Chapter 6
The Impact of Transposable Elements in the Evolution of Plant Genomes: From Selfish Elements to Key Players

Beatriz Contreras, Cristina Vives, Roger Castells and Josep M. Casacuberta

Abstract Transposable elements (TEs) are major components of all eukaryote genomes, and in particular of plant genomes. Whereas these elements have long been considered as selfish 'junk DNA without function', the data accumulated over the years have shown that they are essential components of the genome structure and key players of genome evolution. Here, we summarize the recent advancement in the field and we discuss the role of TEs in the light of the new data coming from whole plant genome sequences and next-generation sequencing (NGS) data on resequencing of plant varieties and lines.

6.1 Transposable Elements, a Major Component of Plant Genome

Transposable elements (TEs) are mobile genetic elements that account for an important fraction of virtually all eukaryote genomes. TEs can be classified into two major classes, class I (retrotransposons) and class II (DNA transposons). Class I elements transpose through an RNA intermediate used as a template in a reverse transcription reaction leading to a new DNA copy that can integrate back into the genome. Therefore, class I TEs do not excise during transposition and their copy number increases as a result of their movement. Whereas the transcription of the element is catalysed by the host's polymerase (Pol II), its reverse transcription and integration are catalysed by enzymatic activities encoded by the retrotransposon itself, in case of autonomous elements, or by a related element, in case of non-autonomous elements. Class II elements transpose via a DNA intermediate, which results from the excision of the element from its chromosomal location and that can be integrated elsewhere into the genome. Both the excision and integration reactions are catalysed by a transposase which is encoded by the mobilized TE in

B. Contreras · C. Vives · R. Castells · J.M. Casacuberta (✉)
Centre de Recerca En Agrigenòmica, CRAG (CSIC-IRTA-UAB-UB), Barcelona, Spain
e-mail: josep.casacuberta@cragenomica.es

P. Pontarotti (ed.), *Evolutionary Biology: Biodiversification from Genotype to Phenotype*, DOI 10.1007/978-3-319-19932-0_6

case of autonomous elements or by a related element in case of a defective TE copy. There are, however, some DNA transposons that move through a different mechanism. This is the case of *Helitrons*, which transpose via a rolling-circle mechanism similar to that of some bacterial TEs. Both class I and class II TEs can be further classified into families and subfamilies depending on their structure, encoded proteins and mechanism of transposition (Wicker et al. 2007).

Whereas TEs are commonplace in eukaryotes, and most eukaryotes contain elements belonging to all major types and classes, their prevalence differs from genome to genome. TEs account for a major but variable fraction of plant genomes (Bennetzen and Wang 2014), with LTR retrotransposons and miniature inverted-repeat transposable elements (MITEs) tending to be the most represented types of TEs (Casacuberta and Santiago 2003). The variability in TE content is huge in plants. For instance, as much as 85 % of maize genome or 70 % of Norway spruce genome (Nystedt et al. 2013) has been annotated as transposons, whereas transposon annotations make only the 21 % of the more compact *Arabidopsis thaliana* genome (Ahmed et al. 2011). These numbers are not directly comparable as the methods and the parameters used to perform the annotations are different, and this may have an important impact on the sensibility and specificity of the detection. Indeed, analyses in *A. thaliana* have shown that there is a continuum between repetitive elements and unannotated genomic dark matter, making it somehow arbitrary to define a frontier (Maumus and Quesneville 2014). However, in spite of these limitations, there seem to be a direct relationship between genome size and percentage of TEs within the genome. Analyses of closely related species, for example of the *Oryza* genus (Chénais et al. 2012), suggest that TE activity and polyploidization are the two main mechanisms responsible for genome size increase during evolution (Panaud et al. 2014). The relationship between genome duplication and transposition is interesting. On the one hand, gene duplication can allow genomes to tolerate a higher TE activity, as their mutagenic capacity is buffered by having extra copies of essential genes, but on the other hand, the lack of gene duplications may force the genome to explore other sources of innovations such as transposition. In this respect, it is interesting to note that gymnosperms, that in contrast to angiosperms do not seem to have suffered recent whole-genome duplications, present extremely big genomes with a very high content of TEs (De La Torre et al. 2014).

The effect of TE activity in genome size may be quite dramatic over short periods of time, as suggested by the high activity of TEs associated to the genome size doubling of *Oryza australiensis,* a wild relative of rice, during the last three million years (Zhao and Ma 2013). However, although TEs may be responsible for rapid genome size changes, their activity is not constant during evolution. Indeed, TEs seem to alternate periods where they are relatively quiescent with burst of transposition where their copy number increases significantly (Vitte et al. 2014). This evolutionary behaviour of transposons as a whole can be in part explained by the results obtained analysing the regulation of particular transposons and genomes. All the data accumulated so far indicate that transposons are heavily silenced in genomes by different mechanisms, and in particular by epigenetic mechanisms

(Ito and Kakutani 2014). Silent TEs of different classes, including both DNA transposons and retrotransposons, can be reactivated in mutated genetic backgrounds showing reduced DNA methylation (Ito and Kakutani 2014), which shows that the silenced TEs retain their capacity to be activated. In fact, TEs can be activated in wild-type plants in particular situations or developmental stages. TEs are de-repressed in the gametophytes and their expression may allow the production of sRNAs to ensure the maintenance of the epigenetic silencing of TEs in the following generation, although alternative explanations of this phenomenon are also possible (Martínez and Slotkin 2012). In addition, over the years, data have accumulated on the stress-related activation of different TEs. This includes the well-studied activation of the tobacco retrotransposon *Tnt1* by biotic and abiotic stresses (Grandbastien et al. 2005), the cold and salt activation of the rice MITE *mPing* (Naito et al. 2009) and the heat activation of the Arabidopsis *ONSEN* retrotransposons (Cavrak et al. 2014). Similarly, it is known that in vitro culture, which can be considered as a complex stress, can reactivate TEs in rice and maize (Hirochika 1997; Kaeppler et al. 2000). Plants are subjected to stress in nature, and this may lead to reactivation of TEs in certain cells. In most cases, the somatic activation of TEs will not lead to germinal transpositions and therefore will not be inherited by the successive generations. However, in particular situations, a general release of the control mechanism may lead to a general activation of TEs leading to a burst of transposition. It is interesting to note that it has been shown that interspecific crosses or polyploidization events may lead to global epigenetic changes and activation of TEs (Parisod et al. 2009; Yaakov and Kashkush 2011). As these phenomena are commonplace in plant evolution, this may give the opportunity to TE amplification bursts to occur and accompany speciation events.

6.2 Transposable Elements in Genome Structure

TEs are usually not homogeneously distributed along chromosomes. They concentrate in pericentromeric regions, while they are less abundant in chromosome arms, in a pattern that is usually complementary to that of genes. These pattern of TEs can be the consequence of both a preferential insertion into these regions, as demonstrated for yeast retroelements, or the effect of selection cleaning up the more frequently deleterious TE insertions in gene-rich regions (Neumann et al. 2011; Peterson-Burch et al. 2004). Selection against insertion within genes, which are not homogeneously distributed along chromosomes, and the recombination rate, which is also different in different chromosomal regions and greatly influences TE elimination, explains in part the distribution of TEs (Bennetzen and Wang 2014). However, it has been shown that some TEs indeed have a preferential insertion into certain genomic regions. In general, *Copia*-like TEs show some preference for gene-rich regions, whereas *Gypsy*-like TEs are supposed to target preferentially the heterochromatic pericentromeric regions (Peterson-Burch et al. 2004). As an example, the tobacco *Tnt1* and the rice *Tos17 Copia* elements preferentially insert

into gene-rich regions (Miyao et al. 2003; Le et al. 2007), whereas in cereals, there are some families of *Gypsy* retrotransposons that are almost exclusively located in the centromeres, suggesting a high preference for insertion into these regions (Gao et al. 2009; Wolfgruber et al. 2009; Langdon et al. 2000; Li et al. 2013; Jiang et al. 2003). However, there are exceptions to this rule, and some *Gypsy* elements such as the low-copy-number *LORE1* retrotransposon from *Lotus japonicus* seem to target gene-rich regions (Madsen et al. 2005) and some *Copia*-like retrotransposons such as the *Tal1* element from *Arabidopsis lyrata* target the centromere for integration (Tsukahara et al. 2012).

The fact that TEs, and in particular high-copy-number retrotransposons, tend to concentrate in gene-poor heterochromatic regions, does not imply that they do not impact on genome function. Indeed, TE insertions in the pericentromeric regions probably have a profound impact on the structure and dynamics of genomes. The main mechanism to control the activity of TEs is their epigenetic silencing. As a consequence of their silencing, TE sequences tend to be heavily methylated and are associated with expression-repressive histone modifications (Ito and Kakutani 2014). Therefore, the concentration of TEs in the centromere also concentrates certain epigenetic marks in these regions, leading to a particular chromatin structure that is essential for heterochromatin compaction and function in the centromeres (Wong and Choo 2004). It has been proposed that TEs, and in particular LTR retrotransposons sitting in the centromere, may transcribe flanking repeats and other centromeric sequences leading to the production of double-stranded RNA which would direct their particular heterochromatic structure (Lippman et al. 2004). In fact, studies on the formation of neocentromeres have shown that it is the epigenetic nature of centromere elements, and not their sequence, which ensures its functionality (Zhang et al. 2013). Therefore, there is probably a dynamic interplay between retrotransposons and heterochromatin where some TEs target heterochromatin for integration (in the case of *Gypsy*-like elements through the chromodomains of their integrases that are known to interact with some heterochromatic epigenetic marks) and help thereafter to maintain heterochromatin by directing their epigenetic modification (Gao et al. 2008).

6.3 Transposable Elements as a Source of New Functions

TEs impact on genome and gene evolution in many ways. Perhaps, the most obvious is the generation of null mutations by transposing into a gene. Some of these null mutations have been selected by humans during plant domestication such as the waxy and sticky varieties of foxtail millet (*Setaria italica*), or Mendel's wrinkled peas (Lisch 2013). For TEs that transpose by a cut-and-paste mechanism (e.g. most class II TEs), the excision of the element may result in function recovery giving rise to mosaic phenotypes as exemplified by the kernel colour of maize cobs. Nevertheless, in some cases the excision may leave behind parts of the element that are not removed and can modify the coding capacity of the gene, and in some cases

provide new gene functions (Lisch 2013; Oliver et al. 2013). This process by which a TE, or a part of it, is established in a specific region and gains a cellular function is known as molecular domestication (Kajihara et al. 2012).

There is an important number of plant genes with a transposon origin (Oliver et al. 2013; Bennetzen and Wang 2014). In particular, several important transcription factors derive from class II transposases. For example, *Daysleeper*, a transcription factor that regulates the morphogenetic development in *A. thaliana,* is derived from a *hAT* transposase (Bundock and Hooykaas 2005), or the light response FHY3 and FAR1 transcription factors that are ancient *Mutator* transposases (Hudson et al. 2003; Lin et al. 2007).

Transposons can also capture, duplicate and mobilize genes or gene fragments, creating new opportunities for gene evolution. Retrotransposons duplicate host genes or gene fragments through the reverse transcription of their mRNAs generating what is called a retrogene. The retroposed gene fragments can be fused to host genes to generate new chimeric proteins (Elrouby and Bureau 2010), and retroposed retrogenes can be regulated differently to the original genes (Abdelsamad and Pecinka 2014), which can be a source of gene innovation. Class II transposons can also transduplicate genes. Pack-MULEs, for example, are *Mutator*-like TEs that carry fragments of genes in different plants and were proposed as important mediators of gene evolution in plants (Jiang et al. 2004). The fact that an important fraction of rice Pack-MULEs is transcribed and show signs of purifying selection suggested that indeed these elements have a role in gene evolution in plants (Hanada et al. 2009). A part from MULEs, other class II TEs, such as *CACTA* elements, have been shown to transduplicate host gene fragments in different plants (Benjak et al. 2008; Morgante 2006). But probably the TEs that seem to capture more actively, amplify and mobilize gene fragments are the rolling-circle transposing elements *Helitrons*. More than one-third of the thousands *Helitrons* of maize genome carry at least one host gene fragment (Du et al. 2009). Therefore, TEs have a great potential to generate new gene structures by shuffling host genome sequences (Bennetzen 2005; Morgante 2006).

6.4 Impact of Transposable Elements in Gene Regulation

In addition to their effect on the coding capacity of the host genome, TEs can impact on host genes in many ways. As already explained, the expression of TEs is tightly regulated, both because they are the main target of the silencing machinery and also because they usually have stress-related promoters that are only active under particular situations. For this reason, in addition to being able to modify host gene expression by interrupting gene regulatory regions upon insertion, for example in the case of the *Vgt1* regulatory locus of maize (Salvi et al. 2007), TEs can modify the expression of host genes located nearby by contributing their own regulatory elements or by attracting the silencing machinery.

There are several examples of insertions of TEs that induce new transcriptional regulations to host genes. This is the case of the insertion of a *Hopscotch* TE some 50 Kb upstream of the *theosinte branched* 1 (*tb1*) gene, which represses branching in maize, which results in its overexpression and the apical dominant phenotype of modern maize (Studer et al. 2011) or the insertion of an LTR retrotransposon upstream of the *Ruby* gene in oranges which confers to this gene a developmental regulation and cold inducibility resulting in the blood orange phenotype (Butelli et al. 2012).

MITEs are a particular type of transposons present in high copy numbers in plant genomes (Casacuberta and Santiago 2003). They are relatively small, which may help them avoiding to generate complete knockout phenotypes, and although they do not need to be expressed to transpose, they can contain transcriptional regulatory sequences. For example the rice *mPing* MITE contains stress-responsible transcriptional regulatory elements that upregulate neighbouring genes under cold and salt stress conditions (Yasuda et al. 2013; Naito et al. 2009). The high copy number of MITEs makes them particularly suited to modify the expression of groups of genes, making it possible to create, or to extend, transcriptional regulatory networks. The fact that some transcription factors derive from transposases (see above), and that the sequences bound by transposases (e.g. the TIRs) can be mobilized throughout the genome, was proposed as a potential mechanisms to create and modify transcriptional regulatory networks (Feschotte 2008). In the recent years, evidences that TEs can mobilize transcription factor binding sites and rewire transcriptional networks have accumulated (Rebollo et al. 2012). In plants, a recent report from our laboratory has shown that different families of MITEs have amplified and redistributed the binding sites for the E2F transcription factor during *Brassica* evolution, and the insertion of some of these MITEs may have wired new genes into the E2F transcriptional network (Hénaff et al. 2014).

In spite of the examples explained above that illustrate the potential of TEs to bring new regulatory sequences to host genes, the most frequent effect of a TE insertion within or close a gene promoter is its inactivation. As already explained, TEs are controlled by epigenetic mechanisms that silence them tightly. For this reason, most TEs are heavily methylated and are associated to inactive chromatin, and this can influence genes located nearby that can become silenced by the presence of the TE. A well-studied example of such an effect is the epigenetic silencing of a sex determination gene in melon linked to a TE insertion in its upstream region (Martin et al. 2009). Similarly, the necessary repression of the flowering regulator *FWA* gene in *A. thaliana* is a consequence of the epigenetic silencing of a SINE transposon located in its promoter (Kinoshita et al. 2007). Genome-wide analyses suggest that these effects may be highly relevant. As an example, it has been shown that about 300 genes differentially expressed in maize populations have changes in DNA methylation, and many of these regions are associated with transposons (Eichten et al. 2013). This suggests that polymorphic TE insertions modify the pattern of genome methylation which translates into changes in gene expression.

Silencing of TEs is mediated by siRNAs that target TE sequences which probably originate from the expression of particular TE structures (e.g. inverse repeated elements). Whereas the main target of these siRNAs are TEs, in some cases TEs may produce siRNAs that target host genes (Bennetzen and Wang 2014; McCue and Slotkin 2012). In fact, it has been proposed that TE can be the source of both siRNAs and miRNAs (Li et al. 2011; Piriyapongsa and Jordan 2008) which suggests that the genome has evolved a new layer of gene regulation from its defence mechanisms against TEs.

The expression of TEs may also interfere with host genes creating sense or antisense transcripts that may result in their specific silencing. It has been shown that read-through transcription, due to a leaky transcriptional terminator, is relatively frequent in plant retrotransposons, and this could result in the inclusion of flanking sequences into retrotransposon transcripts. As a consequence, as it has been shown in tobacco (Hernández-Pinzón et al. 2009), the convergent transcription of a retrotransposon located downstream of a host gene could result in the formation of dsRNAs which may potentially regulate the host gene. In addition, TEs insertions in 5' leader region, 3' trailer sequence or introns can modify the sites of RNA processing or polyadenylation affecting gene expression (Bennetzen and Wang 2014).

6.5 Transposable Elements Dynamics and Evolution of Crop Plants

We have seen in the previous sections that TEs can impact on genomes in many ways, from providing new genes or modifying the existing ones or alter their expression, to modify genome or chromosome structure. Because of that TEs are an extraordinary source of novelty useful for evolution (Lisch 2013). In particular, in the last few years, a number of examples of TE insertions leading to important agronomic traits that have been selected during evolution and breeding have accumulated (Lisch 2013). These include the different flesh fruit colour in blood orange (Butelli et al. 2012), the different skin colours in grapevine (This et al. 2007), the nectarine phenotype in peaches (Vendramin et al. 2014) or the seedless phenotype in apples (Yao et al. 2001) (see Fig. 6.1). However, evaluating the impact of TEs in the evolution of eukaryote genomes is not an easy task. In spite of the examples listed above on TEs that gave rise to mutations that have been selected during evolution, a general evaluation is still lacking. There are several reasons for that, as previously pointed out (Vitte et al. 2014). Although the number of plant genomes sequenced is growing rapidly, the quality of the published genomes is not always good enough to allow a proper analysis of the TE content. Indeed, most published genomes contain a variable, and usually important, fraction of unassembled reads which are usually enriched in repetitive sequences including TEs. This precludes a complete genome-wide TE analysis. In addition to the quality of

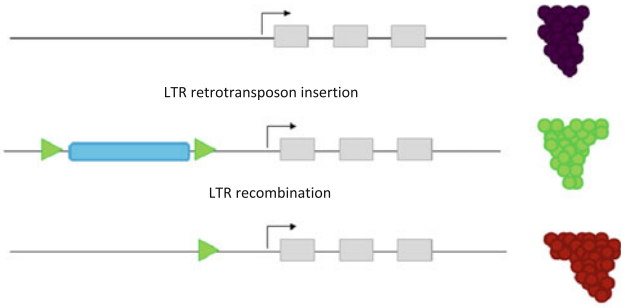
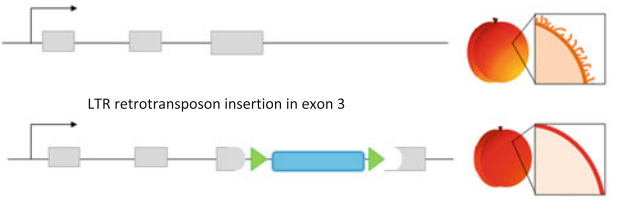
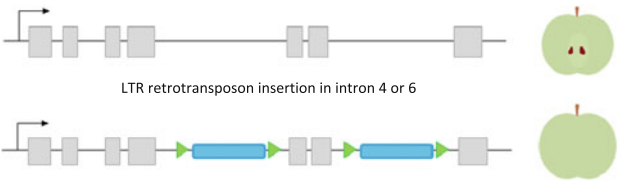
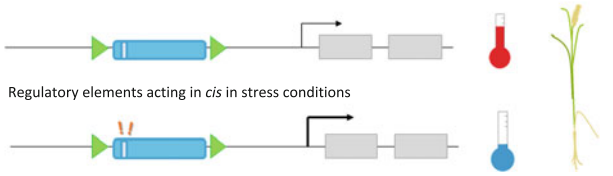
TE related traits in crops	Representation
TE insertion in promoter sequence confers different grape skin colours. (This et al. 2007)	
Peach and nectarine skin phenotype due to a *Copia* retrotransposon insertion in exon sequence. (Vendramin et al. 2014)	
Seedless phenotype in some apple varieties is due to a retrotransposon insertion in an intron. (Yao et al. 2001)	
A TE insertion in regulatory region confers cold specificity regulation in rice. (Yasuda et al. 2013)	

Fig. 6.1 Representation of different important agronomic traits that are due to transposable element insertions. *Grey boxes* represent exons, *blue boxes* represent TE coding region, and *green triangles* represent LTRs

the sequence and assembly, the annotation of the TE content is also highly variable among the sequenced genomes. There are several reasons for that, including the use of different bioinformatics tools and pipelines as well as the thresholds set which determine the sensitivity and specificity of the annotation tools. This makes comparisons of the TE content between genomes a very difficult exercise, and different voices claim that there is a need for an international effort to standardize the methods used for annotating TEs (Hoen, Bureau, Bourke and Blanchette, in

preparation). But even with good genome sequences and TE annotation, reference genomes are only a snapshot, a fixed image, of a genome and analysing the impact of TEs in genome evolution will require sequence variability analysis within a species or among different related species. In the last few years, an important amount of resequencing data of crop varieties and landraces has being accumulated. As an example, 3000 rice varieties have already been sequenced and offer an unprecedented opportunity to search for the genetic bases of a wide range of phenotypic differences (Li et al. 2014). However, in most cases, the analyses of variability are restricted to SNPs, and TE insertion polymorphisms are not analysed. The reason for that is that detecting TE polymorphisms, and in particular TE insertions with respect to the reference genome is far from trivial. There are a number of recent tools that allow detecting TE insertion polymorphisms using paired-end resequencing data, including TEA (Lee et al. 2012), RetroSeq (Keane et al. 2013), VariationHunter (Hormozdiari et al. 2010), TEMP (Zhuang et al. 2014) and Jitterbug (Hénaff et al. submitted), but they are only starting to be used to determine the role of TEs in plant genome evolution (see for example Sanseverino et al., submitted). The use of these tools on the growing amount of resequencing data on plant varieties and accessions will probably allow us in the next future to have a more global and complete view of the impact of TEs in plant genome evolution. In particular, the analysis of crop genomes and the comparison of crop reference genomes with that of, on the one hand, their wild ancestors, and on the other hand, domesticated landraces or elite varieties will shed light on the role of TEs on the evolution of plant genomes during domestication and breeding. In addition, as crop domestication is an excellent model to study genome evolution at large, as it has already been said (Olsen and Wendel 2013), these analyses will probably allow us to better understand the structure and evolution of plant genomes and the key role played by TEs, who once were called junk DNA and now are rediscovered as key factors for genetic innovation.

References

Abdelsamad A, Pecinka A (2014) Pollen-specific activation of Arabidopsis retrogenes is associated with global transcriptional reprogramming. Plant Cell 26:3299–3313. doi:10.1105/tpc.114.126011

Ahmed I, Sarazin A, Bowler C et al (2011) Genome-wide evidence for local DNA methylation spreading from small RNA-targeted sequences in Arabidopsis. Nucleic Acids Res 39:6919–6931. doi:10.1093/nar/gkr324

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

Bennetzen JL (2005) Transposable elements, gene creation and genome rearrangement in flowering plants. Curr Opin Genet Dev 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. Annu Rev Plant Biol 65:505–530. doi:10.1146/annurev-arplant-050213-035811

Bundock P, Hooykaas P (2005) An Arabidopsis hAT-like transposase is essential for plant development. Nature 436:282–284. doi:10.1038/nature03667

Butelli E, Licciardello C, Zhang Y et al (2012) Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. Plant Cell 24:1242–1255. doi:10.1105/tpc.111.095232

Casacuberta JM, Santiago N (2003) Plant LTR-retrotransposons and MITEs: control of transposition and impact on the evolution of plant genes and genomes. Gene 311:1–11. doi:10.1016/S0378-1119(03)00557-2

Cavrak VV, Lettner N, Jamge S et al (2014) How a retrotransposon exploits the plant's heat stress response for its activation. PLoS Genet 10:e1004115. doi:10.1371/journal.pgen.1004115

Chénais B, Caruso A, Hiard S, Casse N (2012) The impact of transposable elements on eukaryotic genomes: from genome size increase to genetic adaptation to stressful environments. Gene 509:7–15. doi:10.1016/j.gene.2012.07.042

De La Torre AR, Birol I, Bousquet J et al (2014) Insights into conifer giga-genomes. Plant Physiol 166:1724–1732. doi:10.1104/pp.114.248708

Du C, Fefelova N, Caronna J et al (2009) The polychromatic Helitron landscape of the maize genome. Proc Natl Acad Sci U S A 106:19916–19921. doi:10.1073/pnas.0904742106

Eichten SR, Briskine R, Song J et al (2013) Epigenetic and genetic influences on DNA methylation variation in maize populations. Plant Cell 25:2783–2797. doi:10.1105/tpc.113.114793

Elrouby N, Bureau TE (2010) Bs1, a new chimeric gene formed by retrotransposon-mediated exon shuffling in maize. Plant Physiol 153:1413–1424. doi:10.1104/pp.110.157420

Feschotte C (2008) Transposable elements and the evolution of regulatory networks. Nat Rev Genet 9:397–405. doi:10.1038/nrg2337

Gao D, Gill N, Kim H-R et al (2009) A lineage-specific centromere retrotransposon in Oryza brachyantha. Plant J 60:820–831. doi:10.1111/j.1365-313X.2009.04005.x

Gao X, Hou Y, Ebina H, et al (2008) Chromodomains direct integration of retrotransposons to heterochromatin. Genome Res 359–369. doi:10.1101/gr.7146408.1

Grandbastien M, Audeon C, Bonnivard E et al (2005) Stress activation and genomic impact of Tnt1 retrotransposons in Solanaceae. Cytogenet Genome Res 110:229–241. doi:10.1159/000084957

Hanada K, Vallejo V, Nobuta K et al (2009) The functional role of pack-MULEs in rice inferred from purifying selection and expression profile. Plant Cell 21:25–38. doi:10.1105/tpc.108.063206

Hénaff E, Vives C, Desvoyes B et al (2014) Extensive amplification of the E2F transcription factor binding sites by transposons during evolution of Brassica species. Plant J 77:852–862. doi:10.1111/tpj.12434

Hénaff E, Zapata L, Casacuberta JM, Ossowski S. (Submitted) Jitterbug: somatic and germline transposon insertion detection at single-nucleotide resolution

Hernández-Pinzón I, de Jesús E, Santiago N, Casacuberta JM (2009) The frequent transcriptional readthrough of the tobacco Tnt1 retrotransposon and its possible implications for the control of resistance genes. J Mol Evol 68:269–278. doi:10.1007/s00239-009-9204-y

Hirochika H (1997) Retrotransposons of rice: their regulation and use for genome analysis. Plant Mol Biol 35:231–240

Hormozdiari F, Hajirasouliha I, Dao P et al (2010) Next-generation VariationHunter: combinatorial algorithms for transposon insertion discovery. Bioinformatics 26:i350–i357. doi:10.1093/bioinformatics/btq216

Hudson M, Lisch D, Quail P (2003) The FHY3 and FAR1 genes encode transposase-related proteins involved in regulation of gene expression by the phytochrome A-signaling pathway. Plant J 453–471

Ito H, Kakutani T (2014) Control of transposable elements in Arabidopsis thaliana. Chromosome Res 22:217–223. doi:10.1007/s10577-014-9417-9

Jiang J, Birchler J a, Parrott W a, Kelly Dawe R (2003) A molecular view of plant centromeres. Trends Plant Sci 8:570–575. doi:10.1016/j.tplants.2003.10.011

Jiang N, Bao Z, Zhang X et al (2004) Pack-MULE transposable elements mediate gene evolution in plants. Nature 431:569–573. doi:10.1038/nature02945.1

Kaeppler S, Kaeppler H, Rhee Y (2000) Epigenetic aspects of somaclonal variation in plants. Plant Mol Biol 179–188

Kajihara D, Godoy F, Hamaji T et al (2012) Functional characterization of sugarcane mustang domesticated transposases and comparative diversity in sugarcane, rice, maize and sorghum. Mol Biol 639:632–639. doi:10.1590/S1415-47572012005000038

Keane TM, Wong K, Adams DJ (2013) RetroSeq: transposable element discovery from next-generation sequencing data. Bioinformatics 29:389–390. doi:10.1093/bioinformatics/bts697

Kinoshita Y, Saze H, Kinoshita T et al (2007) Control of FWA gene silencing in Arabidopsis thaliana by SINE-related direct repeats. Plant J 49:38–45. doi:10.1111/j.1365-313X.2006.02936.x

Langdon T, Seago C, Mende M et al (2000) Retrotransposon evolution in diverse plant genomes. 156(1):313–325

Le QH, Melayah D, Bonnivard E et al (2007) Distribution dynamics of the Tnt1 retrotransposon in tobacco. Mol Genet Genomics 278:639–651. doi:10.1007/s00438-007-0281-6

Lee E, Iskow R, Yang L, Gokcumen O (2012) Landscape of somatic retrotransposition in human cancers. Science 337:967–971. doi:10.1126/science.1222077.Landscape

Li B, Choulet F, Heng Y et al (2013) Wheat centromeric retrotransposons: the new ones take a major role in centromeric structure. Plant J 73:952–965. doi:10.1111/tpj.12086

Li J-Y, Wang J, Zeigler RS (2014) The 3,000 rice genomes project: new opportunities and challenges for future rice research. Gigascience 3:8. doi:10.1186/2047-217X-3-8

Li Y, Li C, Xia J, Jin Y (2011) Domestication of transposable elements into MicroRNA genes in plants. PLoS ONE 6:e19212. doi:10.1371/journal.pone.0019212

Lin R, Ding L, Casola C, Ripoll D (2007) Transposase-derived transcription factors regulate light signaling in Arabidopsis. Science 318:1302–1305

Lippman Z, Gendrel A, Black M (2004) Role of transposable elements in heterochromatin and epigenetic control. Nature. doi:10.1038/nature02724.1

Lisch D (2013) How important are transposons for plant evolution? Nat Rev Genet 14:49–61. doi:10.1038/nrg3374

Madsen LH, Fukai E, Radutoiu S et al (2005) LORE1, an active low-copy-number TY3-gypsy retrotransposon family in the model legume Lotus japonicus. Plant J 44:372–381. doi:10.1111/j.1365-313X.2005.02534.x

Martin A, Troadec C, Boualem A et al (2009) A transposon-induced epigenetic change leads to sex determination in melon. Nature 461:1135–1138. doi:10.1038/nature08498

Martínez G, Slotkin RK (2012) Developmental relaxation of transposable element silencing in plants: functional or byproduct? Curr Opin Plant Biol 15:496–502. doi:10.1016/j.pbi.2012.09.001

Maumus F, Quesneville H (2014) Deep investigation of Arabidopsis thaliana junk DNA reveals a continuum between repetitive elements and genomic dark matter. PLoS ONE 9:e94101. doi:10.1371/journal.pone.0094101

McCue AD, Slotkin RK (2012) Transposable element small RNAs as regulators of gene expression. Trends Genet 28:616–623. doi:10.1016/j.tig.2012.09.001

Miyao A, Tanaka K, Murata K et al (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. doi:10.1105/tpc.012559.ements

Morgante M (2006) Plant genome organisation and diversity: the year of the junk! Curr Opin Biotechnol 17:168–173. doi:10.1016/j.copbio.2006.03.001

Naito K, Zhang F, Tsukiyama T et al (2009) Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. Nature 461:1130–1134. doi:10.1038/nature08479

Neumann P, Navrátilová A, Koblížková A et al (2011) Plant centromeric retrotransposons: a structural and cytogenetic perspective. Mob DNA 2:4. doi:10.1186/1759-8753-2-4

Nystedt B, Street NR, Wetterbom A et al (2013) The Norway spruce genome sequence and conifer genome evolution. Nature 497:579–584. doi:10.1038/nature12211

Oliver KR, McComb JA, Greene WK (2013) Transposable elements: powerful contributors to angiosperm evolution and diversity. Genome Biol Evol 5:1886–1901. doi:10.1093/gbe/evt141

Olsen KM, Wendel JF (2013) A bountiful harvest: genomic insights into crop domestication phenotypes. Annu Rev Plant Biol 64:47–70. doi:10.1146/annurev-arplant-050312-120048

Panaud O, Jackson S, Wendel J (2014) Drivers and dynamics of diversity in plant genomes. New Phytol 202:15–18. doi:10.1111/nph.12633

Parisod C, Salmon A, Zerjal T et al (2009) Rapid structural and epigenetic reorganization near transposable elements in hybrid and allopolyploid genomes in Spartina. New Phytol 184 (4):1003–1015. doi:10.1111/j.1469-8137.2009.03029.x

Peterson-Burch B, Nettleton D (2004) Voytas D (2004) Genomic neighborhoods for Arabidopsis retrotransposons: a role for targeted integration in the distribution of the Metaviridae. Genome Biol 5:R78

Piriyapongsa J, Jordan I (2008) Dual coding of siRNAs and miRNAs by plant transposable elements. Rna 814–821. doi:10.1261/rna.916708.ferred

Rebollo R, Romanish MT, Mager DL (2012) Transposable elements: an abundant and natural source of regulatory sequences for host genes. Annu Rev Genet 46:21–42. doi:10.1146/annurev-genet-110711-155621

Salvi S, Sponza G, Morgante M et al (2007) Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. Proc Natl Acad Sci U S A 104:11376–11381

Sanseverino W, Hénaff E, Vives C, et al. (submitted) The contribution of transposon insertion polymorphisms and nucleotide variability to the evolution of the melon genome

Studer A, Zhao Q, Ross-Ibarra J, Doebley J (2011) Identification of a functional transposon insertion in the maize domestication gene tb1. Nat Genet 43:1160–1163. doi:10.1038/ng.942.Identification

This P, Lacombe T, Cadle-Davidson M, Owens CL (2007) Wine grape (Vitis vinifera L.) color associates with allelic variation in the domestication gene VvmybA1. Theor Appl Genet 114:723–730. doi:10.1007/s00122-006-0472-2

Tsukahara S, Kawabe A, Kobayashi A (2012) Centromere-targeted de novo integrations of an LTR retrotransposon of Arabidopsis lyrata. Genes Dev 26:705–713. doi:10.1101/gad.183871.111 Epub 2012 Mar 19

Vendramin E, Pea G, Dondini L et al (2014) A unique mutation in a MYB gene cosegregates with the nectarine phenotype in peach. PLoS ONE 9:e90574. doi:10.1371/journal.pone.0090574

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

Wicker T, Sabot F, Hua-van A et al (2007) A unified classification system for eukaryotic transposable elements. Nat Rev Genet 8:973–982

Wolfgruber TK, Sharma A, Schneider KL et al (2009) Maize centromere structure and evolution: sequence analysis of centromeres 2 and 5 reveals dynamic Loci shaped primarily by retrotransposons. PLoS Genet 5:e1000743. doi:10.1371/journal.pgen.1000743

Wong LH, Choo KHA (2004) Evolutionary dynamics of transposable elements at the centromere. Trends Genet 20:611–616. doi:10.1016/j.tig.2004.09.011

Yaakov B, Kashkush K (2011) Massive alterations of the methylation patterns around DNA transposons in the first four generations of a newly formed wheat allohexaploid. Genome 54:42–49. doi:10.1139/G10-091

Yao J, Dong Y, Morris B (2001) Parthenocarpic apple fruit production conferred by transposon insertion mutations in a MADS-box transcription factor. Proc Natl Acad Sci U S A 98:1306–1311

Yasuda K, Ito M, Sugita T et al (2013) Utilization of transposable element mPing as a novel genetic tool for modification of the stress response in rice. Mol Breed 32:505–516. doi:10.1007/s11032-013-9885-1

Zhang B, Lv Z, Pang J et al (2013) Formation of a functional maize centromere after loss of centromeric sequences and gain of ectopic sequences. Plant Cell 25:1979–1989. doi:10.1105/tpc.113.110015

Zhao M, Ma J (2013) Co-evolution of plant LTR-retrotransposons and their host genomes. Protein Cell 4:493–501. doi:10.1007/s13238-013-3037-6

Zhuang J, Wang J, Theurkauf W, Weng Z (2014) TEMP: a computational method for analyzing transposable element polymorphism in populations. Nucleic Acids Res 42:6826–6838. doi:10.1093/nar/gku323