Metadata of the chapter that will be visualized online

Series Title		
Chapter Title	Plant Transposable Elements: Biology and Evolution	
Chapter SubTitle		
Copyright Year	2012	
Copyright Holder	Springer-Verlag Wien	

Corresponding Author	Family Name	Kejnovsky
	Particle	
	Given Name	**Eduard**
	Suffix	
	Division	Institute of Biophysics
	Organization	ASCR
	Address	Kralovopolska 135, 612 00, Brno, Czech Republic
	Email	kejnovsk@ibp.cz
Author	Family Name	Hawkins
	Particle	
	Given Name	**Jennifer S.**
	Suffix	
	Division	Department of Biology
	Organization	West Virginia University
	Address	53 Campus Drive, 26506, Morgantown, WV, USA
	Email	jhawkins@uga.edu
Author	Family Name	Feschotte
	Particle	
	Given Name	**Cédric**
	Suffix	
	Division	Department of Biology
	Organization	University of Texas
	Address	76019, Arlington, TX, USA
	Email	cedric@uta.edu

Abstract	Beginning with the pioneering work in the 30s and 40s of Barbara McClintock, R.A. Brink, Rollins Emerson, Marcus Rhoades, and other prominent maize geneticists, transposable elements (TEs) have come to occupy a central position in the study of plant genomes. Not only did McClintock's discovery of the *Activator*/*Dissociation* (*Ac*/*Ds*) system of maize change forever our appreciation of the dynamic nature of chromosomes, her seminal characterization of the regulatory influence of 'controlling elements' (such as *Ac*/*Ds* and later the *Enhancer*/*Suppressor-Mutator* (*En*/*Spm*) system) on adjacent gene expression paved the way for decades of exciting research on the control, both genetic and epigenetic, of gene regulation in plants and other eukaryotes.

Plant Transposable Elements: Biology and Evolution 2

Eduard Kejnovsky, Jennifer S. Hawkins, and Cédric Feschotte

Contents

2.1 Introduction

Beginning with the pioneering work in the 30s and 40s of Barbara McClintock, R.A. Brink, Rollins Emerson, Marcus Rhoades, and other prominent maize geneticists, transposable elements (TEs) have come to occupy a central position in the study of plant genomes. Not only did McClintock's discovery of the *Activator/Dissociation* (*Ac/Ds*) system of maize change forever our appreciation of the dynamic nature of chromosomes, her seminal characterization of the regulatory influence of 'controlling elements' (such as *Ac/Ds* and later the *Enhancer/Suppressor-Mutator* (*En/Spm*) system) on adjacent gene expression paved the way for decades of exciting research on the control, both genetic and epigenetic, of gene regulation in plants and other eukaryotes.

It took four decades after McClintock's groundbreaking discoveries and the rise of recombinant DNA technology for the first TEs to be cloned and sequenced in the 1980s. One of the surprises from these early molecular studies was the striking similarity in structure, genetic organization, and even sometimes nucleotide sequence, among the first TEs characterized in maize, snapdragon, *Drosophila* and bacteria (Green 1980; Fedoroff et al. 1983; Levis et al. 1984; Saedler et al. 1984). At that time, and over the next two decades, the biology of TEs was assessed primarily on the basis of the mutations they engendered. Myriad mutant alleles caused by insertions and/or rearrangements of transposons were collected by geneticists in the field, the greenhouse and the fly room, and meticulously analyzed at the molecular level in the lab. Although this era furnished many crucial insights regarding the mechanistic underpinnings and mutagenic capabilities of transposition (for review, Berg and Howe 1989), it yielded little information regarding the abundance and diversity of TEs, much less the long-term evolutionary impact of TE activity.

The advent of large-scale DNA sequencing over the last two decades, combined with advances in functional

E. Kejnovsky (✉)
AU1 Institute of Biophysics, ASCR, Kralovopolska 135, 612 00 Brno, Czech Republic
e-mail: kejnovsk@ibp.cz

J.F. Wendel et al. (eds.), *Plant Genome Diversity Volume 1*,
DOI 10.1007/978-3-7091-1130-7_2, © Springer-Verlag/Wien 2012

genomics and bioinformatics, has transformed the study of TE biology. This "genomics revolution" has resulted in a greater understanding of the many ways that TEs influence the function and evolution of genes and genomes, and consequently, their host organisms. In particular the genomics era has revealed that, although only a tiny fraction of TEs are transpositionally active, most eukaryotic genomes, and especially plant genomes, are packed with a plethora of seemingly dormant or inactivated TE families (Feschotte et al. 2002). Given the inherent mutagenic potential of active transposition, it should come as no surprise that the majority of these TEs are either defective, fossilized copies or potentially active copies that are restrained by host silencing systems; however, active transposition, as evidenced by instances of mutagenic (yet potentially evolutionarily significant) insertions, has been demonstrated. For example, TEs have been shown to silence or alter expression of genes adjacent to insertion sites, become integrated into functional genes as newly acquired exons (exapted), acquire host gene sequences and insert them into new genomic locations, contribute to chromosomal rearrangements via recombination, epigenetically alter regional methylation patterns, and provide template sequences for RNA interference (Feschotte et al. 2002; Bennetzen 2005; Morgante et al. 2007; Weil and Martienssen 2008; and see Slotkin et al. 2012, this volume). This diverse functional impact of TEs, and their intrinsic contribution to genomic plasticity, suggests that these elements play a major role in molecular diversification and, ultimately, species divergence.

In this chapter, we provide the reader with the fundamentals of TE biology, with an emphasis on plant elements. We begin with an overview of TE classification and transposition mechanisms, followed by an examination of the extensive variability in both inter- and intra-specific TE content across diverse plant taxa. Finally, we explore some of the general principles characterizing and influencing the genomic distribution, activity and evolution of TEs.

2.2 Transposable Element Classification

TEs can be broadly defined as DNA segments capable of chromosomal movement, either via replicative or conservative (cut-and-paste) mechanisms (discussed in more detail below). The TE classification system that we present here is similar to the one proposed by Wicker et al. (2007) and to the one implemented in Repbase, the most popular database of repetitive DNA sequences (http://www.girinst.org/). At the highest level, eukaryotic TEs comprise two major classes, and each class can be divided into subclasses based on their mechanism of chromosomal integration, which is reflective of the protein-coding capabilities and organizational structure of each class and subclass of elements (Figs. 2.1, 2.2). Class I elements, also known as retrotransposons, transpose via an RNA intermediate, which must be reverse transcribed prior to integration into the genome, while Class II elements transpose via a DNA intermediate (Finnegan 1989). Transposition of both classes of elements may result in a heritable increase in genomic copy number; hence, individual TE types are found in multiple copies (often referred to as a TE family) and comprise the majority of the repetitive fraction of eukaryotic genomes (e.g. Adams et al. 2000; The Arabidopsis Genome Initiative 2000; Lander et al. 2001; International Rice Genome Sequencing Project 2005). TEs have been found in virtually every organism studied to date (with few exceptions, such as *Plasmodium falciparum* and other Apicomplexa), although significant qualitative and quantitative variation abounds, even among closely related organisms (see below for a comparison among selected plant species).

The genomes of plants are saturated with many and diverse TEs, and continue to serve as excellent models to yield some of the most significant advances in the field of transposon biology. The vast majority of repetitive DNA in the nuclear genomes of plants is derived from the proliferation of TEs, most often Class I RNA elements (Fig. 2.1) (e.g. SanMiguel et al. 1996; Vicient et al. 1999; Hawkins et al. 2006; Neumann et al. 2006; Vitte and Bennetzen 2006). Two major subclasses of Class I elements have been identified in plants: (1) Long terminal repeat (LTR) retrotransposons, whose reverse-transcription and subsequent integration as double-stranded DNA is mediated by an element-encoded reverse transcriptase and integrase, respectively, (2) non-LTR retrotransposons (sometimes called retroposons), which include long and short interspersed elements (LINEs and SINEs) and use target-primed reverse transcription, a mechanism coupling reverse transcription and integration. DIRS-like elements (named after *Dictyostelium* intermediate repeat sequence) represent a third subclass of retrotransposons integrated through an element-encoded tyrosine recombinase. They are relatively common in animals and fungi, but have yet to be found in flowering plants. Class II elements have been identified in every plant genome that has been thoroughly examined, and these can be divided in two major subclasses: (1) classic 'cut-and-paste' DNA transposons, characterized by terminal inverted repeats (TIRs), which are excised and reintegrated as double-stranded DNA by the action of an element-encoded transposase and (2) *Helitrons*, or rolling-circle transposons, which most likely transpose via a replicative mechanism involving a single-stranded DNA intermediate and which encode recombinase with Replicator initiator motif (Rep) and DNA Helicase domains (Fig. 2.1).

In plants, Class I elements (particularly LTR retrotransposons) make up the largest fraction of the TE complement (SanMiguel et al. 1996, 1998; Vicient et al. 1999;

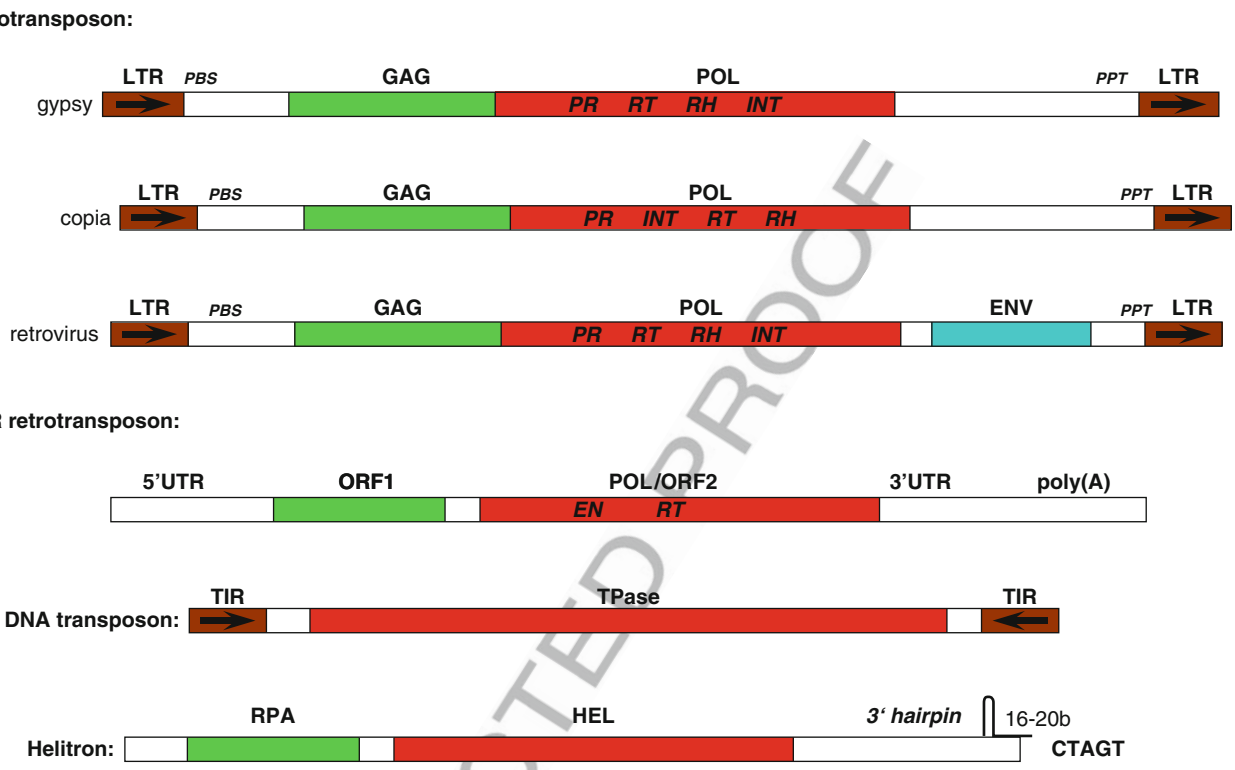

Fig. 2.1 Structure of main types of transposable elements. GAG and POL genes of LTR retrotransposons, ORF1 of non-LTR retrotransposons, transposase (TPase) of DNA transposons and replicative protein A (RPA) and helicase (HEL) of Helitrons are marked. Long terminal repeats (LTRs), primer-binding site (PBS) and polypurine tract (PPT) of LTR retrotransposons, 5′ UTR, 3′ UTR and poly(A) of non-LTR retrotransposons, terminal inverted repeats (TIR) of DNA transposons and 3′ hairpin of Helitrons are labeled. LTR retrotransposons are exemplified by *gypsy*, *copia* and retrovirus superfamilies. Protease (PR), reverse transcriptase (RT), RNaseH (RH), integrase (INT) and endonuclease (EN) domains are marked

Hawkins et al. 2006; Neumann et al. 2006; Vitte and Bennetzen 2006). The LTRs flanking a retrotransposon can range from just a few hundred base pairs to as much as 6 kb, and usually begin with 5′-TG-3′ and end with 5′-CA-3′. The LTR retrotransposons typically contain GAG and POL protein coding ORFs, which encode several enzymes (reverse transcriptase – RT; protease – PR; RNaseH – RH; integrase – INT) responsible for reverse transcription and integration of daughter sequences into new chromosomal locations. Two major superfamilies of LTR retrotransposons are found in plants, *gypsy*-like and *copia*-like (also known as *Metaviridae* and *Pseudoviridae*, respectively). Both types of LTR retrotransposons contain the same protein coding domains, but these are arranged in a different order. Their ancient origin is evidenced by the fact that they form deeply diverged monophyletic clades in phylogenetic analyses of reverse transcriptases (Eickbush and Malik 2002; Havecker et al. 2004). Non-LTR retrotransposons (LINEs and SINEs) are, as their name indicates, not flanked by LTRs, but complete LINEs can reach several thousand base pairs in length, contain coding sequences responsible for transposition, and often display a stretch of adenines or a simple sequence repeat at their 3′ end (Figs. 2.1, 2.2c).

Class II DNA elements are found in most eukaryotes, and despite their conservative transposition mechanism, have been capable of attaining relatively high copy numbers in some plants (see Sect. 2.3.1, Feschotte and Pritham 2007). Class II elements encode the machinery to facilitate their own transposition, usually in the form of a transposase (TPase) encoded by a single gene. "Cut-and-paste" transposition is associated with Subclass 1 DNA transposons, and occurs via TPase binding to the terminal inverted repeats (TIRs) of the element (Fig. 2.1), followed by excision and reintegration of the transposon at a new chromosomal location (Craig et al. 2002). The transposition mechanism of *Helitrons* has not been investigated in functional detail, but these elements are believed to employ a mechanism where only one DNA strand is cut, displaced and which serves as a template for replication of the element at a new locus (Kapitonov and Jurka 2007).

Both Class I and Class II TEs may be further divided into autonomous or non-autonomous elements dependent upon their ability to encode the enzymatic machinery responsible for movement. Non-autonomous elements may still be mobilized *in trans* if they retain the capacity to be recognized by the enzymes encoded by autonomous

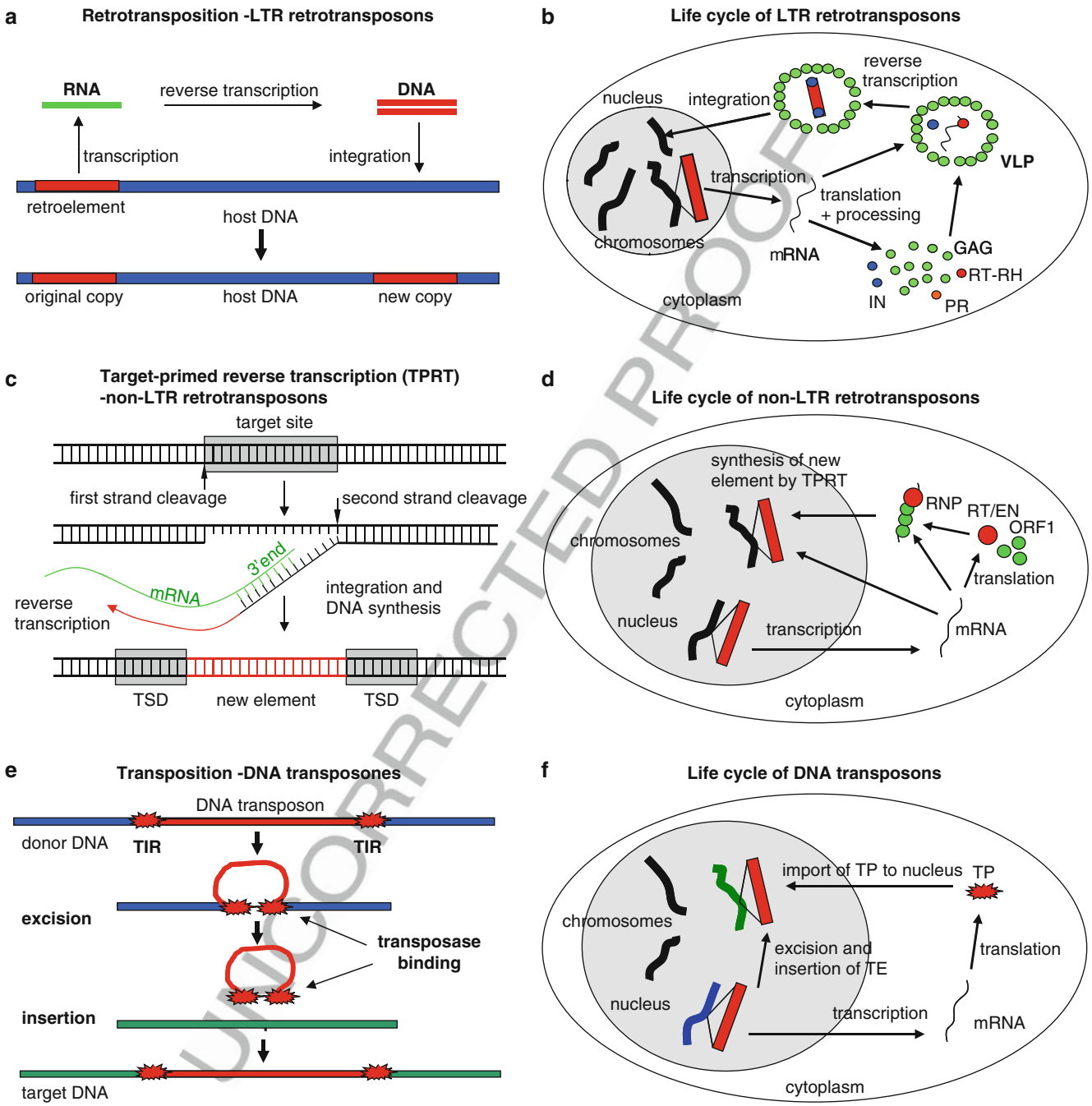

Fig. 2.2 Transpositional mechanism of main types of transposable elements. (**a**) Schematic retrotransposition of LTR retrotransposons and (**b**) their life cycle in the cell has a "copy and paste" character. Target-primed reverse transcription of non-LTR retrotransposons where cDNA is synthesized *in situ* (**c**) and the life cycle of non-LTR retrotransposons in the cell (**d**). Transposition of DNA transposons using "cut and paste" mode (**e**) and their life cycle in the cell (**f**). GAG gene, reverse transcriptase (RT), endonuclease (EN), integrase (INT), protease (PR) domains, transposase (TP), terminal inverted repeat (TIR), target-site duplication (TSD), ribonucleoparticle (RNP) and virus-like particle (VLP) are marked

elements located elsewhere in the genome. Although this concept was initially described for classic, two-component DNA transposon systems, such as *Ac/Ds* in maize, it seems that virtually all types of TEs may include both autonomous elements and non-autonomous counterparts that are movable *in trans* (Feschotte et al. 2002; Wicker et al. 2007). Non-autonomous Class I elements in plants include SINEs (short interspersed elements, Deragon and Zhang 2006), TRIMs (terminal repeat retrotransposons in miniature, Witte et al. 2001) and LARDs (large retrotransposon derivates, Kalendar et al. 2004). MITEs (miniature inverted-repeat elements, Bureau and Wessler 1992)

represent the most abundant type of non-autonomous DNA transposon in plant genomes thus far examined.

Non-autonomous elements may originate in a variety of ways. Most commonly, they derive from autonomous copies that have suffered mutations (substitutions or insertions/deletions) disabling their coding capabilities. For example, most *Ds* elements are directly derived from *Ac* by internal deletions (Yan et al. 1999). Note, however, that autonomous and non-autonomous elements need not share extensive sequence similarity to form a functional pair. Indeed, the original *Ds1* element from maize, which is recognized and mobilized by the *Ac*-encoded transposase, shares only the outermost 11 nucleotides of its TIRs (terminal inverted repeats) with *Ac* (Kunze and Starlinger 1989). Likewise, many families of high-copy number MITEs are not always directly related to autonomous elements present in the same genome. Nonetheless, there is evidence that some MITEs, such as *Stowaway*, can be mobilized with high efficiency by distantly related autonomous transposons (*mariner*-like elements in the case of *Stowaway*, Yang et al. 2009). With respect to the origin of such 'orphan' MITE families, it remains possible that their progenitors are direct derivatives of autonomous elements that did not reach fixation or are no longer recognizable in the genome (Feschotte et al. 2003). Alternatively, some may have arisen 'de novo', by juxtaposition of sequences that were fortuitously recognized by transposition enzymes produced in *trans*. This scenario has been documented at least once in *Drosophila* (Tsubota and Huong 1991), but to our knowledge, never in plants.

SINEs represent another atypical category of non-autonomous elements that derive from non-coding genes transcribed by RNA polymerase III (pol III), most commonly tRNA genes (Deragon and Zhang 2006). The simplest SINE families are equivalent to amplified tRNA retrogene families, which apparently result from accidental *trans*-recognition by the enzymatic machinery of autonomous LINEs. The use of an internal promoter (retained after retroposition) coupled to the short length and perhaps also the cellular localization of pol III transcripts may explain the recurrent amplification of tRNA genes by retroposition. More complex SINEs are formed either by multimerization, duplication and/or fusion with the 3′ terminus of a LINE (Deragon and Zhang 2006). Such chimeric SINEs may become highly efficient at hijacking the machinery of their partner LINEs. Perhaps the best-known SINE is the *Alu* element of primates, which is present in over a million copies per haploid human genome (Lander et al. 2001). SINEs have been identified in a wide range of plant species and individual families may attain several thousand copies (Deragon and Zhang 2006), but due to their short size they tend to make up a relatively small fraction of the repetitive DNA content of plant genomes (Fig. 2.5).

2.3 Transposable Elements Biology: Intrinsic Factors of Transposon Proliferation

Although the total quantitative amount of TEs varies tremendously among (and possibly within) plant species, every genome analyzed so far has been found to harbor representatives of both Class I and Class II TEs (Figs. 2.4, 2.5). As mentioned above, however, the relative qualitative contribution of the two classes and their subclasses to the total TE population varies substantially among species. For example, LTR-retrotransposons predominate in the genomes of cotton and maize (Hawkins et al. 2006; Vitte and Bennetzen 2006), but less so in the genomes of rice or *Lotus japonicus*, where DNA transposons are as (or more) successful than other TE types, as measured by copy numbers (Holligan et al. 2006). Additionally, there are differences in the chromosomal distribution of Class I and Class II elements in the genome (e.g. Peterson-Burch et al. 2004; International Rice Genome Sequencing Project 2005; Baucom et al. 2009). These variations correspond, in part, to the disparate histories of TE invasion experienced by different plant lineages, in addition to how an organism copes with these invasions, which is greatly influenced by host biology, as discussed in Sect. 2.4. In the present section, we examine how the biology and properties of the TEs themselves may lead to significant variation in TE composition among species, and possibly, within the genome.

2.3.1 Mechanisms of Transposition

The mechanism by which a TE family is amplified may determine, in part, their pattern of proliferation and diversification in the genome. Part of the proliferative success of Class I retrotransposons in many taxonomic groups (particularly plants) is ensured by their replicative mode of transposition, where in principle, a small number of 'master' copies can produce hundreds or thousands of 'daughter' copies during a single amplification event (Fig. 2.2a, b). Evidence of such transpositional "bursts" comes from phylogenetically informed analyses in both rice and *Gossypium*, where comparative sequence analyses within various TE families indicates waves of TE accumulation surrounded by periods of relative quiescence (Piegu et al. 2006; Hawkins et al. 2008, 2009). Additionally,

diversification can be accomplished via "template-switching" during reverse transcription, first described in retroviruses (Pathak and Hu 1997), in which two different RNA molecules co-localized in a virus-like particle combine, leading to a new, chimeric element. Recent data point to this mechanism as an important force driving the evolution of maize LTR-retrotransposons (Sharma et al. 2008). Template-switching may also occur during the transposition of non-LTR retrotransposons (Garcia-Perez et al. 2007), and this mechanism may explain the chimeric structure and modular evolution of SINEs (Deragon and Zhang 2006). Exchange of sequences is also possible at the DNA level and may promote the diversification of DNA transposons, including Helitrons (Yang and Bennetzen 2009), providing a mechanism for the acquisition of host gene fragments by various plant TEs (Bureau et al. 1994; Jiang et al. 2004; see also Chap. 2). In fact, template-switching or other forms of inter-element recombination may be viewed as a primitive form of sex, promoting the genetic diversification of TEs.

Class II DNA transposons are mobilized by a cut-and-paste mechanism where the element is excised from one locus and re-inserted elsewhere in the genome (Fig. 2.2e, f). This process, by itself, does not result in an increase in copy number, as the element is not replicated; however, increases in Class II element copy number can occur through two known mechanisms, both of which are dependent upon host cellular activities. First, upon excision of a Class II element, the consequential double-stranded DNA break can be repaired by homologous recombination using the transposon copy located on the homologous chromosome as a template (Engels et al. 1990), or alternatively, the sister chromatid if excision takes place during S phase. DNA replication offers a second opportunity for duplication: when a transposon jumps ahead of a replication fork, from a post- to a pre-replicated region, it can effectively be replicated twice (Ros and Kunze 2001). Nevertheless, each of these mechanisms produces a net gain of only one copy per transposition event.

In spite of the conservative nature of cut-and-paste transposition, it is clear that DNA transposons can amplify to very high copy numbers (up to several thousands per family), as documented by the explosive bursts of MITEs in many angiosperms (for a spectacular example of MITE amplification in 'real-time', see Naito et al. 2009). How MITEs could achieve such high copy number has remained a mystery for nearly two decades, but some important clues have surfaced recently, thanks to the study of actively transposing MITE families discovered in the rice genome. It is now established that MITEs rely on a transposase encoded by larger, autonomous elements (Feschotte and Mouchès 2000; Zhang et al. 2001; Feschotte et al. 2003). Furthermore, the data point to a typical cut-and-paste mechanism involving excision and re-insertion similar to that of other eukaryotic DNA transposons (Petersen and Seberg 2000; Nakazaki et al. 2003; Yang et al. 2006, 2009). One key to the mystery of MITE amplification seems to lie in the complexity of their interactions with transposases, as revealed by functional studies of *Stowaway* MITEs and their partner *Osmar* transposases in rice. First, a single source of *Osmar* transposase is capable of interaction and mobilization of a diversity of *Stowaways* having different origins, even in the absence of extensive sequence similarity between *Osmar* and *Stowaway* termini (Feschotte et al. 2005; Yang et al. 2009). Second, some *Stowaway* elements possess the inherent ability to excise at higher efficiency in response to *Osmar* transposase than other substrates, including the cognate *Osmar* element providing the source of transposase. Sequence-swapping experiments indicate that the excision hyperactivity of the MITE stems from a combination of properties, including short size, the absence of *cis*-elements present in the autonomous *Osmar* element that repress transposition, and conversely the presence of *cis*-elements in the MITE internal sequence that enhance transposition (Yang et al. 2009). Thus, multiple, overlaying rampant amplification of MITEs in plant genomes.

2.3.2 Targeting Strategies

2.3.2.1 Transposable Elements Occupy Different Genomic Niches

Genomes can be partitioned into a variety of "chromosomal niches" that are colonized by various repetitive sequences (Kidwell and Lisch 2001). In particular, constitutive heterochromatic regions of the plant genome, such as pericentromeric regions, knobs, and subtelomeres, represent chromosomal niches heavily occupied by LTR-retrotransposons (Miller et al. 1998; Lippman et al. 2004; Kejnovsky et al. 2006b). By contrast, most DNA transposons, and MITEs especially, are found at higher density in euchromatic regions where they often reside within or in close proximity to genes (Bureau and Wessler 1992; International Rice Genome Sequencing Project 2005).

To account for the chromosomal distribution of TEs, one must consider the action of several, non-mutually exclusive forces acting at the time of insertion and often long after insertion. Some of these forces are inherent to the transposition machinery of the elements that confer insertion preference for certain chromosomal or sequence features. Also, natural selection will favor the fixation of beneficial insertions and the elimination of deleterious ones from the population. Finally, an array of indirect forces may act more gradually, influencing the decay of the elements or their removal by deletion or recombination (for review, Pritham 2009). These latter forces include rates of substitution, deletion and recombination, which can vary dramatically along

chromosomes (e.g. low recombination in peri-centromeric regions) and also among species. It is often difficult to discern the relative importance of these many forces on TE accumulation differentially over time. The effect of insertion preference tends to be more apparent for younger TE insertions, while recombination and deletional processes become more significant as TEs become older and accumulate in the genome.

With respect to TE insertion preference targeting of TEs into specific chromosomal locations, such as heterochromatin where they likely have less deleterious effects, represents a mechanism minimizing the negative impact of TEs on the host. Thus different targeting strategies are likely to evolve among TEs to occupy diverse genomic niches and thereby contribute to their evolutionary persistence. The biased TE populations of two diverged yeast species, *S. cerevisiae* and *S. pombe*, provide an extreme example. In these streamlined genomes, only a handful of LTR retrotransposon families co-exist, and remarkably, all have adopted different targeting strategies. Ty1 and Ty3 of *S. cerevisiae* preferentially insert upstream of tRNA genes and other units transcribed by Pol III (Ji et al. 1993), while Ty5 targets the silent chromatin located in subtelomeric regions and around the mating loci (Zou et al. 1996). In *S. pombe*, Tf elements preferentially insert upstream of Pol II-transcribed genes (Bowen et al. 2003).

In plants, there is evidence that the accumulation of *Arabidopsis* LTR retrotransposons in pericentromeric regions and other highly heterochromatic chromosomal compartments is the result of both active targeting and selective retention over time (Pereira 2004; Peterson-Burch et al. 2004). Comparison of the age and chromosomal distribution of TEs in *Arabidopsis* indicate that *copia*-like elements are integrated fairly randomly into the genome while *gypsy*-like elements preferentially insert into the pericentromeric heterochromatin (Pereira 2004). In maize, high-copy-number LTR retrotransposon families are found to primarily accumulate in gene-poor regions, while LINEs, SINEs, and low-copy-number LTR retrotransposons show biased insertion in gene-rich regions (Baucom et al. 2009). The accumulation of LTR elements in heterochromatic regions is also evident in rice, while conversely, MITEs and most other DNA transposons are found in higher density close to or within protein-coding genes (International Rice Genome Sequencing Project 2005). An examination of a large number of *de novo* insertions of *Mutator* DNA elements in rice and maize (Dietrich et al. 2002; Liu et al. 2009; Jiang et al. 2011) and *mPing* MITEs in *Arabidopsis* and rice (Yang et al. 2007; Naito et al. 2009) demonstrate that these transposons actively target genes, with a preference for insertion in their 5′ upstream region. A preference for insertion within or near the same family of elements (self-preference) was also observed for *Tourist* MITEs in maize and rice (Jiang and Wessler 2001) and for *Helitrons* in maize (Yang and Bennetzen 2009).

The molecular mechanisms underlying the targeting of plant TEs remain poorly understood, but may involve the recognition of specific DNA motifs (Zhang et al. 2001), their position relative to the nucleosome (Jiang and Wessler 2001), or the epigenetic state of the insertion sites (Brady et al. 2008). Indeed, emerging evidence suggests that mobile elements may, in some cases, possess the inherent capacity to target their integration toward particular chromatin domains. For example, chromoviruses are *gypsy*-like retrotransposons that contain 40–50 amino acid "chromodomains" at the C-terminus of their integrase (Kordis 2005; Novikova 2009). These chromodomains are thought to direct integration into heterochromatic regions via interaction with methylated histone residues, thereby facilitating targeted insertion into relatively gene-poor regions (Gao et al. 2008).

2.3.2.2 The Special Relationship of Plant Retrotransposons with Centromeres

In spite of their conserved function, centromeres are highly dynamic at the sequence level (see Hirsch and Jiang 2012, this volume). The major components of plant centromeres are large arrays of tandem repeat, called satellite DNA (Jiang et al. 2003), as exemplified by rice CentO (Cheng et al. 2002) and maize CentC satellites (Ananiev et al. 1998). In most plants examined, centromeric satellites are intermingled with a particular group of *gypsy*-like elements called centromeric retrotransposons (*CRs*). *CRs* were originally found in many grass species, such as *CRM* in maize (Zhong et al. 2002; Nagaki et al. 2003), *CRR* and *RIRE7* in rice (Kumekawa et al. 2001; Cheng et al. 2002; Nagaki et al. 2005), *CEREBA* in barley (Presting et al. 1998), *CRW* in wheat (Liu et al. 2008), *CRS* in sugarcane (Nagaki and Murata 2005), and *Bilby* in rye (Francki 2001). However *CRs* are not restricted to grasses, as they were recently discovered in *Arabidopsis*, soybean (Du et al. 2010) and many other eudicot species (Neumann et al. 2011). These findings suggest that *CRs* colonized centromeres (or pericentromeres) before the divergence of monocots and eudicots and have been stable components of angiosperm genomes ever since (Du et al. 2010). Consistent with this scenario, phylogenetic analyses revealed that *CRs* represent a deeply rooted, monophyletic clade of *gypsy*-like elements (Gorinsek et al. 2004; Kordis 2005; Neumann et al. 2011). There are exceptions, nonetheless, such as in *Oryza brachyantha*, where *CentO* satellites and *CRR* retrotransposons have disappeared from functional centromeres and were subsequently replaced by *FRetro3*, a retrotransposon belonging to a different lineage of *gypsy*-like elements (Gao et al. 2009).

The ancient origin, vertical persistence and relatively high level of sequence conservation across species set *CRs* apart from other plant LTR retrotransposons, and these characteristics prompted several investigators to hypothesize that *CRs* could have been co-opted for a cellular function (Zhong et al. 2002). One possibility is that that they provide an abundant source of promoters for the transcription of satellite repeats, which may be important for the establishment of centromere identity and/or chromosome segregation (May et al. 2005). Indeed, plant centromeric satellites and *CRs* themselves are often transcribed (Topp et al. 2004; Neumann et al. 2007), and there is evidence that the transcripts of rice *CRR* elements are partially processed into small RNAs through the RNA interference (RNAi) pathway (Neumann et al. 2007). It is tempting to speculate that *CR*-derived small RNAs are implicated in the formation and/or maintenance of centromeric chromatin (Neumann et al. 2007), akin to the mechanism underlying the formation of pericentromeric heterochromatin in fission yeast which are also initiated by transcription and RNAi-dependent processing of repetitive elements (Volpe et al. 2002; Grewal and Jia 2007). Furthermore maize *CRM* DNA and, surprisingly, *CRM*-derived transcripts, both interact with the centromeric histone CENH3 (Zhong et al. 2002; Topp et al. 2004) and at least one subfamily of *CRM* elements (*CRM2*) exhibit tightly phased positioning on CENH3-containing nucleosomes (Gent et al. 2011). Together these data point at a functional association of *CRs* with centromeric chromatin, although further experiments are needed to clarify the role of *CRs* in plant centromere biology.

2.4 Influence of Host Biology on Transposable Element Proliferation

Factors acting at the level of the host and affecting the likelihood of fixation of TE insertions, subsequent decay (via nucleotide substitutions or indels), or their physical removal (via large deletions and other recombination events) may have a significant influence on shaping TE content over time. We highlight here three of these forces, effective population size, sexual reproduction and recombination rate that are likely to have prominent effects on TE persistence and accumulation. We also describe the role of recombination in shaping TE proliferation and distribution on sex chromosomes.

2.4.1 Effective Population Size

For long-term persistence any TE family must, on average, give rise to at least one daughter element for each element inactivated by mutation or eliminated by deletion. Strongly deleterious TE insertions are eliminated by selection (Le Rouzic et al. 2007), and the efficiency of selection is proportional to the host effective population size. The persistence of TEs in species with large effective populations, like in most unicellular species, is rare (Wagner 2006), especially in the absence of sex or horizontal transfer. The critical effective population size above which eukaryotic populations appear to be immune to retrotransposon proliferation is suggeste to be ~7×10^7, whereas for DNA transposons it is ~2×10^7 (Lynch and Conery 2003). Additionally, total genome size is inversely correlated with long-term effective population size because TEs form a significant part of the genomes of multicellular eukaryotes having generally smaller effective population sizes (Lynch and Conery 2003). Long-term effective population size reduction then probably enabled increases in genome sizes as well as organism sizes. In plants, it was suggested that species with small population sizes should purge TE insertions less efficiently and hence accrue DNA more rapidly (Lockton et al. 2008); however, a recent study of 205 species of seed plants determined no relationship between effective population size and genome size, suggesting that effective population size is not an especially significant factor in the relative level of proliferation and persistence of TEs in plants (Whitney et al. 2010).

2.4.2 Breeding System

Almost three decades ago Hickey (1982) suggested that the potential for TE proliferation is related to the rate of outcrossing in a given host species. Population genetics and mathematical modeling predict that obligatory out-crossing species should contain a larger number of and more active TEs than self-fertilizing or facultative sexual species, while at the other end of the spectrum obligate asexuals and uniparental organelle genomes should rapid purge active TEs and essentially be free of selfish genetic elements, unless they have recently re-entered by horizontal transfer (Hickey 1982; Bestor 1999; Schön and Martens 2000). Although these theoretical arguments were grounded in arguments of population genetics, they have proven difficult to test empirically (but see Zeyl et al. 1996; Arkhipova and Meselson 2000; Schaack et al. 2010a,b).

Plants, which include closely related selfing and outcrossing species, offer a valuable system to investigate these questions because the genetics of selfing species resemble that of asexuals. Consistent with this theory, the outcrossing *Arabidopsis lyrata* displays higher transposition frequency, stronger selection against new TE insertions, and faster removal of insertions by ectopic recombination than in the selfing *A. thaliana* (Wright et al. 2001, 2003; Lockton and Gaut 2010; Hollister and Gaut 2007). Perhaps consequently, the diversity of TEs is greater in *A. lyrata* than in

A. thaliana (Lockton et al. 2008). However the difference in breeding system may be only partially or indirectly causative of these patterns. As discussed above, demographic history, such as population bottlenecks, has the power to explain most of these variations and to exert a substantial influence on TE dynamics (Lockton et al. 2008; Tenaillon et al. 2010).

2.4.3 Recombination Rates Shape the Chromosomal Distribution of Transposable Elements

The chromosomal distribution of TEs is influenced by many factors, such as local variation in recombination rates or gene density (as reviewed above). Genomic regions with no or low recombination are represented by most of the Y chromosome, B chromosomes, or (peri)centromeres. In particular, the non-recombining Y chromosome is subject to a suite of processes leading to the accumulation of deleterious mutations (Charlesworth and Charlesworth 2000). These processes include (1) Muller's ratchet, (2) genetic hitchhiking and (3) background selection (reviewed in Bachtrog 2006). There are two predictable consequence of these processes, namely, the degeneration of genes and the accumulation of selfish genetic elements, including TEs (Charlesworth et al. 1994). Two models have been evoked to explain the accumulation of TEs in gene-poor regions with low or no recombination, such as the Y chromosome or the peri-centromeric regions of chromosomes (reviewed in Dolgin and Charlesworth 2008). In the "insertion model", there is weaker selection against TEs in gene-poor regions due to the decreased possibility of deleterious insertions, resulting in higher TE abundance in these regions. The "ectopic recombination model" postulates that TEs accumulate in regions of low recombination because ectopic recombination between copies, which is a powerful deletional force, is less frequent in these regions than in regions with high recombination rate (Langley et al. 1988).

The accumulation of TEs in non-recombining regions has been observed empirically on the Y chromosome of humans (Erlandsson et al. 2000; Skaletsky et al. 2003), *Drosophila melanogaster* (Pimpinelli et al. 1995), as well as on the neo-Y chromosome of *Drosophila miranda* (Steinemann and Steinemann 1992; Bachtrog 2003). In plants, however, the relationship of recombination rate to TE distribution is not clear. As noted above, in *A. thaliana* and many other angiosperms examined, TEs tend to accumulate in pericentromeric regions. TE distribution in *A. thaliana*, however, does not correlate with recombination rate, but is negatively correlated with gene density (Wright et al. 2003). Shorter LTR retrotransposons and their fragments accumulate in regions with higher recombination rates, indicating that both recombination and gene density can influence the rate and pattern of TE elimination (Swigonová et al. 2005; Tian et al. 2009).

Some dioecious plants possess sex chromosomes that often are in the early stages of evolution (compared to the more ancient mammalian sex chromosomes), where the Y chromosomes have expanded, rather than contracted, compared to X (Vyskot and Hobza 2004). It is often assumed, but not yet demonstrated, that the increased size of plant Y chromosomes results from the accumulation of TEs in non-recombining regions. Consistent with this idea, various types of repetitive DNA are specific to or enriched on the Y chromosome of several plant species, e.g., RAYS tandem repeats in *Rumex acetosa* (Shibata et al. 1999), LINE elements in *Cannabis sativa* (Sakamoto et al. 2000) and *copia*-like elements in *Marchantia polymorpha* (Okada et al. 2001). In papaya, which possesses the youngest studied plant Y chromosome, the male-specific region of this nascent Y chromosome is associated with a high density of various DNA repeats (Liu et al. 2004). In *Silene latifolia*, the most popular dioecious plant model (Kejnovsky and Vyskot 2010), the Y chromosome is strikingly enlarged (Fig. 2.3a). This is due in part to an accumulation of *copia*-like elements but also to chloroplast DNA insertions and to an expansion of tandem repeats (Hobza et al. 2006; Kejnovsky et al. 2006a; Cermak et al. 2008; Kubat et al. 2008). However, not all TEs of *S. latifolia* accumulate on the Y chromosome. For example, *Ogre*-like *gypsy* elements (Fig. 2.3b) are abundant on all chromosomes but virtually absent on the non-recombining parts of the Y chromosome. Several mechanisms might account for this unexpected distribution, including female-specific transposition activity or specific

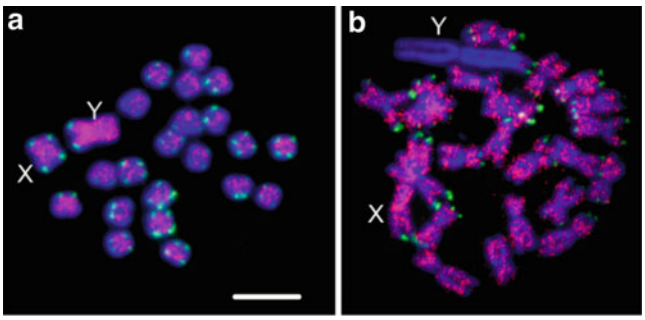

Fig. 2.3 Examples of TE localization on sex chromosomes using FISH in model dioecious plant *Silene latifolia* (*white campion*), species with heteromorphic sex chromosomes. Accumulation of *Copia* elements (in *red*) on the Y chromosome (**a**) in contrast with the *Ogre-like* gypsy retrotransposon (in *red*) that colonizes only recombining parts of genome (**b**). *Ogre-like* elements are ubiquitously distributed on all autosomes and the X chromosome but on the Y chromosome occupy only short pseudoautosomal region while are absent in the large non-recombining parts (**b**). The tandem repeat X-43.1 labels most subtelomeres but on the Y chromosomes only its q-arm (*green signals*). Chromosomes are counterstained by DAPI (*blue*). The X and Y chromosomes are indicated. Bar represents 10 μm. Reproduced by courtesy of Cytogenetic and Genome Research

targeting of recombining regions of genome (Cermak et al. 2008; Kejnovsky et al. 2009a).

The massive aggregation of TEs in early stages of Y chromosome evolution suggests that TEs themselves may be involved in the degeneration of genes located on the Y chromosome, through insertional disruption, rearrangements or post-insertional effects on gene expression (Marais et al. 2008). However, it is still unknown whether the accumulation of retrotransposons causes gene degeneration or whether these elements accumulate on the Y chromosome only after the erosion of most gene content (Steinemann and Steinemann 2005). The comparison of sex chromosomes at different stages of evolution, which should be possible in plants, may provide an opportunity to address this question and evaluate the generality of the models and processes shaping sex chromosomes in plants and animals.

2.5 Transposable Elements and Genome Size Evolution

2.5.1 The C-Value Paradox and Plant TE Composition

The "C-value paradox", a term derived to describe the lack of correlation between morphological complexity and total nuclear DNA content, was resolved in part by the discovery that eukaryotic genomes harbor large and dynamic populations of repetitive sequences, primarily transposable elements. Over the past few decades, numerous studies (summarized in Table 2.1) have described the total TE contribution to genome size and compositional diversity of TEs among and within various plant genomes. These studies have converged upon the conclusion that often the greatest fraction of plant genomes are composed of TEs, particularly in those plants with greater total nuclear content (Zhang and Wessler 2004; Hawkins et al. 2006; Vitte and Bennetzen 2006; Wicker and Keller 2007; Sweredoski et al. 2008; Wicker et al. 2009) (Fig. 2.4). The total TE copy number in plant genomes ranges widely, from as little as a few hundred in those with smaller genome sizes, such as *Arabidopsis*, to hundreds of thousands in their larger genome counterparts (e.g. maize, *Triticum*, *Hordeum*). Notably, this positive correlation between genome size and TE copy number generally holds across a broad range of eukaryotes (Bennett and Leitch 2005).

Copy number of a particular TE family or subfamily is a reflection of its relative success in terms of amplification and subsequent retention in the genome. Comparisons of TE composition across a wide range of plant species suggests that, although the same general TE types are found in all plants, the relative proportions contributed by various classes and subclasses can differ dramatically (Fig. 2.5).

Table 2.1 Genome size and proportion of TEs in plant species

Species	Genome size (Mbp)	Proportion TE (%)	Reference
Arabidopsis thaliana	120	14	The Arabidopsis Genome Initiative (2000)
Fragaria vesca	240	23	Shulaev et al. (2011)
Cucumis sativus	243	24	Huang et al. (2009)
Carica papaya	372	52	Ming et al. (2008)
Medicago truncatula	375	10	Wang and Liu (2008)
Oryza sativa	420	35	Paterson et al. (2009)
Theobroma cacao	430	26	Argout et al. (2011)
Lotus japonicus	470	19	Holligan et al. (2006)
Populus trichocarpa	485	42	Tuskan et al. (2006)
Vitis vinifera	487	17	French-Italian Public Consortium for Grapevine Genome Characterization (2007)
Brassica oleracea	600	20	Qiu et al. (2009)
Sorghum bicolor	740	62	Paterson et al. (2009)
Gossypium raimondii	880	54	Hawkins et al. (2006)
Glycine max	1,100	59	Schmutz et al. (2010)
Zea mays	2,045	76	Paterson et al. (2009)
Hordeum vulgare	5,439	80	Wicker et al. (2009)
Triticum aestivum	16,979	80	Bennett and Smith (1976)
Pinus taeda	21,516	80	Kovach et al. (2010)

Generally speaking, the genomes of eudicots contain fewer transposable elements relative to that of monocots, which have experienced recent and rampant LTR retrotransposon activity (SanMiguel et al. 1998; Vitte and Bennetzen 2006). LTR retrotransposon turnover in monocots appears to be extremely rapid, with both gains and losses of TE sequences occurring over as little as a few million years (see Sect. 2.5.2; Ma et al. 2004). In striking contrast, gymnosperm LTR retrotransposons are distinguished by their high level of decay and significant degree of divergence from angiosperm LTR retrotransposons, indicative of their ancient origin and subsequent long-term retention (Kovach et al. 2010). Additionally, these types of qualitative differences are not necessarily restricted to comparisons among major plant lineages. For example, significant differences in TE content have been observed among species within the genus *Gossypium*, where *gypsy*-like retrotransposons comprise the majority of the TE fraction

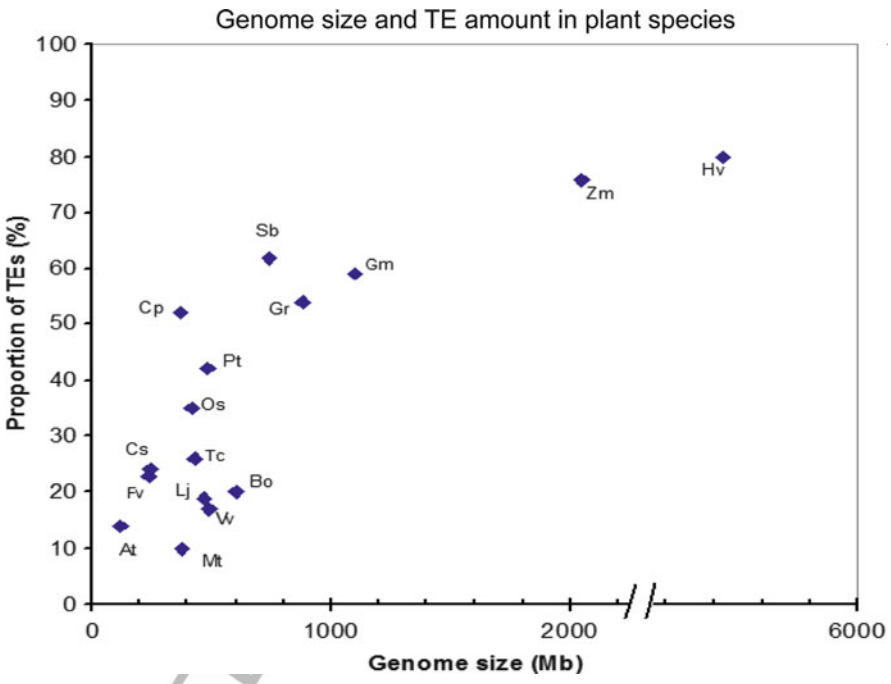

Fig. 2.4 Positive correlation between genome size and TE amount in selected plant species. *Arabidopsis thaliana* (At), *Fragaria vesca* (Fv), *Cucumis sativus* (Cs), *Carica papaya* (Cp), *Medicago truncatula* (Mt), *Oryza sativa* (Os), *Theobroma cacao* (Tc), *Lotus japonicus* (Lj), *Populus trichocarpa* (Pt), *Vitis vinifera* (Vv), *Brassica oleracea* (Bo), *Sorghum bicolor* (Sb), *Gossypium raimondii* (Gr), *Glycine max* (Gm), *Zea mays* (Zm) and *Hordeum vulgare* (Hv)

in species with larger genomes, while *copia*-like elements dominate the TE fraction in species with smaller genomes (Hawkins et al. 2006).

Sequence diversity among individuals within a TE family may also differ substantially, and the extent of divergence is often a direct reflection of a particular TE family's age. TE families that have undergone relatively recent proliferation are usually absent among closely related species, while families that have undergone amplification at more distant time points can be shared among closely related organisms as a product of their shared evolutionary history. An example of the former was demonstrated via comparisons of TE families among wheat and barley, where families that were highly abundant in one species were virtually absent in their close relative (Wicker et al. 2009). The rate at which new families appear (through either vertical or horizontal transfer), as well as the rate at which older TE families decay (by nucleotide mutation leading to sequence erosion or sequence removal via deletion) often differ substantially between species, and will be discussed further in the next section. These forces act to mold genome structure and composition not only at higher levels of taxonomic divergence, but often even at the species level.

2.5.2 Variable TE Insertion and Deletion Rates as a Driving Force in Plant Genome Size Evolution

As outlined above, large-scale amplification of transposable elements can lead to extraordinarily high copy numbers within plant genomes, often over short evolutionary timescales (Bennetzen 2005). One of the best-known and most widely cited examples in plants comes from maize, where repeated bursts of retrotransposon amplification over the past 6 million years have been responsible for generating approximately half of the modern maize genome (SanMiguel et al. 1998; Walbot and Petrov 2001). Similarly, a three-fold increase in the genome size of diploid members of *Gossypium* is due to the accumulation of LTR retrotransposons over the past 5–10 Myr (Hawkins et al. 2006). In *Oryza australiensis*, three LTR retrotransposon families proliferated during the last 3 million years leading to a two-fold increase in genome size compared to that of *Oryza sativa* (Piegu et al. 2006). The two- to threefold higher copy number of TEs in *Arabidopsis lyrata* compared to *A. thaliana* is result of higher expression of TEs caused by their less effective silencing (Hollister et al. 2011). These studies, in addition to several other plant genome surveys, clearly demonstrate that amplification of TEs, together with persistent rounds of genome doubling via polyploidization, are the primary mechanisms responsible for genome size expansion and variation in plants (Vitte and Bennetzen 2006; Kejnovsky et al. 2009b). These examples specifically implicate LTR retrotransposons as the agents most often responsible for massive TE-mediated increases in plant genome size. In contrast, comparative analyses of *A. thaliana* and *Brassica oleracea* indicate that several families of DNA transposons have amplified to high copy number in the lineage of *B. oleracea,* and that this activity has contributed to genome expansion in this lineage (Zhang and Wessler 2004), suggesting that DNA transposons may also play a significant role in shaping genome size in plants.

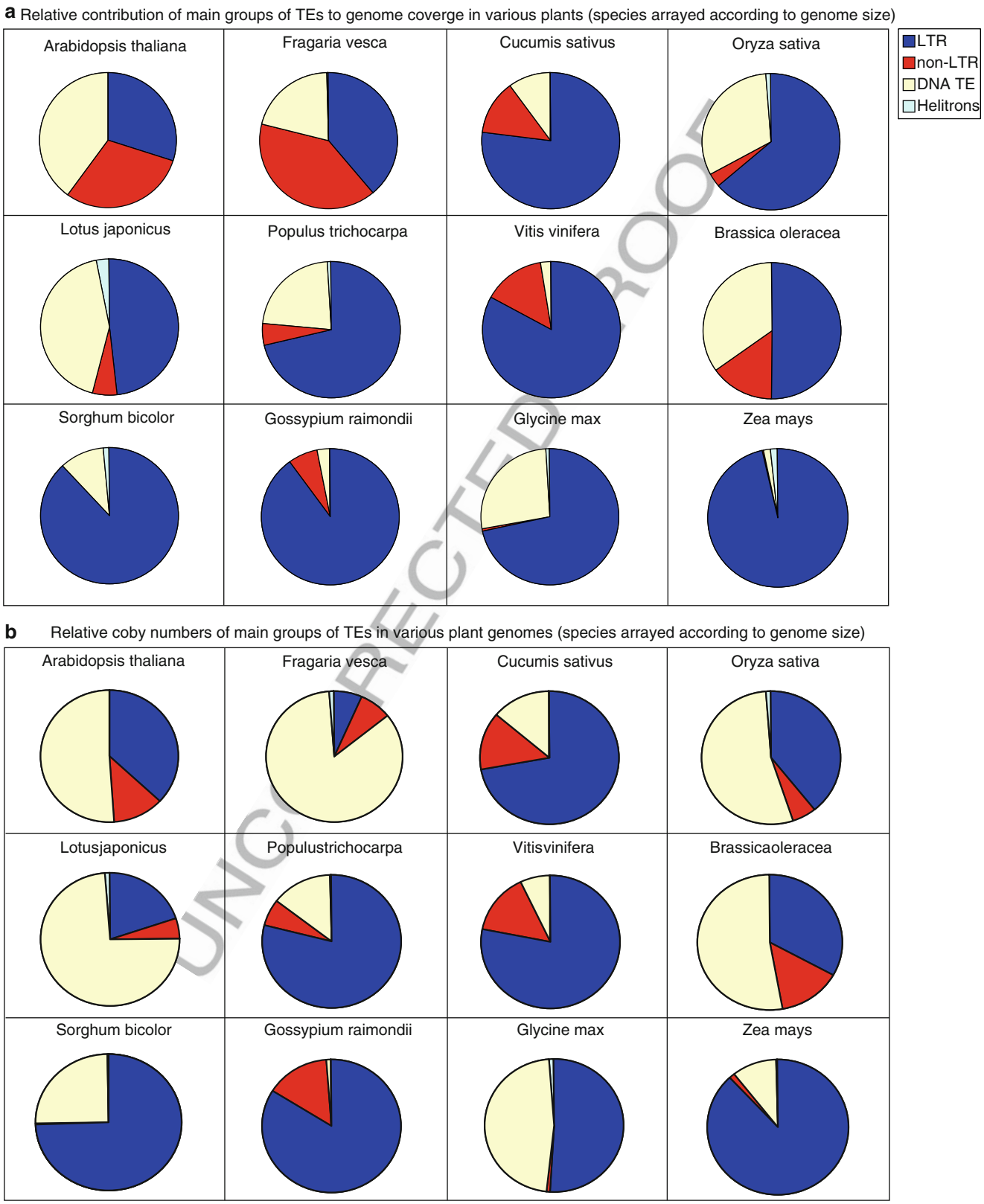

Fig. 2.5 Relative contributions of main groups of transposable elements in 12 plant genomes calculated by either genome coverage (**a**) or TE copy numbers (**b**). LTR retrotransposons (*dark blue*), non-LTR retrotransposons (*red*), DNA transposons (*yellow*) and Helitrons (*white blue*). It is evident that the contribution of LTR retrotransposons to the genome coverage increases with increasing genome size (**a**). DNA transposons are successful in their amplifications in both small and large genomes (**b**) but their contribution to the genome coverage is because their small length not so evident in large genomes (**a**)

The discovery that plant genomes expand via TE amplification combined with a paucity of information regarding mechanisms that might counteract this process and lead to DNA removal raised the question whether the process is a "one-way ticket to genomic obesity" (Bennetzen and Kellogg 1997). Analyses of plant genome size variation across a wide taxonomic range and within a phylogenetic framework show that several species with small genomes are embedded within clades of species characterized by much larger genomes, suggesting that genome downsizing can and does occur (Ma et al. 2004; Leitch et al. 2005; Hawkins et al. 2009). These observations have stimulated significant efforts to discover the genetic mechanisms responsible for DNA removal that might lead to substantial decreases in genome size (Vitte and Panaud 2003, 2005). Presently, two primary mechanisms of genome contraction have been proposed: intra-strand homologous recombination and illegitimate recombination.

Intra-strand homologous recombination, a form of ectopic recombination, is a process in which recombination occurs between non-allelic sequences of high sequence similarity. Such recombination may occur between two different TEs (inter-element recombination) or among highly similar sequences within a single TE, such as the long terminal repeats (LTRs), of the same retroelement (intra-element recombination) (Fig. 2.6). The latter mechanism is straightforward to detect and quantify as it results in the formation of solo LTRs. Since solo LTRs have been identified in virtually all species known to be colonized by LTR retrotransposons, this process appears to occur frequently and it is believed to play a major role in DNA removal in plant genomes. For example, the vast majority of BARE-1 elements in barley are represented by solo LTRs (14,000 full-length and 64,000 solo LTR), indicating massive amplification and subsequent removal of these elements in recent evolutionary history (Vicient et al. 1999). The ratio of full-length elements to solo LTRs is 1:1 in *Arabidopsis*, 2:3 in rice (Devos et al. 2002), 5:1 in maize (SanMiguel et al. 1996) and 1:7–11 in barley (Vicient et al. 1999), suggesting recent amplification of elements in maize as evidenced by the prevalence of intact to partial elements, and conversely, element removal by intra-strand homologous recombination in barley (prevalence of solo LTRs). Intra-strand homologous recombination has been coined a "partial return ticket from genomic obesity" (Vicient et al. 1999) as some portion of the nucleotide sequences involved in recombination (for example, one of the LTRs) is left behind, preventing complete deletion of extraneous DNA.

This type of recombination is expected to operate more strongly (1) to remove TE insertions in regions of high recombination, (2) on larger TE families, and (3) on longer copies of element (Petrov et al. 2003). Because longer elements increase the probability of ectopic recombination, longer TE copies persist in genomes for shorter periods than smaller elements, which is true not only for LTR retrotransposons, but also for Helitrons (Hollister and Gaut 2007). For the same reason, solo LTRs are preferentially formed by TEs with longer LTRs (Du et al. 2010) suggesting selection may occur for shorter LTRs in order to escape these deletions. The mechanisms suppressing genetic recombination may reduce the frequency of the formation of solo-LTRs as was demonstrated in pericentromeric regions of soybean (Du et al. 2010). Additional support for this hypothesis comes from *Oryza sativa*, where short elements accumulate in regions of high recombination while long elements accumulate in regions of low recombination (International

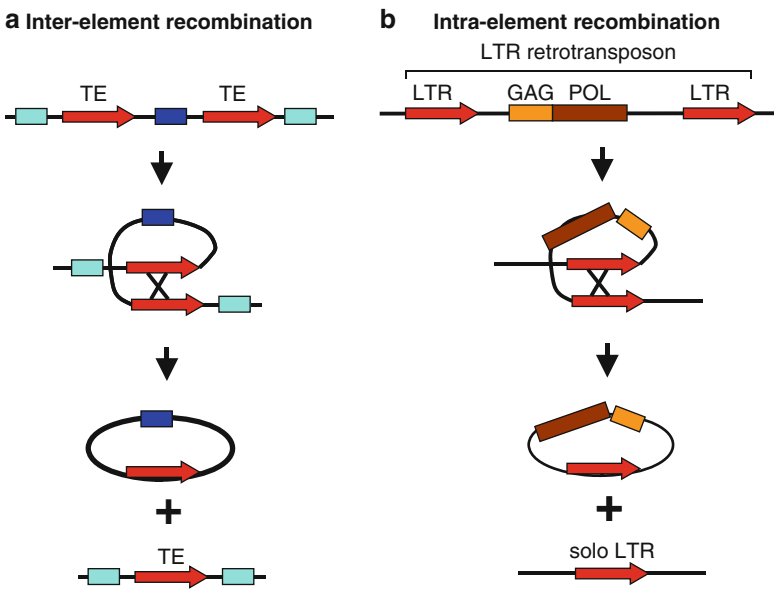

Fig. 2.6 Removal of transposable elements by homologous recombination. (**a**) Recombination between two transposable elements results in deletion of an in-between region. (**b**) Recombination between long terminal repeats (LTRs) of the same retrotransposon results in solo LTRs and deletion of internal region

Rice Genome Sequencing Project 2005). Because the accumulation of substitutions and indels reduce recombination frequency, opportunities for LTR–LTR recombination are rapidly lost in species with high rates of substitution and indels.

Illegitimate recombination is a RecA-independent form of recombination involving sequences of microhomology in which small deletions result due to non-homologous end-joining (NHEJ) or slip-strand mispairing. Sequences of microhomology may be as small as a few nucleotides, and the resulting deletions are often less than 10 bp in length, although they can be much larger. DNA loss leading to genome contraction in *Arabidopsis* and wheat is primarily attributed to illegitimate recombination (Devos et al. 2002; Wicker et al. 2003) but to both intra-strand homologous recombination and illegitimate recombination in rice (Ma et al. 2004). It is unclear at this time as to the evolutionary significance of DNA loss in shaping extant genome size; however, assuming removal rates at a high enough level to counteract genome expansion via TE proliferation, differences in the repair/recombination machinery of the host species might be a driving force in shaping extant genome size (Orel and Puchta 2003).

2.6 Closing

Upon her discovery of transposable elements in the 1950s, Barbara McClintock suggested that these sequences might operate to control gene expression and play a major role in evolution. This suggestion was remarkably prophetic, even though the concept it embodied had to survive several decades of misinterpretation of TEs as mere "junk" or "selfish" prior to emerging in the last decade or more as a central truth about the structure, organization, and function of plant genomes. However our increasing understanding of TE abundance, distribution, and behavior has revealed that the selfish nature of TEs is not incompatible with them playing a significant role in genome evolution at multiple levels, from genome-wide (total nuclear content, chromatin structure, recombination, RNAi, etc.) to local effects (chromosomal rearrangements, regulation of neighboring genes, co-option of individual TE sequences to form new genes, TE-mediated gene duplication, etc.) as summarized above and discussed in more detail in Chap. 2. Thus, after almost 60 years, Barbara McClintock's vision, considered radical at the time and dismissed by most, is receiving growing empirical support. Plant research continues to be at the forefront of TE biology, and the ongoing genomic revolution is promised to yield many more exciting discoveries in the years to come.

Acknowledgments Research on transposable elements in the authors' laboratories has been supported by the Grant Agency of the Czech Republic (grant P305/10/0930), grants No AV0Z50040507 and AV0Z50040702 from the Academy of Sciences of the Czech Republic.

References

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF et al (2000) The genome sequence of *Drosophila melanogaster*. Science 287:2185–2195

Ananiev EV, Phillips RL, Rines HW (1998) Chromosome-specific molecular organization of maize (*Zea mays* L.) centromeric regions. Proc Natl Acad Sci USA 95:13073–13078

Argout X, Salse J, Aury J-M, Guiltinan MJ, Droc G, Gouzy J, Allegre M, Chaparro T, Maximova SN, Abrouk M et al (2011) The genome of *Theobroma cacao*. Nat Genet 43:101–109

Arkhipova I, Meselson M (2000) Transposable elements in sexual and ancient asexual taxa. Proc Natl Acad Sci USA 97:14473–14477

Bachtrog D (2003) Accumulation of Spock and Worf, two novel non-LTR retrotransposons, on the neo-Y chromosome of *Drosophila miranda*. Mol Biol Evol 20:173–181

Bachtrog D (2006) A dynamic view of sex chromosome evolution. Curr Opin Genet Dev 16:578–585

Baucom RS, Estill JC, Chaparro C, Upshaw N, Jogi A, Deragon J, Westerman RP, Sanmiguel PJ, Bennetzen JL (2009) Exceptional diversity, non-random distribution, and rapid evolution of retroelements in the B73 maize genome. PLoS Genet 5:e1000732

Bennett MD, Leitch IJ (2005) Nuclear DNA amounts in angiosperms: progress, problems and prospects. Ann Bot 95:45–90

Bennett MD, Smith JB (1976) Nuclear DNA amounts in angiosperms. Philos Trans Roy Soc B 274:227–274

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

Bennetzen JL, Kellogg EA (1997) Do plants have a one-way ticket to genomic obesity? Plant Cell 9:1509–1514

Berg DE, Howe MM (1989) Mobile DNA. American Society for Microbiology, Washington, DC

Bestor TH (1999) Sex brings transposons and genomes into conflict. Genetica 107:289–295

Bowen NJ, Jordan IK, Epstein JA, Wood V, Levin HL (2003) Retrotransposons and their recognition of pol II promoters: a comprehensive survey of the transposable elements from the complete genome sequence of *Schizosaccharomyces pombe*. Genome Res 13:1984–1997

Brady TL, Fuerst PG, Dick RA, Schmidt C, Voytas DF (2008) Retrotransposon target site selection by imitation of a cellular protein. Mol Cell Biol 28:1230–1239

Bureau TE, Wessler SR (1992) Tourist: a large family of small inverted repeat elements frequently associated with maize genes. Plant Cell 4:1283–1294

Bureau TE, White SE, Wessler SR (1994) Transduction of a cellular gene by a plant retroelement. Cell 77:479–480

Cermak T, Kubat Z, Hobza R, Koblizkova A, Widmer A, Macas J, Vyskot B, Kejnovsky E (2008) Survey of repetitive sequences in *Silene latifolia* with respect to their distribution on sex chromosomes. Chromosome Res 16:961–976

Charlesworth B, Charlesworth D (2000) The degeneration of Y chromosomes. Philos Trans Roy Soc B 355:1563–1572

Charlesworth B, Sniegowski P, Stephan W (1994) The evolutionary dynamics of repetitive DNA in eukaryotes. Nature 371:215–220

Cheng Z, Dong F, Langdon T, Ouyang S, Buell CR, Gu M, Blattner FR, Jiang J (2002) Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon. Plant Cell 14:1691–1704

Craig NL, Craigie R, Gellert M, Lambowitz AM (2002) Mobile DNA II. ASM Press, Washington, DC

Deragon J, Zhang X (2006) Short interspersed elements (SINEs) in plants: origin, classification, and use as phylogenetic markers. Syst Biol 55:949–956

Devos KM, Brown JKM, Bennetzen JL (2002) Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis*. Genome Res 12:1075–1079

Dietrich CR, Cui F, Packila ML, Li J, Ashlock DA, Nikolau BJ, Schnable PS (2002) Maize Mu transposons are targeted to the 5′ untranslated region of the gl8 gene and sequences flanking Mu target-site duplications exhibit nonrandom nucleotide composition throughout the genome. Genetics 160:697–716

Dolgin ES, Charlesworth B (2008) The effects of recombination rate on the distribution and abundance of transposable elements. Genetics 178:2169–2177

Du J, Tian Z, Hans CS, Laten HM, Cannon SB, Jackson SA, Shoemaker RC, Ma J (2010) Evolutionary conservation, diversity and specificity of LTR-retrotransposons in flowering plants: insights from genome-wide analysis and multi-specific comparison. Plant J. http://www.ncbi.nlm.nih.gov/pubmed/20525006. Accessed 13 Sept 2010

Eickbush TH, Malik HS (2002) Origin and evolution of retrotransposons. In: Craig NL, Craigie R, Gellert M, Lambowitz AM (eds) Mobile DNA. ASM Press, Washington, DC, pp 1111–1146

Engels WR, Johnson-Schlitz DM, Eggleston WB, Sved J (1990) High-frequent P element loss in *Drosophila* is homolog dependent. Cell 10:515–525

Erlandsson R, Wilson JF, Pääbo S (2000) Sex chromosomal transposable element accumulation and male-driven substitutional evolution in humans. Mol Biol Evol 17:804–812

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

Feschotte C, Mouchès C (2000) Evidence that a family of miniature inverted-repeat transposable elements (MITEs) from the *Arabidopsis thaliana* genome has arisen from a pogo-like DNA transposon. Mol Biol Evol 17:730–737

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

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

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

Feschotte C, Osterlund MT, Peeler R, Wessler SR (2005) DNA-binding specificity of rice mariner-like transposases and interactions with Stowaway MITEs. Nucleic Acids Res 33:2153–2165

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

Francki MG (2001) Identification of Bilby, a diverged centromeric Ty1-copia retrotransposon family from cereal rye (*Secale cereale* L.). Genome 44:266–274

French-Italian public consortium for grapevine genome characterization (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449:463–468

Gao X, Hou Y, Ebina H, Levin HL, Voytas DF (2008) Chromodomains direct integration of retrotransposons to heterochromatin. Genome Res 18:359–369

Gao D, Gill N, Kim H, Walling JG, Zhang W, Fan C, Yu Y, Ma J, SanMiguel P, Jiang N et al (2009) A lineage-specific centromere retrotransposon in *Oryza brachyantha*. Plant J 60:820–831

Garcia-Perez JL, Doucet AJ, Bucheton A, Moran JV, Gilbert N (2007) Distinct mechanisms for trans-mediated mobilization of cellular RNAs by the LINE-1 reverse transcriptase. Genome Res 17:602–611

Gent JI, Schneider KL, Topp CN, Rodriguez C, Presting GG, Dawe RK (2011) Distinct influences of tandem repeats and retrotransposons on CENH3 nucleosome positioning. Epigenetics Chromatin 4:3

Gorinsek B, Gubensek F, Kordis D (2004) Evolutionary genomics of chromoviruses in eukaryotes. Mol Biol Evol 21:781–798

Green MM (1980) Transposable elements in *Drosophila* and other Diptera. Annu Rev Genet 14:109–120

Grewal SIS, Jia S (2007) Heterochromatin revisited. Nat Rev Genet 8:35–46

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

Hawkins JS, Kim H, Nason JD, Wing RA, Wendel JF (2006) Differential lineage-specific amplification of transposable elements is responsible for genome size variation in *Gossypium*. Genome Res 16:1252–1261

Hawkins JS, Hu G, Rapp RA, Grafenberg JL, Wendel JF (2008) Phylogenetic determination of the pace of transposable element proliferation in plants: copia and LINE-like elements in *Gossypium*. Genome 51:11–18

Hawkins JS, Proulx SR, Rapp RA, Wendel JF (2009) Rapid DNA loss as a counterbalance to genome expansion through retrotransposon proliferation in plants. Proc Natl Acad Sci USA 106:17811–17816

Hickey DA (1982) Selfish DNA: a sexually-transmitted nuclear parasite. Genetics 101:519–531

Hirsch C, Jiang J (2012) Centromeres: sequences, structure, and biology. In: Wendel JF (ed) Plant genome diversity, vol 1, Plant genomes, their residents, and their evolutionary dynamics. Springer, Wien, New York

Hobza R, Lengerova M, Svoboda J, Kubekova H, Kejnovsky E, Vyskot B (2006) An accumulation of tandem DNA repeats on the Y chromosome in *Silene latifolia* during early stages of sex chromosome evolution. Chromosoma 115:376–382

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

Hollister JD, Gaut BS (2007) Population and evolutionary dynamics of Helitron transposable elements in *Arabidopsis thaliana*. Mol Biol Evol 24:2515–2524

Hollister JD, Smith LM, Guo Y-L, Ott F, Weigel D, Gaut BS (2011) Transposable elements and small RNAs contribute gene expression divergence between *Arabidopsis thaliana* and *Arabidopsis lyrata*. Proc Natl Acad Sci USA 108:2322–2327

Huang S, Li R, Zhang Z, Li L, Gu X, Fan W, Lucas WJ, Wang X, Xie B, Ni P et al (2009) The genome of the cucumber, *Cucumis sativus* L. Nat Genet 41:1275–1283

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

Ji H, Moore DP, Blomberg MA, Braiterman LT, Voytas DF, Natsoulis G, Boeke JD (1993) Hotspots for unselected Ty1 transposition events on yeast chromosome III are near tRNA genes and LTR sequences. Cell 73:1007–1018

Jiang N, Wessler SR (2001) Insertion preference of maize and rice miniature inverted repeat transposable elements as revealed by the analysis of nested elements. Plant Cell 13:2553–2564

Jiang J, Birchler JA, Parrott WA, Dawe RK (2003) A molecular view of plant centromeres. Trends Plant Sci 8:570–575

Jiang N, Bao Z, Zhang X, Eddy SR, Wessler SR (2004) Pack-MULE transposable elements mediate gene evolution in plants. Nature 431:569–573

Jiang N, Ferguson AA, Slotkin RK, Lisch D (2011) Pack-Mutator-like transposable elements (Pack-MULEs) induce directional modification of genes through biased insertion and DNA acquisition. Proc Natl Acad Sci USA 108:1537–1542

Kalendar R, Vicient CM, Peleg O, Anamthawat-Jonsson K, Bolshoy A, Schulman AH (2004) Large retrotransposon derivatives: abundant, conserved but nonautonomous retroelements of barley and related genomes. Genetics 166:1437–1450

Kapitonov VV, Jurka J (2007) Helitrons on a roll: eukaryotic rolling-circle transposons. Trends Genet 23:521–529

Kejnovsky E, Vyskot B (2010) *Silene latifolia*: the classical model to study heteromorphic sex chromosomes. Cytogenet Genome Res 129:250–262

Kejnovsky E, Kubat Z, Hobza R, Lengerova M, Sato S, Tabata S, Fukui K, Matsunaga S, Vyskot B (2006a) Accumulation of chloroplast DNA sequences on the Y chromosome of *Silene latifolia*. Genetica 128:167–175

Kejnovsky E, Kubat Z, Macas J, Hobza R, Mracek J, Vyskot B (2006b) Retand: a novel family of gypsy-like retrotransposons harboring an amplified tandem repeat. Mol Genet Genomics 276:254–263

Kejnovsky E, Hobza R, Cermak T, Kubat Z, Vyskot B (2009a) The role of repetitive DNA in structure and evolution of sex chromosomes in plants. Heredity 102:533–541

Kejnovsky E, Leitch IJ, Leitch AR (2009b) Contrasting evolutionary dynamics between angiosperm and mammalian genomes. Trends Ecol Evol 24:572–582

Kidwell MG, Lisch DR (2001) Perspective: transposable elements, parasitic DNA, and genome evolution. Evolution 55:24

Kordis D (2005) A genomic perspective on the chromodomain-containing retrotransposons: chromoviruses. Gene 347:161–173

Kovach A, Wegrzyn JL, Parra G, Holt C, Bruening GE, Loopstra CA, Hartigan J, Yandell M, Langley CH, Korf I et al (2010) The *Pinus taeda* genome is characterized by diverse and highly diverged repetitive sequences. BMC Genomics 11:420

Kubat Z, Hobza R, Vyskot B, Kejnovsky E (2008) Microsatellite accumulation on the Y chromosome in *Silene latifolia*. Genome 51:35356

Kumekawa N, Ohmido N, Fukui K, Ohtsubo E, Ohtsubo H (2001) A new gypsy-type retrotransposon, RIRE7: preferential insertion into the tandem repeat sequence TrsD in pericentromeric heterochromatin regions of rice chromosomes. Mol Genet Genomics 265:48488

AU3 Kunze R, Starlinger P (1989) The putative transposase of transposable element Ac from *Zea mays* L. interacts with subterminal sequences of Ac. EMBO J 8:3173–3185

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W et al (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921

Langley CH, Montgomery E, Hudson R, Kaplan N, Charlesworth B (1988) On the role of unequal exchange in the containment of transposable element copy number. Genet Res 52:223–235

Le Rouzic A, Boutin TS, Capy P (2007) Long-term evolution of transposable elements. Proc Natl Acad Sci USA 104:19371–19380

Leitch IJ, Soltis DE, Soltis PS, Bennett MD (2005) Evolution of DNA amounts across land plants (embryophyta). Ann Bot 95:207–217

Levis R, O'Hare K, Rubin GM (1984) Effect of transposable element insertions on RNA encoded by the white gene of *Drosophila*. Cell 38:471–481

Lippman Z, Gendrel AV, Black M, Vaughn MW, Dedhia N, McCombie WR, Lavine K, Mittal V, May B, Kasschau KD et al (2004) Role of transposble elements in heterochromatin and epigenetic control. Nature 430:471–476

Liu Z, Moore PH, Ma H, Ackerman CM, Ragiba M, Yu Q, Pearl HM, Kim MS, Charlton JW, Stiles JI et al (2004) A primitive Y chromosome in papaya marks incipient sex chromosome evolution. Nature 427:348–352

Liu Z, Yue W, Li D, Wang RR, Kong X, Lu K, Wang G, Dong Y, Jin W, Zhang X (2008) Structure and dynamics of retrotransposons at wheat centromeres and pericentromeres. Chromosoma 117: 445–456

Liu S, Yeh CT, Ji T, Ying K, Wu H, Tang HM, Fu Y, Nettleton D, Schnable PS (2009) Mu transposon insertion sites and meiotic recombination events co-localize with epigenetic marks for open chromatin across the maize genome. PLoS Genet 5:e1000733

Lockton S, Gaut BS (2010) The evolution of transposable elements in natural populations of self-fertilizing *Arabidopsis thaliana* and its outcrossing relative *Arabidopsis lyrata*. BMC Evol Biol 10:10

Lockton S, Ross-Ibarra J, Gaut BS (2008) Demography and weak selection drive patterns of transposable element diversity in natural populations of *Arabidopsis lyrata*. Proc Natl Acad Sci USA 105:13965–13970

Lynch M, Conery JS (2003) The origins of genome complexity. Science 302:1401–1404

Ma J, Devos KM, Bennetzen JL (2004) Analyses of LTR-retrotransposon structures reveal recent and rapid genomic DNA loss in rice. Genome Res 14:860–869

Marais GAB, Nicolas M, Bergero R, Chambrier P, Kejnovsky E, Monéger F, Hobza R, Widmer A, Charlesworth D (2008) Evidence for degeneration of the Y chromosome in the dioecious plant *Silene latifolia*. Curr Biol 18:545–549

May BP, Lippman ZB, Fang Y, Spector DL, Martienssen RA (2005) Differential regulation of strand-specific transcripts from *Arabidopsis* centromeric satellite repeats. PLoS Genet 1:e79

Miller JT, Dong F, Jackson SA, Song J, Jiang J (1998) Retrotransposon-related DNA sequences in the centromeres of grass chromosomes. Genetics 150:1615–1623

Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A, Saw JH, Senin P, Wang W, Ly BV, Lewis KLT et al (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). Nature 452:991–996

Morgante M, De Paoli E, Radovic S (2007) Transposable elements and the plant pan-genomes. Curr Opin Plant Biol 10:149–155

Nagaki K, Murata M (2005) Characterization of CENH3 and centromere-associated DNA sequences in sugarcane. Chromosome Res 13:195–203

Nagaki K, Song J, Stupar RM, Parokonny AS, Yuan Q, Ouyang S, Liu J, Hsiao J, Jones KM, Dawe RK et al (2003) Molecular and cytological analyses of large tracks of centromeric DNA reveal the structure and evolutionary dynamics of maize centromeres. Genetics 163:759–770

Nagaki K, Neumann P, Zhang D, Ouyang S, Buell CR, Cheng Z, Jiang J (2005) Structure, divergence, and distribution of the CRR centromeric retrotransposon family in rice. Mol Biol Evol 22:845–855

Naito K, Zhang F, Tsukiyama T, Saito H, Hancock CN, Richardson AO, Okumoto Y, Tanisaka T, Wessler SR (2009) Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. Nature 461:1130–1134

Nakazaki T, Okumoto Y, Horibata A, Yamahira S, Teraishi M, Nishida H, Inoue H, Tanisaka T (2003) Mobilization of a transposon in the rice genome. Nature 421:170–172

Neumann P, Koblízková A, Navrátilová A, Macas J (2006) Significant expansion of *Vicia pannonica* genome size mediated by amplification of a single type of giant retroelement. Genetics 173:1047–1056

Neumann P, Yan H, Jiang J (2007) The centromeric retrotransposons of rice are transcribed and differentially processed by RNA interference. Genetics 176:749–761

Neumann P, Navrátilová A, Koblížková A, Kejnovský E, Hřibová E, Hobza R, Widmer A, Doležel J, Macas J (2011) Plant centromeric retrotransposons: a structural and cytogenetic perspective. Mob DNA 2:4

Novikova O (2009) Chromodomains and LTR retrotransposons in plants. Commun Integr Biol 2:158–162

Okada S, Sone T, Fujisawa M, Nakayama S, Takenaka M, Ishizaki K, Kono K, Shimizu-Ueda Y, Hanajiri T, Yamato KT et al (2001) The Y chromosome in the liverwort *Marchantia polymorpha* has

accumulated unique repeat sequences harboring a male-specific gene. Proc Natl Acad Sci USA 98:9454–9459

Orel N, Puchta H (2003) Differences in the processing of DNA ends in *Arabidopsis thaliana* and tobacco: possible implications for genome evolution. Plant Mol Biol 51:523–531

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A et al (2009) The *Sorghum bicolor* genome and the diversification of grasses. Nature 457:551–556

Pathak VK, Hu W-S (1997) "Might as well jump!" Template switching by retroviral reverse transcriptase, defective genome formation, and recombination. Semin Virol 8:141–150

Pereira V (2004) Insertion bias and purifying selection of retrotransposons in the *Arabidopsis thaliana* genome. Genome Biol 5:R79

Petersen G, Seberg O (2000) Phylogenetic evidence for excision of Stowaway miniature inverted-repeat transposable elements in triticeae (Poaceae). Mol Biol Evol 17:1589–1596

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

Petrov DA, Aminetzach YT, Davis JC, Bensasson D, Hirsh AE (2003) Size matters: non-LTR retrotransposable elements and ectopic recombination in *Drosophila*. Mol Biol Evol 20:880–892

Piegu B, Guyot R, Picault N, Roulin A, Saniyal A, Kim H, Collura K, Brar DS, Jackson S, Wing RA et al (2006) Doubling genome size without polyploidization: dynamics of retrotransposition-driven genomic expansions in *Oryza australiensis*, a wild relative of rice. Genome Res 16:1262–1269

Pimpinelli S, Berloco M, Fanti L, Dimitri P, Bonaccorsi S, Marchetti E, Caizzi R, Caggese C, Gatti M (1995) Transposable elements are stable structural components of *Drosophila melanogaster* heterochromatin. Proc Natl Acad Sci USA 92:3804–3808

Presting GG, Malysheva L, Fuchs J, Schubert I (1998) A Ty3/gypsy retrotransposon-like sequence localizes to the centromeric regions of cereal chromosomes. Plant J 16:721–728

Pritham EJ (2009) Transposable elements and factors influencing their success in eukaryotes. J Hered 100:648–655

Qiu D, Gao M, Li G, Quiros C (2009) Comparative sequence analysis for *Brassica oleracea* with similar sequences in *B. rapa* and *Arabidopsis thaliana*. Plant Cell Rep 28:649–661

Ros F, Kunze R (2001) Regulation of activator/dissociation transposition by replication and DNA methylation. Genetics 157:1723–1733

Saedler H, Bonas U, Gierl A, Harrison BJ, Klösgen RB, Krebbers E, Nevers P, Peterson PA, Schwarz-Sommer Z, Sommer H (1984) Transposable elements in *Antirrhinum majus* and *Zea mays*. Cold Spring Harb Symp Quant Biol 49:355–361

Sakamoto K, Ohmido N, Fukui K, Kamada H, Satoh S (2000) Site-specific accumulation of a LINE-like retrotransposon in a sex chromosome of the dioecious plant *Cannabis sativa*. Plant Mol Biol 44:723–732

SanMiguel P, Tikhonov A, Jin YK, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z et al (1996) Nested retrotransposons in the intergenic regions of the maize genome. Science 274:765–768

SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL (1998) The paleontology of intergene retrotransposons of maize. Nat Genet 20:43–45

Schaack S, Choi E, Lynch M, Pritham EJ (2010a) DNA transposons and the role of recombination in mutation accumulation in *Daphnia pulex*. Genome Biol 11:R46

Schaack S, Pritham EJ, Wolf A, Lynch M (2010b) DNA transposon dynamics in populations of *Daphnia pulex* with and without sex. Proc Biol Sci 7:2381–2387

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J et al (2010) Genome sequence of the palaeopolyploid soybean. Nature 463:178–183

Schön I, Martens K (2000) Transposable elements and asexual reproduction. Trends Ecol Evol 15:287–288

Sharma A, Schneider KL, Presting GG (2008) Sustained retrotransposition is mediated by nucleotide deletions and interelement recombinations. Proc Natl Acad Sci USA 105:15470–15474

Shibata F, Hizume M, Kuroki Y (1999) Chromosome painting of Y chromosomes and isolation of a Y chromosome-specific repetitive sequence in the dioecious plant *Rumex acetosa*. Chromosoma 108:266–270

Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP et al (2011) The genome of woodland strawberry (*Fragaria vesca*). Nat Genet 43:109–118

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T et al (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423:825–837

Slotkin R, Nuthikattu S, Jiang N (2012) The impact of transposable elements on gene and genome evolution. In: Wendel JF (ed) Plant genome diversity, vol 1, Plant genomes, their residents, and their evolutionary dynamics. Springer, Wien, New York

Steinemann M, Steinemann S (1992) Degenerating Y chromosome of *Drosophila miranda*: a trap for retrotransposons. Proc Natl Acad Sci USA 89:7591–7595

Steinemann S, Steinemann M (2005) Y chromosomes: born to be destroyed. Bioessays 27:1076–1083

Sweredoski M, DeRose-Wilson L, Gaut BS (2008) A comparative computational analysis of nonautonomous helitron elements between maize and rice. BMC Genomics 9:467

Swigonová Z, Bennetzen JL, Messing J (2005) Structure and evolution of the r/b chromosomal regions in rice, maize and sorghum. Genetics 169:891–906

Tenaillon MI, Hollister JD, Gaut BS (2010) A triptych of the evolution of plant transposable elements. Trends Plant Sci 15:471–478

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

Tian Z, Rizzon C, Du J, Zhu L, Bennetzen JL, Jackson SA, Gaut BS, Ma J (2009) Do genetic recombination and gene density shape the pattern of DNA elimination in rice long terminal repeat retrotransposons? Genome Res 19:2221–2230

Topp CN, Zhong CX, Dawe RK (2004) Centromere-encoded RNAs are integral components of the maize kinetochore. Proc Natl Acad Sci USA 101:15986–15991

Tsubota SI, Huong DV (1991) Capture of flanking DNA by a P element in *Drosophila melanogaster*: creation of a transposable element. Proc Natl Acad Sci USA 88:693–697

Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A et al (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313:1596–1604

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

Vitte C, Bennetzen JL (2006) Analysis of retrotransposon structural diversity uncovers properties and propensities in angiosperm genome evolution. Proc Natl Acad Sci USA 103:17638–17643

Vitte C, Panaud O (2003) Formation of solo-LTRs through unequal homologous recombination counterbalances amplifications of LTR retrotransposons in rice *Oryza sativa* L. Mol Biol Evol 20:528–540

Vitte C, Panaud O (2005) LTR retrotransposons and flowering plant genome size: emergence of the increase/decrease model. Cytogenet Genome Res 110:91–107

Volpe TA, Kidner C, Hall IM, Teng G, Grewal SIS, Martienssen RA (2002) Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science 297:1833–1837

Vyskot B, Hobza R (2004) Gender in plants: sex chromosomes are emerging from the fog. Trends Genet 20:432–438

Wagner A (2006) Periodic extinctions of transposable elements in bacterial lineages: evidence from intragenomic variation in multiple genomes. Mol Biol Evol 23:723–733

Walbot V, Petrov DA (2001) Gene galaxies in the maize genome. Proc Natl Acad Sci USA 98:8163–8164

Wang H, Liu J-S (2008) LTR retrotransposon landscape in *Medicago truncatula*: more rapid removal than in rice. BMC Genomics 9:382

Weil C, Martienssen R (2008) Epigenetic interactions between transposons and genes: lessons from plants. Curr Opin Genet Dev 18:188–192

Whitney KD, Baack EJ, Hamrick JL, Godt MJW, Barringer BC, Bennett MD, Eckert CG, Goodwillie C, Kalisz S, Leitch IJ et al (2010) A role for nonadaptive processes in plant genome size evolution? Evolution 64:2097–2109

Wicker T, Keller B (2007) Genome-wide comparative analysis of copia retrotransposons in Triticeae, rice, and *Arabidopsis* reveals conserved ancient evolutionary lineages and distinct dynamics of individual copia families. Genome Res 17:1072–1081

Wicker T, Yahiaoui N, Guyot R, Schlagenhauf E, Liu Z, Dubcovsky J, Keller B (2003) Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and Am genomes of wheat. Plant Cell 15:1186–1197

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

Wicker T, Taudien S, Houben A, Keller B, Graner A, Platzer M, Stein N (2009) A whole-genome snapshot of 454 sequences exposes the composition of the barley genome and provides evidence for parallel evolution of genome size in wheat and barley. Plant J 59:712–722

Witte CP, Le QH, Bureau T, Kumar A (2001) Terminal-repeat retrotransposons in miniature (TRIM) are involved in restructuring plant genomes. Proc Natl Acad Sci USA 98:13778–13783

Wright SI, Le QH, Schoen DJ, Bureau TE (2001) Population dynamics of an Ac-like transposable element in self- and cross-pollinating arabidopsis. Genetics 158:1279–1288

Wright SI, Agrawal N, Bureau TE (2003) Effects of recombination rate and gene density on transposable element distributions in *Arabidopsis thaliana*. Genome Res 13:1897–1903

Yan X, Martínez-Férez IM, Kavchok S, Dooner HK (1999) Origination of Ds elements from Ac elements in maize: evidence for rare repair synthesis at the site of Ac excision. Genetics 152:1733–1740

Yang L, Bennetzen JL (2009) Distribution, diversity, evolution, and survival of Helitrons in the maize genome. Proc Natl Acad Sci USA 106:19922–19927

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

Yang G, Zhang F, Hancock CN, Wessler SR (2007) Transposition of the rice miniature inverted repeat transposable element mPing in *Arabidopsis thaliana*. Proc Natl Acad Sci USA 104:10962–10967

Yang G, Nagel DH, Feschotte C, Hancock CN, Wessler SR (2009) Tuned for transposition: molecular determinants underlying the hyperactivity of a stowaway MITE. Science 325:1391–1394

Zeyl C, Bell G, Green DM (1996) Sex and the spread of retrotransposon Ty3 in experimental populations of *Saccharomyces cerevisiae*. Genetics 143:1567–1577

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

Zhang X, Feschotte C, Zhang Q, Jiang N, Eggleston WB, Wessler SR (2001) P instability factor: an active maize transposon system associated with the amplification of tourist-like MITEs and a new superfamily of transposases. Proc Natl Acad Sci USA 98:12572–12577

Zhong CX, Marshall JB, Topp C, Mroczek R, Kato A, Nagaki K, Birchler JA, Jiang J, Dawe RK (2002) Centromeric retroelements and satellites interact with maize kinetochore protein CENH3. Plant Cell 14:2825–2836

Zou S, Ke N, Kim JM, Voytas DF (1996) The *Saccharomyces* retrotransposon Ty5 integrates preferentially into regions of silent chromatin at the telomeres and mating loci. Genes Dev 10:634–645

Author Queries

Chapter No.: 2

Query Refs.	Details Required	Author's response
AU1	As per style, we have to retain only one author as corresponding author. Please check and confirm.	
AU2	Figure 3 is not clear for print. Please check and provide me revised figure along with your proof corrections.	
AU3	Please check the page range in references "Kunze and Starlinger (1989), Kordis (2005), Lander et al. (2001), Langley et al. (1988), Le Rouzic et al. (2007), Leitch et al. (2005), Levis et al. (1984), Lippman et al. (2004), Walbot and Petrov (2001)".	