

An eruption of LTR retrotransposons in the autopolyploid genomes of *Chrysanthemum nankingense* (Asteraceae)

Haibin Wang (✉ hb@njau.edu.cn)

Nanjing Agricultural University - Weigang Campus: Nanjing Agricultural University <https://orcid.org/0000-0001-9476-5204>

Jun He

Nanjing Agricultural University - Weigang Campus: Nanjing Agricultural University

Zhongyu Yu

Nanjing Agricultural University - Weigang Campus: Nanjing Agricultural University

Jiafu Jiang

Nanjing Agricultural University - Weigang Campus: Nanjing Agricultural University

Sumei Chen

Nanjing Agricultural University - Weigang Campus: Nanjing Agricultural University

Weimin Fang

Nanjing Agricultural University - Weigang Campus: Nanjing Agricultural University

Zhiyong Guan

Nanjing Agricultural University - Weigang Campus: Nanjing Agricultural University

Yuan Liao

Nanjing Agricultural University - Weigang Campus: Nanjing Agricultural University

Fadi Chen

Nanjing Agricultural University - Weigang Campus: Nanjing Agricultural University

Research

Keywords: Autopolyploidization, Genome duplication, Evolution, Chrysanthemum, Asteraceae

Posted Date: April 5th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-382794/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background

Whole genome duplication, associated with the induction of widespread genetic changes, has played an important role in the evolution of many plant taxa. The majority of extant angiosperm species have undergone at least one polyploidization event, forming either an auto- or allopolyploid organism. Compared with allopolyploidization, however, few studies have examined autopolyploidization, and almost no studies have focused on the response of genetic changes to autopolyploidy.

Results

In the present study, newly synthesized *C. nankingense* autotetraploids (Asteraceae) were employed to characterize the genome shock following autopolyploidization. Available evidence suggested that the genetic changes primarily involved the loss of old fragments and the gain of novel fragments, and some novel sequences were potential long terminal repeat (LTR) retrotransposons. As *Ty1-copia* and *Ty3-gypsy* elements represent the two main superfamilies of LTR retrotransposons, the dynamics of *Ty1-copia* and *Ty3-gypsy* were evaluated using RT-PCR, transcriptome sequencing and LTR retrotransposon-based molecular marker techniques. These results suggest that autopolyploidization might also be accompanied by perturbations of LTR retrotransposons, and the emergence retrotransposon insertions might show more rapid homologue divergence, resulting in diploid-like behaviour, potentially accelerating the evolutionary process among progenies.

Conclusions

Our results strongly suggest a need to expand current evolutionary framework to encompass a genetic dimension when seeking to understand genomic shock following autopolyploidization in Asteraceae.

Introduction

Whole genome duplication (WGD) or polyploidization is widely recognized as a significant driver of higher plant evolution[1]. The majority of extant angiosperm species have undergone at least one polyploidization event, forming either an auto- or an allopolyploid species. In addition to polyploidization, some species long believed to be genuine diploids are actually pale-polyploids and have originated multiple times via hybridization and polyploidization throughout their evolutionary history[2, 3]. Arabidopsis[4, 5], rice [6]and cotton[7] are all good examples of pale-polyploids. The question of why polyploids have been so successful has generated important research activity in the past decade. Significant progress towards an understanding of the polyploid genome structure and evolution in both the short-term (synthesized polyploid) and the long-term (natural polyploid) has been made[1, 8–10].

Polyploid species can be formed by one progenitor species (i.e., autopolyploids) or divergent parental lineages (i.e., allopolyploids, via one-step or stepwise hybridization and polyploidization) through diverse mechanisms in natural populations. However, in either case, new combinations of redundant genes originate from different or similar genomes, potentially resulting in novel traits that do not exist in the contributing species, such as increasing resistance to levels of heat/cold, waterlogging/drought and salt/alkali tolerance, disease/pest resistance, and flowering time/organ size changes that might enable the offspring to thrive in harsh environments[11]. These changes have been detected in several diverse synthetic allopolyploids and might be associated with genomic perturbation at both the genetic and epigenetic levels. Several mechanisms that might affect these (genetic or epigenetic) alterations have occurred following polyploidization, for which the main manifestations include chromosomal rearrangement, the gain and loss of chromosome segments,

gene repression and activation, and changes in the epigenome, particularly with respect to patterns of cytosine methylation[8, 10]. These mechanisms might buffer the disadvantages of redundant genome and might provide starting points for the evolution of new lineages under “genomic shock”, according to McClintock[12, 13].

Retrotransposons, as mobile genetic elements, have attracted considerable attention. These elements can move, and occasionally spread, within the host genome through the reverse transcription of an RNA intermediate[14–16]. In many cases, genome-sized polymorphisms might reflect the activity of retrotransposons, particularly the copy number differences among long terminal repeat (LTR) retrotransposons, which can achieve exceptionally high copy numbers and often occupy over 50% of the nuclear DNA content, although these elements are typically transcriptionally silent[17]. *Ty1-copia* and *Ty3-gypsy* elements represent two major superfamilies of LTR retrotransposons in plants with similar structural features, e.g., flanking LTRs in direct orientation and multiple domains involved in different steps of element transposition[18]. The LTR retrotransposon replication cycle involves mobilization via a reverse transcriptase (RT)-mediated copy-and-paste mechanism, whereby individual elements serve as transcriptional templates for the synthesis of mRNAs, which are reverse transcribed prior to integration at new positions as double-stranded DNA molecules[17].

Because of the repressive chromatin modifications mediated through RNA-directed DNA methylation, LTR retrotransposon activity is typically epigenetically silenced. However, under certain biotic and abiotic phenomena, host repression might fail, resulting in large-scale LTR retrotransposon activation and/or proliferation[19–21]. Recent studies have provided valuable insights into the LTR retrotransposon dynamics in allopolyploids between distantly related genomes, which might be restricted to some specific elements and/or model systems. For example, the mobility of the sunfish transposon was detected in *Arabidopsis thaliana* × *Arabidopsis arenas* synthetic allotetraploids[22], where retrotransposons showed strong transcriptional activation in several hybrid sunflowers (*Helianthus* species), potentially reflecting interspecific hybridization events[23, 24]. In rice (*Oryza sativa*), introgressive hybridization from the wild species *Zizania latifolia* might be correlated with retrotransposon proliferation[25]. Remarkably, the Wis2 retrotransposon exhibits strong transcriptional activation in wheat; however, there was no evidence of Wis2 retrotransposon proliferation in response to allopolyploidization in synthetic allopolyploids (*Aegilops sharonensis* × *Triticum monococcum*)[26] and no evidence of mobility in response to allopolyploidization in natural *Brassica* species[27, 28]. Hence, species differences in various families might affect the proportion and categories of retrotransposons. Studies have shown that genome rearrangement, the gain and loss of chromosome segments, and retrotransposon mobility might be consistent with expectations of conflict resolution in allopolyploids derived from genetically divergent parents in several species groups. However, few autopolyploids have been examined, and almost no studies have focused on the responses of genome changes and retrotransposon polymorphisms in autopolyploidization. As genome shock might expectedly be minimized in young autopolyploids originating from slightly divergent progenitors, this phenomenon has limited effects on the global genome structure and yields limited genome reorganization in some species[29]. The proportion and categories of DNA fragments affected by the species differs in various families under polyploidization. To promote a better understanding of the success of polyploidization, further independent polyploidization events (particularly autopolyploidization) should be analysed in future studies.

The tribe Anthemideae Cass of the Asteraceae is considered an advanced subject at the forefront of evolution. As one of the largest and most diverse families of flowering plants, Anthemideae has at least 25,000 dominant species[30]. Polyploidization is a prominent driver of evolution among angiosperms[1], affecting not only the current distribution and success but also the evolution of these species. Coincidentally, polyploidization generally occurs with greater frequency among Asteraceae species than among other angiosperm families, and this increased frequency also provides an ideal opportunity for investigations of the rapid adaptations associated with polyploidization[31–33]. The Asteraceae genus *Chrysanthemum* harbours several polyploidization species. Considerable variations at the ploidy level have been observed in this genus (from $2n = 2x = 18$, to $2n = 36, 54, 72$, up to 90)[33–35]. In previous studies, we investigated the genomic, transcriptomic and epigenomic alterations in allopolyploids of this genus[33, 36–38]. Genome-wide transcriptomic

alterations were detected in a set of synthetic allohexaploids. Rapid sequence elimination is common in *Chrysanthemum morifolium* × *Leucanthemum paludosum*, *Chrysanthemum nankingense* × *Tanacetum vulgare* and *Chrysanthemum crassum* × *Crossostephium chinense* amphihaploids and amphidiploids[10, 37]. Some polymorphism fragments of these plants were actually microsatellite sequences, suggesting that allopolyploidization might be accompanied by defects in homologous recombination and DNA damage repair mechanisms[38]. Cytosine methylation machinery responses might act as a "genome defence" system in *Chrysanthemum* spp. during allopolyploid formation, potentially reflecting a balance between the increased activity of *MET1* in the higher ploidy genomes and the larger number of CpG dinucleotide sites available for methylation [10, 33]. These findings suggest the need to expand the current evolutionary framework to encompass the genetic/epigenetic dimension when seeking to understand allopolyploidization.

In a previous study, we performed a cytogenetic analysis of these autotetraploids and observed no evidence of any chromosome elimination. Although four autotetraploid species appeared to be cytologically uniform and shared an identical phenotype, previous reports on the simple sequence repeat (SSR) analysis revealed slight polymorphic variations between these plants[38]. To a certain degree, microsatellite changes might reflect the genetic variation following autopolyploidization. Considering that autopolyploidy also plays an important role in the evolution of polyploidization origins and the rate of autopolyploid formation is higher than the number of allopolyploids, few samples were examined to characterize the response of genome changes following autopolyploidization[39]. Here, we used the *C. nankingense* diploid and its autotetraploids to examine rapid genome evolution following autopolyploidization using amplified fragment length polymorphism (AFLP) markers. Subsequently, we detected a *Ty1-copia* type retrotransposon with homology, suggesting that retrotransposons might activate the autopolyploidization of *C. nankingense*. In particular, we examined the characteristics and transcriptional dynamics of the reverse transcription of *Ty1-copia* and *Ty3-gypsy* elements. The activity and quiescence of individual retrotransposon subfamilies were also distinguished using high-throughput sequencing. After further amplification of the specific LTR sequences flanking the two *Ty1-copia* elements, the accumulation of *Ty1-copia* elements in the autotetraploid *C. nankingense* was also evaluated.

Results

Alterations in the genome sequences of the newly synthesized autotetraploids

A total of four individual autotetraploid *C. nankingense* plants were employed in the present study. Ten AFLP primer combinations were used to amplify 422 fragments from diploid *C. nankingense*. The expectation was that autopolyploidization would have no effect on the AFLP genotype of autotetraploid compared with diploid species; however, the AFLP profiles revealed an obvious fragment variation in autotetraploid lines. For example, in autotetraploid line 1 (T1), the AFLP profiles consisted of 415 fragments, of which 6 novel fragments were observed, and 13 fragments present in diploid species were not detected following autopolyploidization (Fig. 1). In addition, the autotetraploid lines 2/3/4 consisted of 418/415/416 fragments, of which 3/5/3 fragments were not present in diploid species, and 7/12/9 fragments were lost in the three autotetraploid lines, respectively.

To obtain additional information, the 17 novel fragments specific to the autotetraploid profile were recycled and sequenced. Among these fragments, seven yielded poor sequence quality, and only ten fragments were successfully sequenced. We compared these ten sequences with known genes for general similarities and the presence of specific motifs in the NCBI database. The blast search revealed that only one fragment (belonging to autotetraploid (T1)) showed high similarity with a known rice gene (E-value = 6e-34) as an LTR retrotransposon (*Ty1-copia* type).

Cloning and analysis of the RT region of LTR retrotransposons

In LTR retrotransposons, the reverse transcriptase (RT) domain, encoding the enzymes responsible for the generation of a DNA copy from the genomic RNA template, is the major regulatory gene of the transposition cycle. Consequently, the replication of many transposons was first characterized through the PCR amplification of the RT domain. These fragments were subsequently shown to transpose. The process of reverse transcription was prone to error, suggesting the vast potential for variability of the RT domain, which also played a role in the evolution of the control of retrotransposon activity (i.e., transcription termination). Using two pairs of degenerated primers, the conserved fragments of the RT domains from *Ty1-copia* and *Ty3-gypsy* elements were PCR-amplified. After differential screening, sequencing and removing redundant sequences, a total of 76 *Ty1-copia*-RT-domain sequences (abbreviated as Ty1RT) and 50 *Ty3-gypsy*-RT-domain sequences (abbreviated as Ty3RT) were obtained from 232 positive clones.

The Ty1RT sequence variable length ranged from 216 to 401 bp. Among these, sequences 263 bp in length were primarily observed, accounting for 48.7% of the observed sequence lengths, followed by sequences 266 bp in length, accounting for one sixth of the former (8.0%). The Ty3RT sequence length ranged from 421 to 433 bp, and the length polymorphism was significantly lower than Ty1RT, with a vast majority of sequences (42/50, 84.0%) 432 bp in length, without considering the nucleotide polymorphisms. This analysis also revealed that all nucleotide sequences were AT-rich, with AT/GC ratios of 1.43 and 1.46 for Ty1RT and Ty3RT, respectively.

When translated into amino acid sequences using a bioinformatics approach, 40.8% (31/76) of Ty1RT and 28% (14/50) of Ty3RT sequences showed common abnormally processed transcripts, resulting in frame-shifts with premature termination through the introduction of stop codons. Further analysis showed that among the 37 total 263-bp, seven 266-bp and one 255-bp sequence Ty1RT sequences examined, only one 263-bp sequence (2.7%) showed early termination. Without exception, all other Ty1RT sequences showed abnormal transcription (Table 1). For Ty3RT, all eight non-432 bp genes harboured one or more stop codons in the sequence, while the proportion of 432-bp sequences was 16.7% (7/42).

Table 1

Analysis of sequence premature transcription termination and the fragment length of TyRT1

Fragment length	Premature transcription termination number/Total number	Percentage
263	1/37	2.7%
216	4/4	100%
221	1/1	100%
225	1/1	100%
226	2/2	100%
231	1/1	100%
232	4/4	100%
250	1/1	100%
254	2/2	100%
255	0/1	0%
261	1/1	100%
264	1/1	100%
265	1/1	100%
266	0/7	0%
274	1/1	100%
276	0/1	0%
302	1/1	100%
317	3/3	100%
332	1/1	100%
386	4/4	100%
401	1/1	100%

The remaining nucleotide sequences could be translated to amino acid sequences without any difficulties. For Ty1RT, 45 sequences had amino acid similarities of 80.1%, and all sequences contained conserved regions of the *Ty1-copia*-RT-domain, including the 5' end of "TAFLHG", the middle of "SLYGLKQ" and the 3' end of "YVDDM". The identity of the amino acid sequences of Ty1RT-20, Ty1RT-21, Ty1RT-23, Ty1RT-35 and Ty1RT-7 was 100%, suggesting that "TAFLHGQLKETVYVSQPDGFVDPEFPNHVYKLNKALYGLKQAPRAWYDKLSSFLIANNFTKGSVDPTLFIQYHGAILIVQIYVDDM" might be the original CnTy1RT sequence. The sequences of Ty1RT-39, Ty1RT-41 and Ty1RT-42 were also similar, indicating that these sequences were also in a primitive state in the *C. nankingense* genome (Figure S1). The sequence similarity of Ty3RT (approximately 90.0%) was higher than that of Ty1RT, and both of these sequences contained "RMCVDY" in the 5' region and "VMPFGL" in the middle region as conserved areas of the *Ty3-gypsy*-RT-domain. Although the length variation of Ty3RT was small, the amino acid variation was higher than that in Ty1RT, and only Ty3RT-14 and Ty3RT-26 were consistent at the amino acid level (Figure S2). The codon substitution and maximum likelihood models of dN (nonsynonymous) and dS (synonymous) were used for the reliable detection of positive selection at individual amino acid sites. Consequently, all Ty1RT sites showed no significant selection pressure, suggesting that these sites might maintain the gene function via the reduction of nonsynonymous mutations during the independent evolutionary events of Ty1RT

(Table S1). For sites 2, 6, 72, 87, 138, 139 and 142 of Ty3RT, the positive selective effects might reflect selection pressure during evolution (Table S2).

A WebLogo plot showed the amino acid frequencies at conserved regions within *C. nankingense*, and although most of these loci shared a highly conservative character and specificity, some sites showed similar stack heights, indicating a similar frequency of occurrence of each amino acid at that position. Comparatively fewer amino acid mutations were observed in CnTy1RT, and the stack height of each site was constant, with only limited thin stacks observed at sites 11 and 18, while in CnTy3RT, thin stacks were more common at sites 14, 36, 51, 70, 113, 114, 130, 131, 142, 143, etc. (Fig. 2a). However, completely different results were observed for the related Ty1RT and Ty3RT sequences from other species in NCBI. The thin stacks were more apparent in Ty1RT than in Ty3RT (Fig. 2b). A phylogenetic analysis of different species showed that all Ty1RT sequences could be divided among four clades, while all Ty3RT sequences could be divided into five clades (Fig. 2c). In the four clades of Ty1RT, the sequence from *C. nankingense* was classified into two groups (Ty1RT-42, -39, -41, -43, -44, -38, -40 and -22 vs. others). These two groups showed the most closely related similarity with the Ty1RT sequences from *Solanum lycopersicum*, *A. thaliana* and *Prunus mume*, indicating that the *Ty1-copia* elements of these species might exist in horizontal transmission (Fig. 2c). A similar pattern was also observed in Ty3RT, and all sequences could be further divided into three groups (Ty3RT-31 vs. Ty3RT-29, -5, -36, -7, -34 and -8 vs. the others) (Fig. 2c). These apparent clades indicated the proliferation of corresponding LTR retrotransposons in genome differentiation. A random analysis of the sequence homology of each group showed that the sequence similarities were ~ 63% for Ty1RT and 79% for Ty3RT. These findings were consistent with the WebLogo plot results.

Transcriptional activation of LTR retrotransposons following autopolyploidization

Many of the plant LTR retrotransposons are transcriptionally activated through various biotic and abiotic stress factors. Indeed, polyploidization is a strong stimulant for plants, causing genome-wide genetic perturbations, particularly in allopolyploids. In the present study, we employed diploid organisms (tissue-cultured and non-tissue-cultured) as controls to investigate the transcriptional activation of the RT domain of *C. nankingense* following autopolyploidization. The PCR amplicons were separated through polyacrylamide gel electrophoresis. The nearly invisible bands in tissue-cultured and non-tissue cultured diploid *C. nankingense* showed a complete turnaround, becoming obvious in all four individual autotetraploid plants, although LTR retrotransposon-like AFLP fragments were not detected in the autotetraploid line 2/3/4, suggesting that tissue culture could not induce the expression of Ty1RT and Ty3RT elements, while the transcriptional activation of LTR retrotransposons might result from autopolyploidization (Fig. 3a).

To further verify the transpositional activity following autopolyploidization, the autotetraploid *C. nankingense* line 1 (T1), which showed fragment polymorphisms in the AFLP genotype (fragment not present in diploid organisms) and had high similarity with known rice *Ty1-copia* elements, was employed to sequence the transcriptome using high-throughput sequencing. Compared with diploid species, the autotetraploid line (T1) showed the marked transcriptional regulation of approximately 60,000 detected genes, and 1/3 of these genes showed more than two-fold up or down regulation at the transcriptional level (Fig. 3b), involved in metabolism, cellular processes, enzyme activities, protein binding and other processes or functions (Figure S3). Further analysis revealed that a total of 285 retrotransposon fragments were detected in the diploid or tetraploid line, including gag protein and gag-pol protein, pol-poly protein, int protein, the RT domain and other proteins with the retrotransposon annotation. These annotations suggest the existence of retrotransposons; however, the type of retrotransposon remains unknown. The sequencing of the rice genome has recently been completed, and this information is valuable for understanding the composition and structural features of the plant genome. Therefore, we re-annotated the transcriptome sequence using rice genome databases and the expressed sequence tag [40] database. These 285 retrotransposon fragments were classed as 139 *Ty1-copia*-like, 118 *Ty3-gypsy*-like, 21 non-LTR type (SINE) elements, and 7 non-specific sequences (Fig. 3c).

A total of 104 *Ty1-copia*-like, 92 *Ty3-gypsy*-like and 13 non-LTR type (SINE) elements were observed in diploid lines, and most of these elements had a low reads per kilobase per million reads mapped (RPKM) values, showing > 100-fold lower levels than housekeeping genes, such as *EF1a* and *GAPDH*. Among the 104 *Ty1-copia*-like elements, the average RPKM value was only 1.4; however, following autopolyploidization, the average RPKM value increased to 4.6, with 49% gene activation and 25% gene silencing. In addition, 35 new *Ty1-copia*-like elements were first observed in autotetraploids. The gene activation was 3.3-fold higher than the gene silencing. A similar result was observed for the *Ty3-gypsy*-like and non-LTR type (SINE) elements, and the gene activation was 3-fold higher than the gene silencing (75/23 and 12/4, respectively). The average RPKM value of the *Ty3-gypsy*-like elements in autotetraploid (T1) increased to 2.2, a value 1.6-fold higher than that observed in diploid species (Fig. 3c).

LTR region analysis of *Ty1-copia*-like elements

Based on the transcriptome database, we selected three *Ty1-copia*-like elements with higher up-regulation expression alterations and high homology with known retrotransposons in the autotetraploid (T1) as target genes for cloning and sequence analysis. Unfortunately, only two of these target genes were successfully amplified, reflecting the high polymorphism of this subtype of retrotransposon. The 4,598 and 4,616 bp sequences were renamed as *CnMp1* and *CnMp2*, respectively.

Sequence analysis revealed that the *CnMp1* and *CnMp2* sequences contained conserved poly A sites in the LTR region. The boundary sequence (CA) of R-U5 was also detected among the 40 amino acid sequences of the LTR region. The PBS region also harboured the characteristic base sequence of "TGGT", followed closely by LTR regions at 3 and 0 bp in distance, respectively. For the coding regions, the GAG-PR-INT-RT-RNaseH genes were closely interlinked, and the GAG region contained the conserved nucleic acid-binding sequence "C-C-H-C". The PR region followed, and *CnMp1* contained the conserved "D (T/S) G" sequence, while the last amino acid of *CnMp2* was "A", which was not conserved. The INT regions of these two sequences had a conserved zinc finger domain (H-H-C-C). Three motifs, "AFLHG", "YGLKQ", and "YVDDM", in the RT region and two motifs, "KHID" and "KHID", in the RNaseH region were all relatively conserved (Fig. 4). The analysis revealed that the cloned sequence was a *Ty1-copia*-type retrotransposon.

SSAP analysis of the insertion sites in *CnMp1* and *CnMp2*

The LTR region was sequence-specific for a given retrotransposon. The insertion of a retrotransposon into the plant genome, as opposed to DNA transposons, did not result in sequence divergences, rather showing stable insertions in the genome that could be exploited using recently developed marker systems. The retrotransposon-based sequence-specific amplification polymorphism (SSAP) is a novel molecular marker technique based on insertion polymorphisms generated through the specificity of an oligonucleotide primer anchored on the terminal sequences of a retrotransposon, typically in the LTR region[41].

In the present study, the primers were derived from the LTR regions of *CnMp1* and *CnMp2*, reflecting the actual loci of *CnMp1* and *CnMp2* in the autotetraploid (T1) genome. A total of fourteen primer combinations (seven combinations for each gene) produced a multi-fragment SSAP profile. Most of the fragments of diploid species inherited by the autotetraploid (T1) line; however, five of the *CnMp1* primer combinations generated at least one polymorphic fragment, while only two *CnMp2* primer combinations showed certain polymorphisms, generating up to 12 polymorphic fragments, of which 8 fragments (seven fragments for *CnMp1* and one fragment for *CnMp2*) were inserted into new loci, and 4 fragments (three fragments for *CnMp1* and one fragment for *CnMp2*) showed the loss of the original loci, suggesting the activation and insertion of the retrotransposon following autopolyploidization (Fig. 5a and 5b). Notably, although the transcript abundance of *CnMp1* and *CnMp2* in the autotetraploid (T1) was similar (RPKM 6.9 vs. 7.6), and *CnMp2* was also transcribed in the diploid organism (RPKM = 3.0), the total and polymorphic numbers of SSAP fragments of *CnMp1*

were higher than those of *CnMp2*, suggesting that *CnMp1* might be a relatively young retrotransposon, and a mechanism for the inhibition of *CnMp2* insertion might exist in *C. nankingense*.

Discussion

Effect on the genome sequence following autopolyploidization

Polyploidization has been considered a pervasive force in plant evolution. The subsequently formed polyploids are divided into two categories: autopolyploids, resulting from duplication of the same genome; and allopolyploids, resulting from hybridization and duplication between two related genomes, merging into a common nucleus[1]. The available data indicate considerable genomic rearrangements, including exchanges between genomes and gene loss, following allopolyploidization [1]. In the early stages of the artificial synthesis of allopolyploids between wheat and its near relatives, sequence elimination is a major immediate response of the wheat genome to hybridization or allopolyploidization[42]. Similar measurable degrees of sequence elimination were also generated in cucumber[43], cabbage[44] and salsify[45]. Although remaining elusive and receiving little attention compared with allopolyploidization, the limited currently available evidence suggests that autopolyploidization might also effect genome structure via a series of genome-wide genomic perturbations, which might rapidly change and play an influential role in genome reorganization. In the present study, the *C. nankingense* autotetraploid showed 1.6%-3.0% fragment elimination and 0.7%-1.4% novel fragment emergence. Although this finding was inconspicuous and not larger than the scope of *Chrysanthemum* allopolyploidization in our previous investigation[10, 37], the results also reflect rapid genetic changes in the *C. nankingense* autotetraploid to a certain extent.

Compared with the predominant formation of bivalents in allopolyploids (AABB), autopolyploids (AAAA) might easily show multivalent formation during meiosis. Clausen et al.[46] and Stebbins[47] proposed that the abnormal meiosis (e.g., multivalent formation) and reduced fertility compared with the diploid progenitor in autopolyploids might represent an evolutionary disadvantage. Instead, the successful range expansion and radiation demonstrated in various natural autopolyploids were much more common than traditionally anticipated, suggesting genome multiplication[48]. The benefit was obvious for autopolyploidization; thus, we hypothesized that the rapid genetic changes might provide a short-term evolutionary channel to increase the genome flexibility of autopolyploids via the accumulation of minor differences in the chromosome during early generation, and the cumulative effects of variations in genetic redundancy (potential reversion to a diploid-like status) might represent an evolutionary advantage on a longer evolutionary timescale.

Genetic heterogeneity of the LTR retrotransposon

The phenomenon of retrotransposon activation has been widely observed in plant evolution. In the present study, we cloned the RT domains of *Ty1-copia* and *Ty3-gypsy* elements. The transposition of retrotransposons in most eukaryotic organisms requires RNA intermediates via DNA-RNA-DNA models[17]. In this process, retrotransposons must encode reverse transcriptases and replicate through RNA intermediates, providing a key to identify the evolutionary relationships among different retrotransposons, as RTs do not have rigorous proofreading functions (substitution frequencies of approximately 10^{-4} per base) and could be influenced through many factors, including mutations, retrotransposition cycles, pressure selection, differentiation, population size, competition, etc., resulting in high base mutations, deletions, insertions, substitutions, frameshifts and genetic rearrangements, particularly in the RT domain[40]. Among the 121 Ty1RT clones and 102 Ty3RT clones obtained in the present study, the length of the Ty1RT sequence ranged from 216–401 bp, where 48.7% of sequences were 263 bp in length, and the length of the Ty3RT sequence ranged from 421–433 bp, where 84.0% of the sequences were 432 bp in length, showing a large sequence length variation of Ty1RT compared with Ty3RT.

In addition, 40.8% of Ty1RT and 28% of Ty3RT sequences showed premature termination, indicating that early termination explained the reverse transcription inhibition. In addition, the likelihood of the premature termination of the 263-bp Ty1RT

sequence was low (2.7%), and all seven 266-bp Ty1RT sequences and only one 255-bp Ty1RT sequence were observed, while the other Ty1RT sequences harboured one or more stop codons (Table 1). Similarly, all non-432-bp Ty3RT sequences showed abnormal transcription; however, the probability of the premature termination of the 432-bp sequence was 16.7%, 6.4-fold higher than that observed for the 263-bp Ty1RT sequences. The lower length variation (51.3% non-263 bp Ty1RT vs. 16.0% non-432 bp Ty3RT) and high base mutation in the Ty3RT sequence indicated that mutations contribute to Ty3RT premature termination. Based on the maximum likelihood codon neutrality analysis of Ty3RT, a total of seven codons showed evolutionary selection, i.e., Ty3RT might suffer certain selection pressure and undergo mutations during the evolutionary process[49], resulting in premature termination, and among these basic mutations, sequence length variations might directly reflect the heterogeneous nature of the two retrotransposon types.

Retrotransposons are more ancient transposable elements, widely distributed throughout the plant kingdom, with a high degree of heterogeneity¹⁶. As the prevalent component of most plant genomes, these elements can be inherited from parent to offspring (vertical transmission) and become transferred between different species (transverse transfer)[50]. Cluster analysis revealed that either Ty1RT or Ty3RT series harbour a certain similarity with other species, suggesting the lateral transfer of retrotransposons with a common origin among these species. Comparatively, the amino acid polymorphisms in Ty1RT were fewer than those in Ty3RT in *C. nankingense*; however, within different species, Ty3RT was more conserved than Ty1RT. This considerable variation of Ty1RT in different species might reflect sequence diversification after original species differentiation, while for Ty3RT, the amino acid variations might not be as dynamic as that of Ty1RT during *C. nankingense* evolution, resulting in populations of highly similar sequences of quiescent sequence lengths, particularly under a certain pressure selection.

Transcriptional activity of retrotransposons

Retrotransposons significantly affect the genetic stability of the plant genome, particularly regarding the contribution of these elements to the large-scale structure of complex genomes[15, 51]^{15,49}. Although, silenced in the long-term in a vast majority of cases, many of the plant retrotransposons could be transcriptionally activated through various factors. Polyploidization was an extremely strong shock and could activate quiescent transposons/retrotransposons, with some mutations in allopolyploid species[52, 53]. For example, the *copia*-like wheat retrotransposon Wis2-1A can be induced through allotetraploidization, affecting the expression of adjacent genes[26]. New insertion sites of retrotransposon *TNT1* were detected in synthetic allotetraploid tobacco[54]. Long fragment genome sequencing analyses of maize ancient polyploid genomes showed that the activation and amplification of LTR-retrotransposons resulted in an explosive increase in the maize genome size from approximately 1200 to 2400 Mb in the past three million years[55], while in cotton and the allotetraploid *Triticum dicoccoides*, a high level of intergenomic invasion of retrotransposons in the early stages of polyploid formation might reflect general tendencies in speciation and the stabilization of the allopolyploid genome[56, 57]. In the present study, polymorphic AFLP fragments matched the *Ty1-copia* type rice retrotransposon sequence, suggesting that autopolyploidization might also accompany the perturbations of retrotransposons.

Using RT-PCR and transcriptome sequencing, the retrotransposon activity in the *C. nankingense* autotetraploid was examined in the present study. All detected retrotransposons could be divided into three categories and classed according to the rice genome, primarily showing *Ty1-copia*-like elements, followed by *Ty3-gypsy*-like elements and several non-LTR type (lines) elements. Retrotransposons are typically silenced; however, studies have shown the transcription of a large proportion of retrotransposons in the human genome, of which at least 120 retrotransposon elements have functional implications or have evolved into bona fide genes[58]. As an indicator of functionality, we used the ratio of nonsynonymous substitutions over synonymous substitutions per site (dN/dS) to evaluate the positive selection at individual amino acid sites. We observed that the Ty1RT sites had no significant selection pressure, suggesting that these sites were predominantly shaped through purifying selection throughout evolution, representing potential transcriptional activity and the evolution of functional genes under selective constraints. In the present study, although with a low expressive abundance, some transcriptional retrotransposons were observed in diploid lines, whether these elements profit

from the genomic environment for transcription and functionality without generating transcriptional "noise" warrants further study. More importantly, following autopolyploidization, the retrotransposon activation (approximately 3-fold higher than the silenced condition) and increases in RPKM in *Ty1-copia*-like or *Ty3-gypsy*-like elements and the number and proportion of *Ty1-copia*-like elements were higher than those of the *Ty3-gypsy* element. Therefore, retrotransposons might participate in the genome evolution of autotetraploid *C. nankingense*, thereby confirming initial hypotheses.

Dynamics of retrotransposons following autopolyploidization

Retrotransposons are reverse transcribed through a self-encoded reverse transcriptase, and the resulting DNA is integrated into the genome. This "copy and paste" mode of transposition increases the copy number of these elements, likely providing direct evidence of retrotransposon-mediated variations in genome size within the plant genome[15, 51]. The SSAP results suggested that activated *CnMp1* and *CnMp2* insert into the new loci of the autotetraploid *C. nankingense* genome, indicating that autopolyploidization leads to retrotransposon activation and insertion (Fig. 5). It has recently been suggested that retrotransposons influence genome evolution through a high degree of adaptive or evolutionary potential in angiosperms, resulting in an extraordinary array of genetic changes, including gene modifications, duplications, altered expression patterns, and exaptation to create novel genes, with occasional gene disruptions[15, 59, 60]. Compared with allopolyploids, the bottleneck of abnormal meiosis for autopolyploids was amplification, and to restore gamete fertility, the emergence retrotransposon insertion might show faster homologue divergence, resulting in diploid-like meiotic behaviour realized in limited generation, whereas the evolutionary potential of retrotransposon insertion can be benefited for thousands or millions of years[15].

Asteraceae have high frequencies of hybridization and polyploidization (often at the same time), and the offspring of hybridization and polyploidization frequently show retrotransposon activation-insertion and beneficial character variations[30]. In *Achillea borealis* (Asteraceae), following polyploidization, tetraploid species showed ecological tolerance and occupied new habitats previously unassailable by diploid species[48]. Among three diploid sunflower (*Helianthus*) hybrids, retrotransposons were enhanced and inserted into the offspring genome following hybridization, resulting in a 50% increase in drought and salt tolerance compared with either diploid parent^{21,22}. We previously observed higher resistance to cold, drought and salt stress in the autotetraploid *C. nankingense* compared with diploid lines[61], while in the present study, the SSAP profiles supported dramatic retrotransposon insertions in the genome evolution of the autotetraploid *C. nankingense*. As retrotransposons are abundant and dispersed throughout genomes, these elements might result in the rapid shuffling of newly formed Asteraceae polyploid species via the contribution of abundant raw material at a critical moment in adaptive evolution, ultimately leading to the spectacular success of Asteraceae. Accordingly, future studies addressing the evolutionary consequences of polyploidization should focus on the specific evolutionary impact of retrotransposons.

Materials And Methods

Plant material and DNA extraction

The diploid ($2n = 2x = 18$) and four autotetraploid ($2n = 4x = 36$) *C. nankingense* lines were maintained at the Chrysanthemum Germplasm Resource Preserving Centre, Nanjing Agricultural University, China (32°05'N, 118°8'E, 58 m elevation). All plants were propagated through cuttings, and the medium contained a 2:2:1(v/v) mixture of perlite, vermiculite and leaf mould, respectively. Rooted seedlings were grown in a greenhouse at 22°C during the day and at least 15°C during the night, with a relative humidity of 70–75% under natural light. The experiments included three biological replications.

The genomic DNA of diploid and autotetraploid *C. nankingense* species was extracted using a cetyltrimethyl ammonium bromide (CTAB)-based method[62]. The DNA was treated with RNase to remove any RNA contamination. The

concentration and purity of the DNA preparations were measured using a NanoDropHND-1000 spectrophotometer (Nanodrop Technologies, USA), and the OD260/OD280 ratio was calculated.

AFLP fingerprinting

The DNAs were subjected to AFLP profiling according to Vos *et al* [63]. Approximately 500 ng of DNA was digested overnight at 37°C with 10 U each of *EcoRI* and *MseI* (New England Biolabs, Beijing, China). Subsequently, adaptor sequences were ligated onto the ends of the digested fragments at 16°C for 4 h, and the ligation mix contained the digested fragments, 5 pmol *EcoRI*, 50 pmol *MseI* adaptors (Table S3) and 4 U T4 DNA polymerase (NEB). The enzymes were subsequently inactivated at 70°C for 15 min. The amplicons generated from this template based on *EcoRI* and *MseI* primers lacking any selective bases (Table S3) were diluted 1:30 in ddH₂O and were used as templates for subsequent selective amplifications based on *EcoRI* and *MseI* primers carrying three selective bases, totalling nine primer combinations of *EcoRI* selective primer #2 plus *MseI* selective primer #5 (abbreviated 'E2 + M5'), E2 + M6, E3 + M2, E4 + M3, E4 + M8, E6 + M7, E7 + M3, E8 + M3 and E8 + M7 (Table S3). The *EcoRI* primers were labelled with 5-FAM. Each PCR and electrophoretic separation was repeated twice. The amplicons were separated using an ABI 377 DNA Analyzer, and only fragments ranging from 80 to 450 bp were considered.

Isolation and sequencing of informative AFLP amplicons

Informative AFLP amplicons were also separated through electrophoresis on 6% denaturing polyacrylamide gels, followed by silver staining[10, 63]. Autotetraploid-specific AFLP fragments were cut from the gel and eluted in 50 µL of ddH₂O after boiling for 10 min. After centrifugation (12,000 g, 5 min), an aliquot of the supernatant was collected and served as the template. The fragment was re-amplified using the same primer combination and PCR conditions described above. The resulting amplicon was inserted into the cloning vector pMD19 TA (Takara) and sequenced. The similarity of these sequences to those in public databases was obtained using BLAST analysis (<http://www.ncbi.nlm.nih.gov/BLAST/>).

Isolation of the RT region and transcriptional analysis of the LTR retrotransposon

Total RNA from diploid (tissue cultured and non-tissue cultured) and autotetraploid *C. nankingense* species was isolated using the Total RNA Isolation System (Takara) according to the manufacturer's instructions. Prior to reverse transcription, 30 ng of RNA was treated with 10 U of RNase-free DNAase I (Takara) at 37°C for 30 min to remove any contaminating genomic DNA. A 1-µL aliquot of the resulting RNA (containing approximately 600 ng) was used as the template for the synthesis of first-strand cDNA using SuperScriptIII Reverse Transcriptase (Takara), primed with random hexamer primers. The subsequent PCR involved two degenerate primer pairs targeting the RT regions of *Ty1-copia* (F: ACNGCNTTYTNCAYGG, R: ARCATRTRCRTCACRTA)[17] and *Ty3-gypsy* (F: AGMGRATGTGYGTSGAYTAT, R: CAMACCMRAAMWCACAMTT)[64] retrotransposons. Subsequent phylogenetic (homologue sequences of other species were the most related sequences in the NCBI database according to the BlastX results) and dN (nonsynonymous)/dS (synonymous) analyses were performed using MEGA 5.0 software (<http://www.megasoftware.net/mega.php>). Diploid DNA was used as a positive control, and the transcription profiling of the RT regions of *Ty1-copia* and *Ty3-gypsy* retrotransposons in diploids (tissue cultured and non-tissue cultured) and autotetraploids was performed using denaturing polyacrylamide gels.

Total RNA from diploid and selected autotetraploid *C. nankingense* species was isolated using the Total RNA Isolation System (Takara) according to the manufacturer's instructions. The RNA quality (RNA Integrity Number (RIN) > 8.5 and 28S:18S > 1.5) was verified using a 2100 Bioanalyser RNA Nanochip (Agilent, Santa Clara, CA), and the concentration was determined using an ND-1000 Spectrophotometer (NanoDrop, Wilmington, DE). The applied standards were 1.8 < OD260/280 < 2.2 and OD260/230 > 1.8. Illumina (San Diego, CA) sequencing was performed using an Illumina HiSeq™ 2000 device at the Beijing Genomics Institute (Shenzhen, China; <http://www.genomics.cn/index.php>) according to the

manufacturer's protocols. Transcriptome assembly was achieved using the short-read assembly program Trinity[65]. Functional annotation was assigned using the protein (Nr and Swiss-Prot), Clusters of Orthologous Groups (COG) and Gene Ontology (GO) databases. BLASTX was employed to identify related sequences in the protein databases based on E-values of less than 10^{-5} [66]. In addition, all transcriptome sequences were re-annotated using the rice genome databases or the EST database.

Isolation of LTR regions and SSAP fingerprinting

Three *Ty1-copia*-like elements with higher up-regulation expression alterations and high homology with the known retrotransposons in the autotetraploid (T1) were selected as target genes to clone and analyse the sequence using thermal asymmetric interlaced PCR (TAIL-PCR) [67] and ligation-mediated PCR (LM-PCR) methods[68]. The SSAP protocol was similar to the AFLP procedure of Vos et al. [63], except for the LTR-specific primers (LSPs) used in the second amplification, which were designed using Primer 5.0. The sequences for CnMp1 and CnMp2 were 5'-TAGACCAAATTACTAACTACCC-3' (LSP1) and 5'-AGTCTTTTTCTTTCTCATTTCTTCTC-3' (LSP2), respectively. In the present study, LSP1 and LSP2 were labelled with 5-FAM and combined with the standard AFLP *Mse*I selective primers (M2, M3, M4, M5, M6, M7 and M8), respectively. Each PCR and electrophoretic separation was repeated twice. The amplicons were separated using an ABI 377 DNA Analyzer, and only fragments ranging from 80 to 450 bp were considered.

Declarations

Author Contributions

HW and FC planned and designed the research. HW, XQ, JH, JJ, SC, and WF performed experiments. HW, ZG and YL analyzed data. HW wrote the manuscript.

Funding

This research was supported by the National Key Research and Development Program of China (2018YFD1000401, 2020YFD1001105), the Natural Science Fund of Qinghai Province, China (2018-HZ-819), a Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

Availability of data and materials

Please contact author for data requests for the datasets used and/or analyzed during the current study

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable

Competing interests

The authors declare that they have no conflict of interest.

References

1. Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS *et al.* Ancestral polyploidy in seed plants and angiosperms. *Nature* 2011, **473**(7345):97-100.

2. Garmendia A, Ferriol M, Benavent D, Ferrer-Gallego PP, Merle H. Intra- and inter-specific crosses among *Centaurea aspera* L. (Asteraceae) polyploid relatives-influences on distribution and polyploid establishment. *Plants (Basel)* 2020, **9**(9):1-16.
3. Soltis DE, Soltis PS. Polyploidy: recurrent formation and genome evolution. *Trends Ecol Evol* 1999, **14**(9):348-352.
4. Mandakova T, Joly S, Krzywinski M, Mummenhoff K, Lysak MA. Fast diploidization in close mesopolyploid relatives of *Arabidopsis*. *Plant Cell* 2010, **22**(7):2277-2290.
5. Quesneville H. Twenty years of transposable element analysis in the *Arabidopsis thaliana* genome. *Mob DNA* 2020, **11**(1):28.
6. Kamara N, Jiao Y, Lu Z, Aloryi KD, Wu J, Liu X, Shahid MQ. Cytological observations and bulked-segregant analysis coupled global genome sequencing reveal two genes associated with pollen fertility in tetraploid rice. *Int J Mol Sci* 2021, **22**(2).
7. Chen ZJ, Sreedasyam A, Ando A, Song Q, De Santiago LM, Hulse-Kemp AM, Ding M, Ye W, Kirkbride RC, Jenkins J *et al*. Genomic diversifications of five *Gossypium* allopolyploid species and their impact on cotton improvement. *Nat Genet* 2020, **52**(5):525-533.
8. Adams KL, Wendel JF. Polyploidy and genome evolution in plants. *Current Opinion in Plant Biology* 2005, **8**(2):135-141.
9. Van de Peer Y, Ashman TL, Soltis PS, Soltis DE. Polyploidy: an evolutionary and ecological force in stressful times. *Plant Cell* 2021, **33**(1):11-26.
10. Wang H, Jiang J, Chen S, Qi X, Fang W, Guan Z, Teng N, Liao Y, Chen F. Rapid genetic and epigenetic alterations under intergeneric genomic shock in newly synthesized *Chrysanthemum morifolium* x *Leucanthemum paludosum* hybrids (Asteraceae). *Genome Biol Evol* 2014, **6**(1):247-259.
11. Chen ZJ. Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annu Rev Plant Biol* 2007, **58**:377-406.
12. Leitch AR, Leitch IJ. Genomic plasticity and the diversity of polyploid plants. *Science* 2008, **320**(5875):481-483.
13. McClintock B. The significance of responses of the genome to challenge. *Science* 1983, **226**:792-801.
14. Huang Y, Chen H, Han J, Zhang Y, Ma S, Yu G, Wang Z, Wang K. Species-specific abundant retrotransposons elucidate the genomic composition of modern sugarcane cultivars. *Chromosoma* 2020, **129**(1):45-55.
15. Stapley J, Santure AW, Dennis SR. Transposable elements as agents of rapid adaptation may explain the genetic paradox of invasive species. *Mol Ecol* 2015, **24**(9):2241-2252.
16. Vicent C, Casacuberta J. Additional ORFs in Plant LTR-Retrotransposons. *Frontiers in Plant Science* 2020, **11**.
17. Kumar A, Bennetzen JL. Plant retrotransposons. *Annu Rev Genet* 1999, **33**:479-532.
18. Ramachandran D, McKain MR, Kellogg EA, Hawkins JS. Evolutionary dynamics of transposable elements following a shared polyploidization event in the tribe Andropogoneae. *G3 (Bethesda)* 2020, **10**(12):4387-4398.
19. Mhiri C, Parisod C, Daniel J, Petit M, Lim KY, Dorlhac de Borne F, Kovarik A, Leitch AR, Grandbastien MA. Parental transposable element loads influence their dynamics in young *Nicotiana* hybrids and allotetraploids. *New Phytol* 2019, **221**(3):1619-1633.
20. Pikaard CS, Mittelsten Scheid O. Epigenetic regulation in plants. *Cold Spring Harb Perspect Biol* 2014, **6**(12):a019315.
21. Belyayev A, Josefiowa J, Jandova M, Mahelka V, Krak K, Mandak B. Transposons and satellite DNA: on the origin of the major satellite DNA family in the *Chenopodium* genome. *Mob DNA* 2020, **11**:20.
22. Madlung A, Henkhaus N, Jurevic L, Kahsai EA, Bernhard J. Natural variation and persistent developmental instabilities in geographically diverse accessions of the allopolyploid *Arabidopsis suecica*. *Physiol Plant* 2012, **144**(2):123-133.
23. Kawakami T, Strakosh SC, Zhen Y, Ungerer MC. Different scales of *Ty1/copia*-like retrotransposon proliferation in the genomes of three diploid hybrid sunflower species. *Heredity (Edinb)* 2010, **104**(4):341-350.

24. Ungerer MC, Strakosh SC, Zhen Y. Genome expansion in three hybrid sunflower species is associated with retrotransposon proliferation. *Curr Biol* 2006, **16**(20):R872-873.
25. Wang N, Wang H, Wang H, Zhang D, Wu Y, Ou X, Liu S, Dong Z, Liu B: Transpositional reactivation of the *Dart* transposon family in rice lines derived from introgressive hybridization with *Zizania latifolia*. *BMC Plant Biol* 2010, **10**:190.
26. Kashkush K, Feldman M, Levy AA. Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat Genet* 2003, **33**(1):102-106.
27. Alix K, Heslop-Harrison JS. The diversity of retroelements in diploid and allotetraploid *Brassica* species. *Plant Mol Biol* 2004, **54**(6):895-909.
28. Alix K, Joets J, Ryder CD, Moore J, Barker GC, Bailey JP, King GJ, Pat Heslop-Harrison JS. The CACTA transposon Bot1 played a major role in *Brassica* genome divergence and gene proliferation. *Plant J* 2008, **56**(6):1030-1044.
29. Parisod C, Holderegger R, Brochmann C. Evolutionary consequences of autopolyploidy. *New Phytol* 2010, **186**(1):5-17.
30. Bremer K, Humphries C. Generic monograph of the Asteraceae-Anthemideae. *Bulletin of the Natural History Museum Botany* 1993, **23**.
31. Kane NC, Barker MS, Zhan SH, Rieseberg LH. Molecular evolution across the Asteraceae: micro- and macroevolutionary processes. *Mol Biol Evol* 2011, **28**(12):3225-3235.
32. Pellicer J, Garcia S, Canela MA, Garnatje T, Korobkov AA, Twibell JD, Valles J. Genome size dynamics in *Artemisia* L. (Asteraceae): following the track of polyploidy. *Plant Biol (Stuttg)* 2010, **12**(5):820-830.
33. Wang H, Qi X, Chen S, Fang W, Guan Z, Teng N, Liao Y, Jiang J, Chen F. Limited DNA methylation variation and the transcription of MET1 and DDM1 in the genus *Chrysanthemum* (Asteraceae): following the track of polyploidy. *Front Plant Sci* 2015, **6**:668.
34. Chen X, Wang H, Yang X, Jiang J, Ren G, Wang Z, Dong X, Chen F. Small-scale alpine topography at low latitudes and high altitudes: refuge areas of the genus *Chrysanthemum* and its allies. *Hortic Res* 2020, **7**(1):184.
35. He J, Lin S, Yu Z, Song A, Guan Z, Fang W, Chen S, Zhang F, Jiang J, Chen F *et al*. Identification of 5S and 45S rDNA sites in *Chrysanthemum* species by using oligonucleotide fluorescence *in situ* hybridization (Oligo-FISH). *Mol Biol Rep* 2021, **48**(1):21-31.
36. Wang H, Chen S, Jiang J, Zhang F, Chen F. Reference gene selection for cross-species and cross-ploidy level comparisons in *Chrysanthemum* spp. *Sci Rep* 2015, **5**:8094.
37. Wang H, Jiang J, Chen S, Fang W, Guan Z, Liao Y, Chen F. Rapid genomic and transcriptomic alterations induced by wide hybridization: *Chrysanthemum nankingense* x *Tanacetum vulgare* and *C. crassum* x *Crossostephium chinense* (Asteraceae). *BMC Genomics* 2013, **14**:902.
38. Wang H, Qi X, Gao R, Wang J, Dong B, Jiang J, Chen S, Guan Z, Fang W, Liao Y *et al*. Microsatellite polymorphism among *Chrysanthemum* sp. polyploids: the influence of whole genome duplication. *Sci Rep* 2014, **4**:6730.
39. Ramsey J, Schemske DW. Pathways, mechanisms, and rates of polyploid formation in flowering Plants. *Annual Review of Ecology and Systematics* 1998, **29**(1):467-501.
40. Liu Y, Tahir Ul Qamar M, Feng JW, Ding Y, Wang S, Wu G, Ke L, Xu Q, Chen LL. Comparative analysis of miniature inverted-repeat transposable elements (MITEs) and long terminal repeat (LTR) retrotransposons in six *Citrus* species. *BMC Plant Biol* 2019, **19**(1):140.
41. Syed NH, Sureshsundar S, Wilkinson MJ, Bhau BS, Cavalcanti JJ, Flavell AJ. *Ty1-copia* retrotransposon-based SSAP marker development in cashew (*Anacardium occidentale* L.). *Theor Appl Genet* 2005, **110**(7):1195-1202.
42. Shaked H, Kashkush K, Ozkan H, Feldman M, Levy AA. Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. *Plant Cell* 2001, **13**(8):1749-1759.

43. Zhai Y, Yu X, Zhou J, Li J, Tian Z, Wang P, Meng Y, Zhao Q, Lou Q, Du S *et al.* Complete chloroplast genome sequencing and comparative analysis reveals changes to the chloroplast genome after allopolyploidization in *Cucumis*. *Genome* 2021.
44. Schiessl SV, Mason AS. Ancient and recent polyploid evolution in *Brassica*. In: *Brassica Improvement*. 2020: 49-66.
45. Tate JA, Ni Z, Scheen AC, Koh J, Gilbert CA, Lefkowitz D, Chen ZJ, Soltis PS, Soltis DE. Evolution and expression of homeologous loci in *Tragopogon miscellus* (Asteraceae), a recent and reciprocally formed allopolyploid. *Genetics* 2006, **173**(3):1599-1611.
46. Clausen J, Hiesey W. The nature of species. (Scientific Books: Experimental studies on the nature of species. II. Plant evolution through amphiploidy and autopoloidy with examples from the Madiinae). *Science* 1945, **102**:674-675.
47. Stebbins G. Chromosomal evolution in higher plants. 1971.
48. Ramsey J. Polyploidy and ecological adaptation in wild yarrow. *Proc Natl Acad Sci U S A* 2011, **108**(17):7096-7101.
49. Hanada K, Vallejo V, Nobuta K, Slotkin RK, Lisch D, Meyers BC, Shiu SH, Jiang N. The functional role of pack-MULEs in rice inferred from purifying selection and expression profile. *Plant Cell* 2009, **21**(1):25-38.
50. Doolittle RF, Feng DF, Johnson MS, McClure MA. Origins and evolutionary relationships of retroviruses. *Q Rev Biol* 1989, **64**(1):1-30.
51. Feschotte C, Jiang N, Wessler SR. Plant transposable elements: where genetics meets genomics. *Nat Rev Genet* 2002, **3**(5):329-341.
52. Lisch D. Epigenetic regulation of transposable elements in plants. *Annu Rev Plant Biol* 2009, **60**:43-66.
53. Lisch D, Bennetzen JL. Transposable element origins of epigenetic gene regulation. *Curr Opin Plant Biol* 2011, **14**(2):156-161.
54. Petit M, Guidat C, Daniel J, Denis E, Montoriol E, Bui QT, Lim KY, Kovarik A, Leitch AR, Grandbastien MA *et al.* Mobilization of retrotransposons in synthetic allotetraploid tobacco. *New Phytol* 2010, **186**(1):135-147.
55. SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL. The paleontology of intergene retrotransposons of maize. *Nat Genet* 1998, **20**(1):43-45.
56. Belyayev, Raskina, Korol A, Nevo AB. Coevolution of A and B genomes in tetraploid *Triticum dicoccoides*. *Genome* 2000.
57. Zhao XP, Si Y, Hanson RE, Crane CF, Price HJ, Stelly DM, Wendel JF, Paterson AH. Dispersed repetitive DNA has spread to new genomes since polyploid formation in cotton. *Genome Res* 1998, **8**(5):479-492.
58. Vinckenbosch N, Dupanloup I, Kaessmann H. Evolutionary fate of retroposed gene copies in the human genome. *Proc Natl Acad Sci U S A* 2006, **103**(9):3220-3225.
59. Feschotte C. Transposable elements and the evolution of regulatory networks. *Nat Rev Genet* 2008, **9**(5):397-405.
60. Oliver KR, McComb JA, Greene WK. Transposable elements: powerful contributors to angiosperm evolution and diversity. *Genome Biol Evol* 2013, **5**(10):1886-1901.
61. Liu S, Chen S, Chen Y, Guan Z, Yin D, Chen F. *In vitro* induced tetraploid of *Dendranthema nankingense* (Nakai) Tzvel. shows an improved level of abiotic stress tolerance. *Scientia Horticulturae* 2011, **127**(3):411-419.
62. Stewart CN, Via L. A rapid CTAB DNA isolation technique useful for RAPD fingerprinting and other PCR amplifications. *BioTechniques* 1993, **14**:748-750.
63. Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M *et al.* AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* 1995, **23**(21):4407-4414.
64. Kumekawa N, Ohtsubo H, Horiuchi T, Ohtsubo E. Identification and characterization of novel retrotransposons of the *gypsy* type in rice. *Mol Gen Genet* 1999, **260**(6):593-602.
65. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 2011, **29**(7):644-

652.

66. Cameron M, Williams HE, Cannane A. Improved gapped alignment in BLAST. *IEEE/ACM Trans Comput Biol Bioinform* 2004, **1**(3):116-129.
67. Liu YG, Mitsukawa N, Oosumi T, Whittier RF. Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J* 1995, **8**(3):457-463.
68. Courtial B, Feuerbach F, Eberhard S, Rohmer L, Chiapello H, Camilleri C, Lucas H. *Tnt1* transposition events are induced by in vitro transformation of *Arabidopsis thaliana*, and transposed copies integrate into genes. *Mol Genet Genomics* 2001, **265**(1):32-42.

Figures

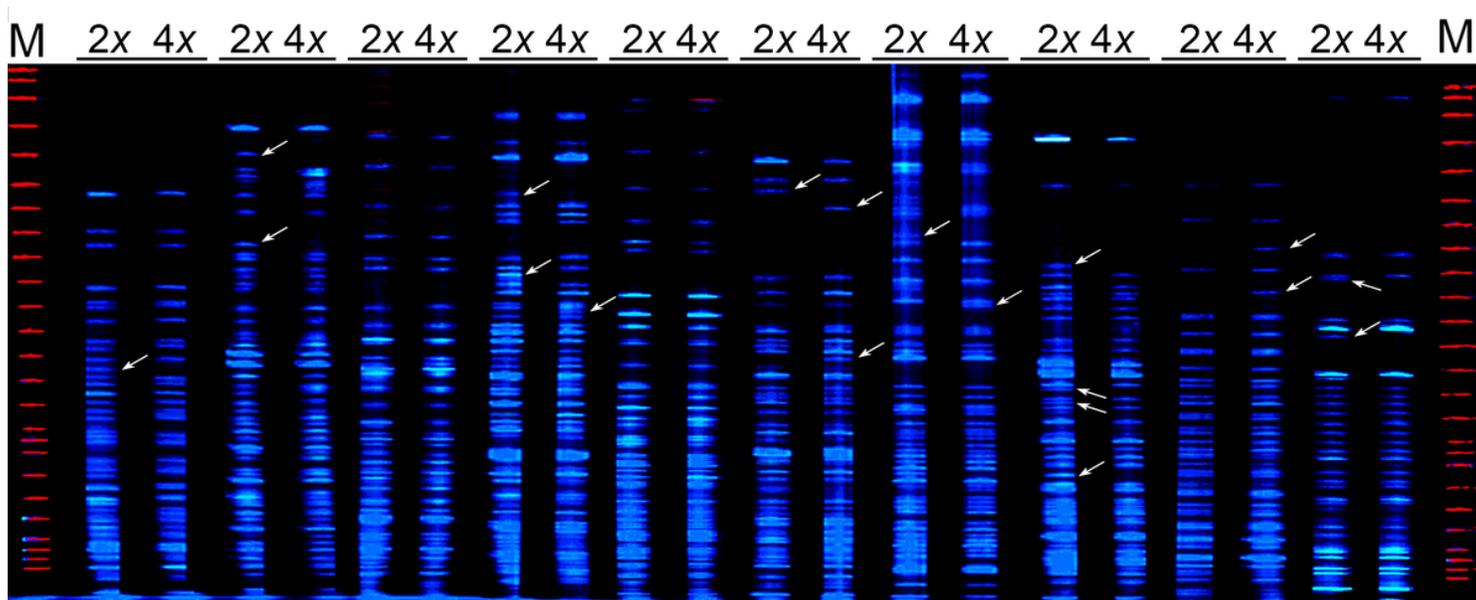


Figure 1

DNA-AFLP profiling of diploid and autotetraploid line 1 (T1) of *C. nankingense*. The arrows indicate genome changes.

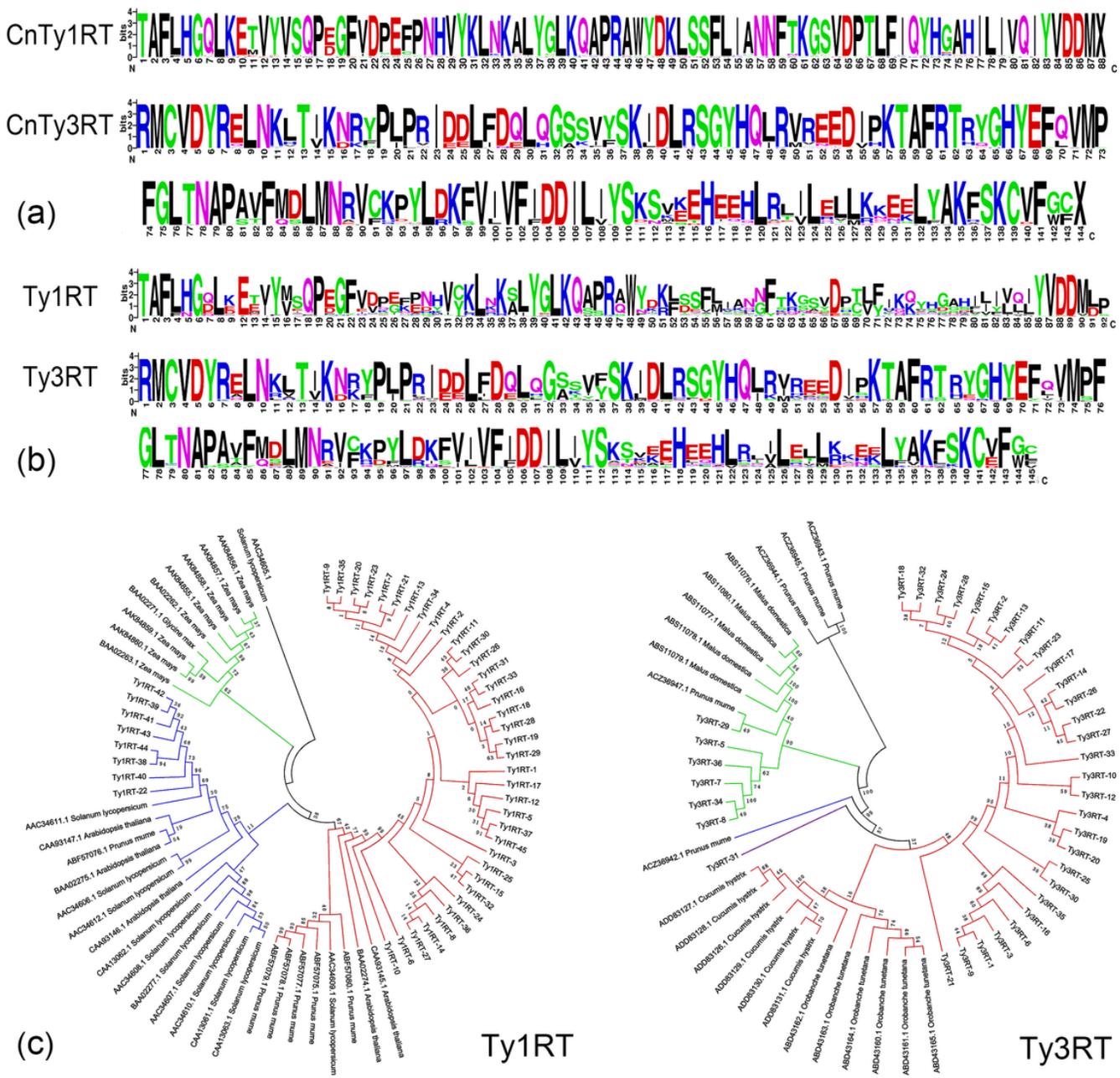


Figure 2

Amino acid analysis of the RT region of the LTR retrotransposon. (a) WebLogo plot of amino acid frequencies at conserved regions of Ty1RT and Ty3RT of *C. nankingense*. Positions are relative to the 5' end of the target site. N, length of the target site. Logos were generated using WebLogo. (b) WebLogo plot of amino acid frequencies at conserved regions of the Ty1RT and Ty3RT sequences of different species. Positions are relative to the 5' end of the target site. N, length of target site. Logos were generated using WebLogo. (c) The phylogenies of Ty1RT and Ty3RT. According to the BlastX results, homologous sequences from other species were the most related sequences in the NCBI database.

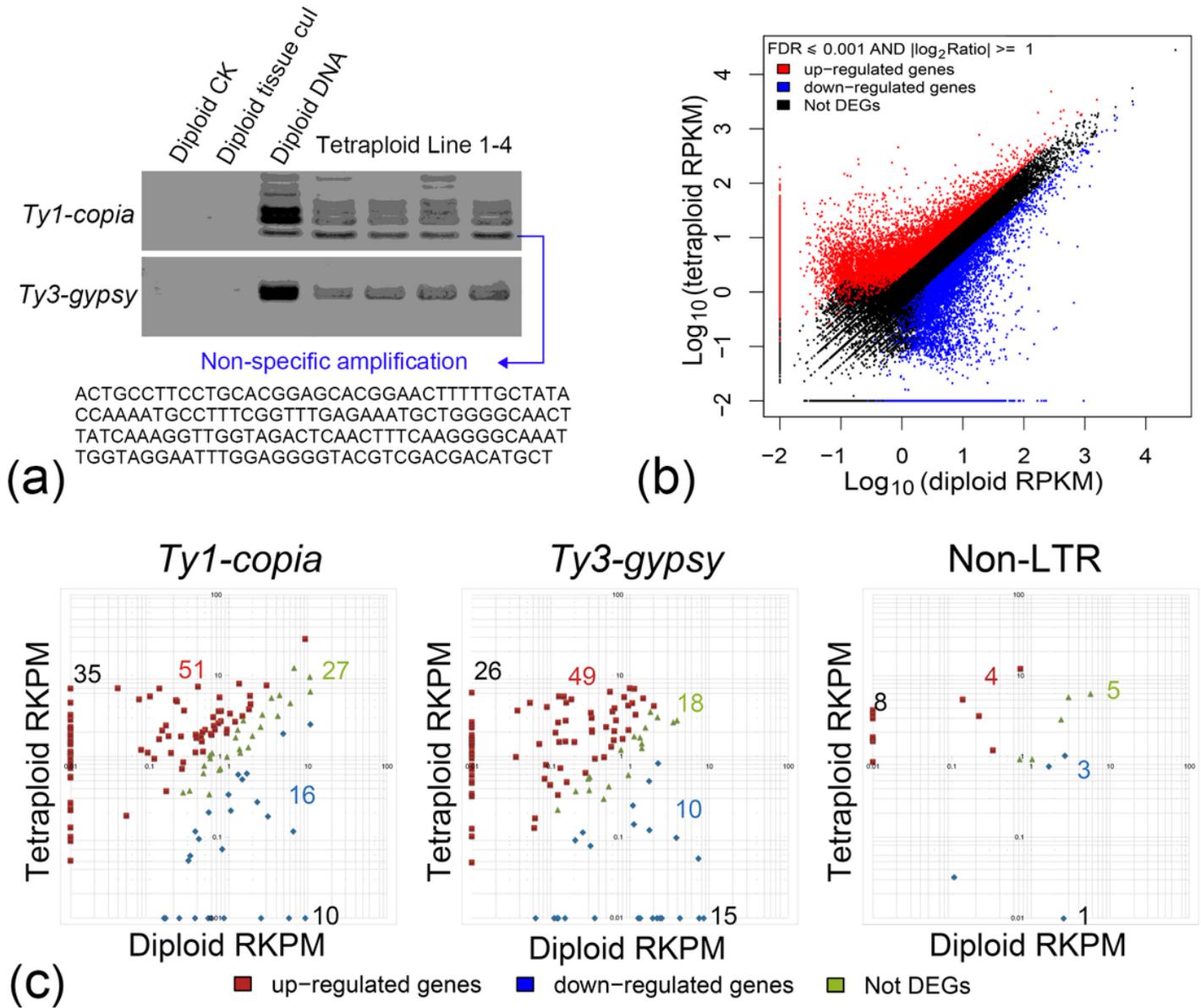


Figure 3

Transcriptional activation of LTR retrotransposons following autopolyploidization. (a) Polyacrylamide gel electrophoresis of the Ty1RT and Ty3RT sequences of diploid (tissue-cultured and non-tissue-cultured) and tetraploid species of *C. nankingense*. The Ty1RT bands at the bottom of the gel show non-specific amplification. (b) Analysis of the differentially expressed genes of diploid and tetraploid species of *C. nankingense*. (c) Analysis of the Ty1-copia, Ty3-gypsy and non-LTR-type retrotransposons of diploid and tetraploid species of *C. nankingense*.

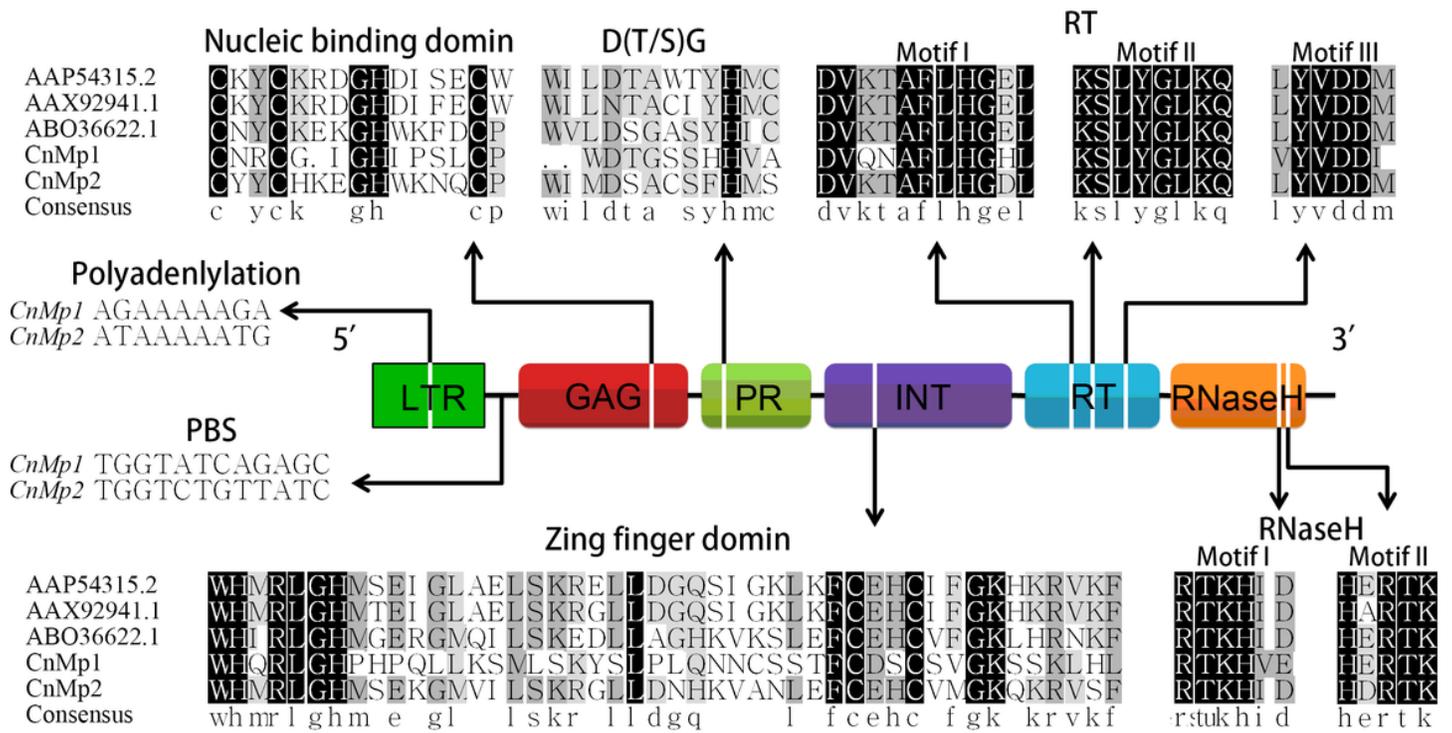


Figure 4

The conservative motifs of CnMp1 and CnMp2. Entire conserved signatures among retroelements are indicated with black shading, and partially conserved signatures are indicated with grey shading.

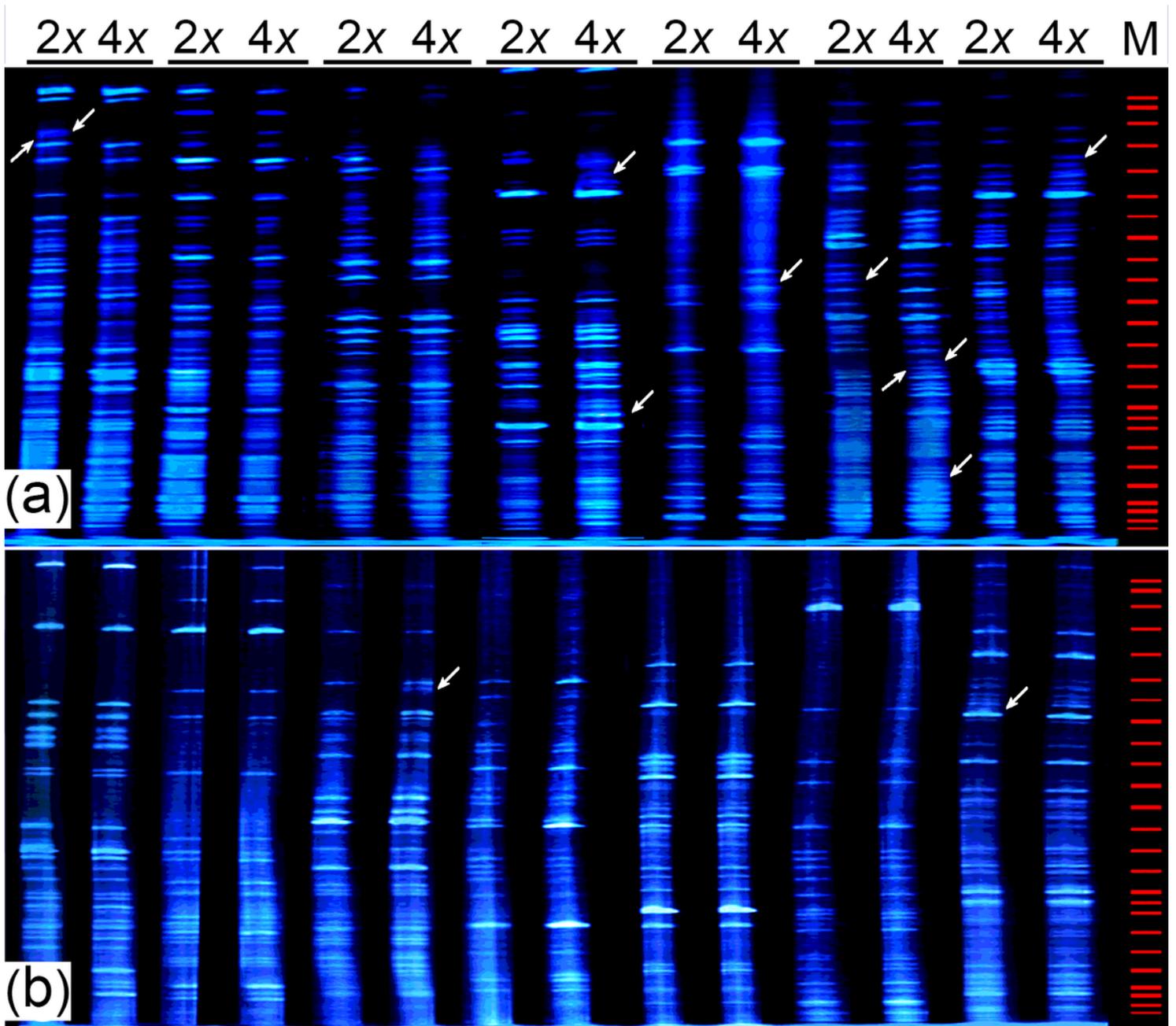


Figure 5

SSAP profiles of diploid and tetraploid (T1) species of *C. nankingense*. The LTR-specific primers were designed from CnMp1(a) and CnMp2 (b).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Fig.S1.tif](#)
- [Fig.S2.tif](#)
- [Fig.S3.tif](#)
- [TableS1.docx](#)

- [TableS2.docx](#)
- [TableS3.docx](#)