA methylation in plants. Nat Rev Mol Cell Biol 19:489–506

Zhao Y, Xu T, Shen C-Y, Xu G-H, Chen S-X, Song L-Z, Li M-J, Wang L-L, Zhu Y, Lv W-T (2014) Identification of a retroelement from the resurrection plant *Boea hygrometrica* that confers osmotic and alkaline tolerance in *Arabidopsis thaliana*. PLoS ONE 9:e98098

Zou J, Huss M, Abid A, Mohammadi P, Torkamani A, Telenti A (2018) A primer on deep learning in genomics. Nat Genet 51:12–18