Mar Biotechnol
DOI 10.1007/s10126-017-9764-6

ORIGINAL ARTICLE

Chimeras Linked to Tandem Repeats and Transposable Elements in Tetraploid Hybrid Fish

Lihai Ye[1,2] · Ni Jiao[1,2] · Xiaojun Tang[1,2] · Yiyi Chen[1,2] · Xiaolan Ye[1,2] · Li Ren[1,2] · Fangzhou Hu[1,2] · Shi Wang[1,2] · Ming Wen[1,2] · Chun Zhang[1,2] · Min Tao[1,2] · Shaojun Liu[1,2]

Received: 9 February 2017 / Accepted: 9 June 2017

Abstract The formation of the allotetraploid hybrid lineage (4nAT) encompasses both distant hybridization and polyploidization processes. The allotetraploid offspring have two sets of sub-genomes inherited from both parental species, and therefore, it is important to explore its genetic structure. Herein, we construct a bacterial artificial chromosome library of allotetraploids, and then sequence and analyze the full-length sequences of 19 bacterial artificial chromosomes. Sixty-eight DNA chimeras are identified, which are divided into four models according to the distribution of the genomic DNA derived from the parents. Among the 68 genetic chimeras, 44 (64.71%) are linked to tandem repeats (TRs) and 23 (33.82%) are linked to transposable elements (TEs). The chimeras linked to TRs are related to slipped-strand mispairing and double-strand break repair while the chimeras linked to TEs benefit from the intervention of recombinases. In addition, TRs and TEs can also result in insertions/deletions of DNA segments. We conclude that DNA chimeras accompanied by TRs and TEs coordinate a balance between the sub-genomes derived from the parents. It is the first report on the relationship between formation of the DNA chimeras and TRs and TEs in the polyploid animals.

Keywords Chimeras · Tandem repeats · Transposable elements · Tetraploid hybrid fish

Introduction

It is certain that recombination between DNA sequences from various sources with various functions could result in the formation of DNA chimeras. Recombination can provide the raw materials for biological evolution, and it enables the reconstruction and rearrangement of genome to eliminate deleterious mutations. The main driving factor of recombination between any two sequences is homology (Gaeta and Chris 2010). Researches show that recombination can occur among repeats within the same chromosome, on homologous chromosomes (Jelesko et al. 2004), and even among nonhomologous chromosomes that share some degree of homology (Mezard et al. 2007). Recombination initiates at double-stranded DNA breaks and at single-stranded DNA gaps; thus, it is closely linked with DSB (double-strand break) repair, and it is useful for the restoration of broken replication forks (Heyer and Ehmsen 2010). What we concern is how the recombinations occur and what causes the DNA chimeras in polyploids. Recombination in allopolyploids plants may occur ectopically among paralogous (Jelesko et al. 2004) or homoeologous (Qi et al. 2007) sequences because of a lack of diploid pairing fidelity. In newly formed allopolyploids of *Brassica napus*, homoeologous recombinations are deeply entwined with reciprocal exchanges and gene conversions, and are responsible for many of genetic changes (Gaeta and Chris 2010). In addition, nonreciprocal homoeologous exchanges have occurred throughout polyploid divergence and

Lihai Ye, Ni Jiao, Xiaojun Tang, and Yiyi Chen contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s10126-017-9764-6) contains supplementary material, which is available to authorized users.

✉ Shaojun Liu
lsj@hunnu.edu.cn

[1] State Key Laboratory of Developmental Biology of Freshwater Fish, Hunan Normal University, Changsha 410081, Hunan, People's Republic of China

[2] College of Life Sciences, Hunan Normal University, Changsha 410081, People's Republic of China

Published online: 05 July 2017

speciation in allopolyploid cotton (*Gossypium*) (Salmon et al. 2010). Considering that polyploidy increases the number of duplicated sequences resident in the genome, homology-dependent chimeras could also be increased correspondingly. However, there is no doubt that there are some other factors that could cause chimeras in polyploid.

Although our laboratory has conducted research on chimeras in allotetraploid fish (Liu et al. 2016), the repetitive elements contained within those chimeras were not even noticed. There are investigations that implicate that repetitive elements play a role in the generation of chimeras (Jiang et al. 2004; Kapitonov et al. 2006). Repetitive elements have historically been called "junk DNA"; however, they are important forces for recombination and genome evolution (Shapiro and Sternberg 2005). Major repetitive elements include tandem repeats (TRs) and transposable elements (TEs). TRs comprise repeat units that are directly adjacent to each other, and because of their particular structure, the DNA sequences around them are unstable (Bichara et al. 2006). Two mechanisms have been proposed to explain this instability: unequal recombination and DNA polymerase slippage (Debrauwere et al. 1997). There are differences in content of TRs among different species (Toth and Gaspari 2002), among different chromosomes within the same species (Katti et al. 2001), and between coding and noncoding regions (Catasti et al. 1999; Cox and Mirkin 1997). The variable TRs can change the gene structure or even influence gene expression and evolution (Gemayel et al. 2010). The TE is a large category of repetitive DNA, which can transfer to new genome sites and reproduce themselves during the process. According to the mechanisms of transposition (Finnegan 1992) and structural features in the genome, TEs are divided into two types: transposons and retrotransposons. Transposons consist of one target site repeat (TSR) on each side of the transposase coding an enzyme catalysis transposition. Retrotransposons can also be classified into two groups: one group contains protease (PR), integrase (INT), and reverse transcriptase (RT) and is flanked by long terminal repeats (LTRs), however, there are no LTR and PR found in the other groups of retrotransposons, except for (A)n, GAG, INT, and RT. Different types of TEs have different roles in genome evolution. They intersperse in the genomes of plants and animals, occupying a large proportion; thus, they may alter not only the individual gene structure, but also the genome structure and function (Bennetzen 2000). In addition, they have vital influences on the structural diploidization of genomic DNA (Lim et al. 2007; Bruggmann et al. 2006) and the early stage of the evolution of allopolyploids (Oliver and Greene 2009). Undoubtedly, they are an important component of an organism's genome and a key player of biological evolution, genetic heredity, and variation. Thus, repetitive elements provide a great platform for the study of genetic variation in polyploids.

Polyploids are organisms with three or more complete sets of chromosomes, and many plants and animals have experienced polyploidization one or more times during their evolution (Song et al. 2012; Masterson 1994). The essence of allopolyploid formation is the fusion of genomes of two species, which develops into a hybrid whose phenotype and genotype are different from the parental species. However, the key mechanism of their genome reconstruction is homoeologous pairing and recombination (Gaeta and Chris 2010). Homoeologous recombination is observed in allopolyploids of *Lolium multiflorum* × *Festuca* by genomic in situ hybridization (Zwierzykowski et al. 1998). In the genome of resynthesized *B. napus* allopolyploids, many genomic changes are detected, such as recombinations, deletions, replication, and translocations (Gaeta and Chris 2010). But the allopolyploid formation is very rare in animals, and even less reports have focused on the recombination in animal polyploids. The allotetraploid fish hybrid (4nAT) is created by our laboratory via an artificial intergeneric cross between red crucian carp (RCC) (*Carassius auratus* red var., ♀, 2n = 100) and common carp (CC) (*Cyprinus carpio* L., ♂, 2n = 100) (Liu et al. 2001), and to date, the 25th generation of 4nAT has been formed through successive self-breeding. The 4nAT represents the foundation for the speciation of a new tetraploid fish and provides perfect materials for studying the effects of distant hybridization and polyploidization on genome evolution. However, the allotetraploids has two sets of genomes inherited from diploid progenitors; what genetic changes have happened to them and what elements are supposed to be involved in the process? These are the questions that we seek to answer in the present study. Herein, we construct a bacterial artificial chromosome (BAC) library of allotetraploid offspring and compared the BAC sequences with their parental genomes to analyze the relationship between the mechanisms of genetic changes, such as gene recombination, insertions/deletions (indels), and repetitive elements, which will provide new impetus for research on the genome evolution of polyploids.

Materials and Methods

The Constructing of 4nAT BAC Library

Blood samples (10 ml) from five allotetraploid individuals from the 20th generation were used to construct a BAC library according to a previously described method (Wang et al. 2015). The allotetraploid fish were obtained from the Engineering Research Center of Polyploid Fish Breeding at Hunan Normal University (Changsha, China), and the ploidy of five samples was confirmed by metaphase chromosome assay of kidney cells. The genomic DNA of allotetraploids was digested with the restriction enzyme *Hind* III, ligated into the CopyControl pCC1BAC vector, and then transformed into

Escherichia coli DH10B. The recombinant bacterial clones were screened through blue-white screening. Nineteen BAC clones were picked randomly and sequenced by Shanghai Majorbio Bio-Pharm Technology Co. Ltd. The sequencing of the 19 BAC clones was performed by Illumina next-generation sequencing technology and PacBio RS platform; the assembly work of the BACs was done by professional software SOAPdenovo/Celera, and ABI 3730XL was used for closing gaps. The coverage of the overall sequencing data is more than 100×, and the average error rate of single base is lower than 1/100,000.

Data Processing

NCBI-VecScreen was used to remove the vector sequences from the obtained BAC sequences (http://www.ncbi.nlm.nih.gov/tools/vecscreen/), and RepeatMasker (http://www.repeatmasker.org/) was used to calculate the types and contents of repetitive elements in the allotetraploid BAC sequences, using *Danio rerio* as the reference organism. We compared the BAC sequences that include repetitive elements with the genomes of both progenitors (using NCBI-BLAST-2.3.0+), and homologous sequences were identified with an *E* value <1e–5 and similarity >95%. The maternal progenitor genome sequencing has been completed by our laboratory together with Yunnan University, and the database has yet to be published (http://rd.biocloud.org.cn/). The genome of paternal progenitor can be searched from the NCBI database (http://www.ncbi.nlm.nih.gov/) and the Common Carp Genome Database (http://www.carpbase.org/). It was confirmed by PCR and sequencing that the DNA chimeras really exist in the hybrid fish genome.

Models of Chimeras Based on BAC Analysis

We calculated the number of chimera DNA linked to repetitive elements in the allotetraploid BAC sequences and then surveyed the types and presence of the repetitive elements in the parental homologous sequences. Almost all the recombinations (except for the *fibronectin 1* (*FN1*) gene loci) detected in this study are related with repetitive elements, and they can be divided into four composite patterns according to the presence of repetitive elements in the parental homologous sequences. The four composite patterns are shown in Fig. 1: repetitive elements existing only in homologous sequences of maternal progenitor (Fig. 1a), repetitive elements existing only in homologous sequences of paternal progenitor (Fig. 1b), repetitive elements existing in both parental homologous sequences (Fig. 1c), and repetitive elements that do not exist in either parental homologous sequence (Fig. 1d).

Significance Difference on Statistical Analysis

To better explore the mechanism of chimeras that were resulted from repetitive elements, significance tests of the incidence of chimeras under situations a, b, c, and d were conducted. In addition, we analyzed the differences among the incidence of chimeras linked to mono-, di-, tri-, and tetra-nucleotide repeats. The formula involved in the different significance tests in this article is listed in the following (Du 2003):

$$u = \frac{\frac{x_1}{n_1} - \frac{x_2}{n_2}}{\sqrt{\overline{p}\left(1-\overline{p}\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

$$\overline{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

In the formulas, n_1 and n_2 represent total number of samples and x_1 and x_2 represent DNA chimera numbers in total samples.

Results

Statistics of Repetitive Elements in 4nAT BAC

The BAC library of allotetraploid hybrids has been constructed, and 19 full-length BAC clones (NCBI accession numbers: KF758440–KF758444, KJ424354–KJ424362, and KT726912–KT726916) are obtained with a total length of 752,904 bp. The GC content of the 19 BAC clone sequences ranged from 34.17% to 41.40%, with an average of 37.31%, and the content of repetitive elements ranged from 5.12% to 54.96%, with an average of 16.88%, according to RepeatMasker datasets. The repetitive DNA content of homologous sequences of the maternal RCC and paternal CC genome is 12.32% and 15.50%, respectively.

We identify 480 TRs, of which 44 (9.17%) are related to recombination (Table 1). The results show that the most abundant type of TRs is A/T, whose total number is much larger than that of G/C. The total content of the TRs in 4nAT is negatively correlated with the length of the repeat unit. The statistics confirm that the numbers of chimeras in 4nAT linked to mono-, di-, tri-, and tetra-nucleotide repeats are 11, 24, 2, and 2, respectively. Among the chimeras linked to TRs, the numbers of chimeras under composite patterns a, b, c, and d are 8, 12, 14, and 10, respectively. The total number of TRs under situation c has the largest number, and the corresponding number of DNA chimeras is the least (Table 1). The different significance tests of rates of chimeras under the above four situations showed that the incidence of chimeras under situation c (in which both the homologous sequences of the parental progenitors shared the TR) is much lower than the

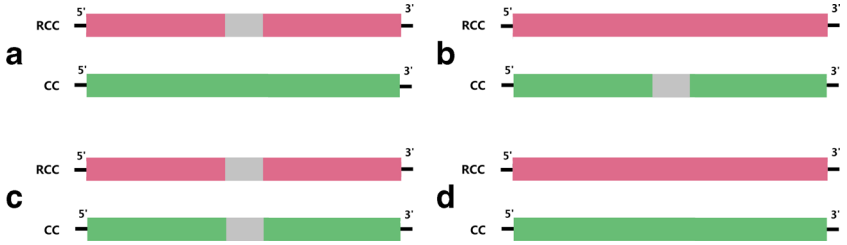

Fig. 1 Four composite patterns of parental homologous sequences. **a** Repetitive elements existing only in homologous sequences of maternal progenitor, **b** repetitive elements existing only in homologous sequences of paternal progenitor, **c** repetitive elements existing in both parental homologous sequences, and **d** repetitive elements that do not exist in either parental homologous sequence. The above four composite patterns are divided according to the existent of repetitive element in homologous sequences of parental progenitors. The *red color* represents genomic DNA inherited from red crucian carp (RCC), and the *green color* represents genomic DNA inherited from common carp (CC). The repetitive elements are represented in *gray*

rates of the other three situations, while the incidences of the three other situations showed no remarkable differences (Fig. 2a). The different significance tests of the incidences of chimeric genes linked to TRs of different repeat unit length show that the incidence of chimeric genes related to di-nucleotide repeats is significantly higher than the incidences related to the other three types of TRs and there are no remarkable differences among the three other incidences (Fig. 2b).

We obtain 217 TEs in the 19 BAC sequences, including 139 (64.06%) DNA transposons and 78 (35.94%) retrotransposons. The analysis shows that the distribution densities of TEs in different BAC sequences are different and the maximum density (4nAT-150D6) is 6.9 times greater than the minimum (4nAT-150B4) (Fig. 3). The numbers of chimeras related to TEs under situations a, b, c, and d are 2, 2, 18, and 1, respectively, and the incidence of DNA chimeras in the above four situations shows no significant differences ($u < u_{0.05} = 1.645$, Fig. 3).

Models of Chimeras

In this study, we confirm that repetitive elements are closely linked to the incidence of chimeras. In each of the composite patterns a, b, c, and d (Fig. 1), there are four chimera models (Fig. 4) that can be identified in allotetraploid offspring according to their genetic source of the two ends: model 1: the 5′ end is inherited from RCC while the 3′ end is from CC, model 2: the 5′ end is inherited from CC while the 3′ end is from RCC, model 3: both the 5′ end and the 3′ end are inherited from different scaffolds of RCC, and model 4: both the 5′ end and the 3′ end are inherited from different scaffolds of CC.

Chimeras Linked to TRs and TEs

Forty-four (9.17%) out of 480 TR loci are linked to chimeras, and among them, the largest number of chimeras is linked to TA repeats, followed by mono-nucleotide A/T repeats. Notably, no DNA chimeras occur at mono-nucleotide G/C repeat loci. Chimeras linked to TRs have four models shown in Fig. 4, and their specific structures are shown in Figs. S1, S2, S3, and S4. In addition, there are small pieces of DNA inserted into the recombination locus (see Figs. S1, S2, and S4), and according to our statistics, 29.55% of the chimeras have these “little tails” of various lengths.

Table 1 The statistical parameters of different tandem repeats

Repeat types	*x*/*n* (a)	*x*/*n* (b)	*x*/*n* (c)	*x*/*n* (d)	Total number
A/T	6/46	2/38	2/76	1/8	11/173
G/C	0/2	0/0	0/3	0/0	
TA	5/20	2/8	2/19	3/14	24/162
TG	1/10	4/12	3/46	2/7	
TC	0/4	1/5	0/14	1/3	
TAT	0/1	0/2	0/4	0/0	2/49
GTT	1/1	0/1	0/2	0/1	
AAT	0/6	0/4	1/18	0/2	
Other tri-nucleotide repeats	0/1	0/2	0/4	0/0	
ATTT	0/3	1/5	0/7	0/0	2/46
ATAA	1/1	0/2	0/7	0/2	
Other tetra-nucleotide repeats	0/2	0/4	0/11	0/2	
The remaining repeats	0/6	2/7	0/29	3/8	5/50
	14/103	12/90	8/240	10/47	44/480

The *x* and *n* represent the same statistical parameters with the formulas above

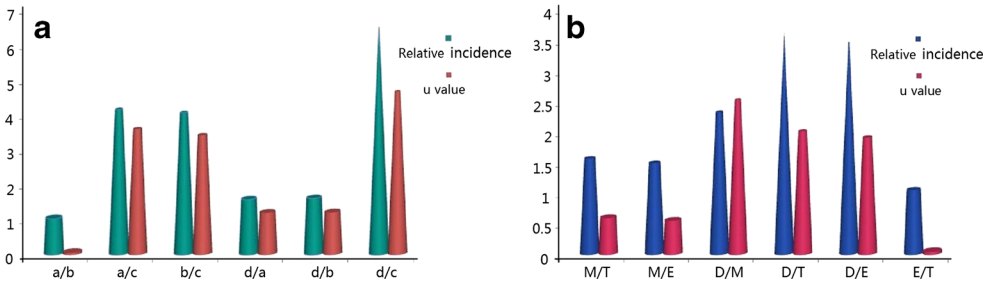

Fig. 2 The statistics of tandem repeats. **a** Relative incidence of chimeras and *u* value of significant test among situations a, b, c, and d, where the *u* values of a/c, b/c, and d/c are greater than $u_{0.01}$ (2.337). **b** Relative incidence of chimeras and *u* value of significant test among mono-nucleotide (*M*), di-nucleotide (*D*), tri-nucleotide (*T*), and tetra-nucleotide (*E*) repeat loci, where the *u* values of M/D (mono-nucleotide repeat versus di-nucleotide repeat), D/T (di-nucleotide repeat versus tri-nucleotide repeat), and D/E (di-nucleotide repeat versus tetra-nucleotide repeat) are greater than $u_{0.05}$ (1.645)

Twenty-three (10.60%) of 217 TE loci are detected linked with chimeras. Among them, 12 are DNA transposons and 11 are retrotransposons. Chimeras related to TEs also have four models shown in Fig. 4, and their specific structures are shown in Figs. S5, S6, S7, and S8. Similarly, there are small DNA segments inserted into the recombination locus too (see Figs. S5 and S6), and according to our statistics, 24.67% of the chimeras have been inserted with DNA pieces of various lengths.

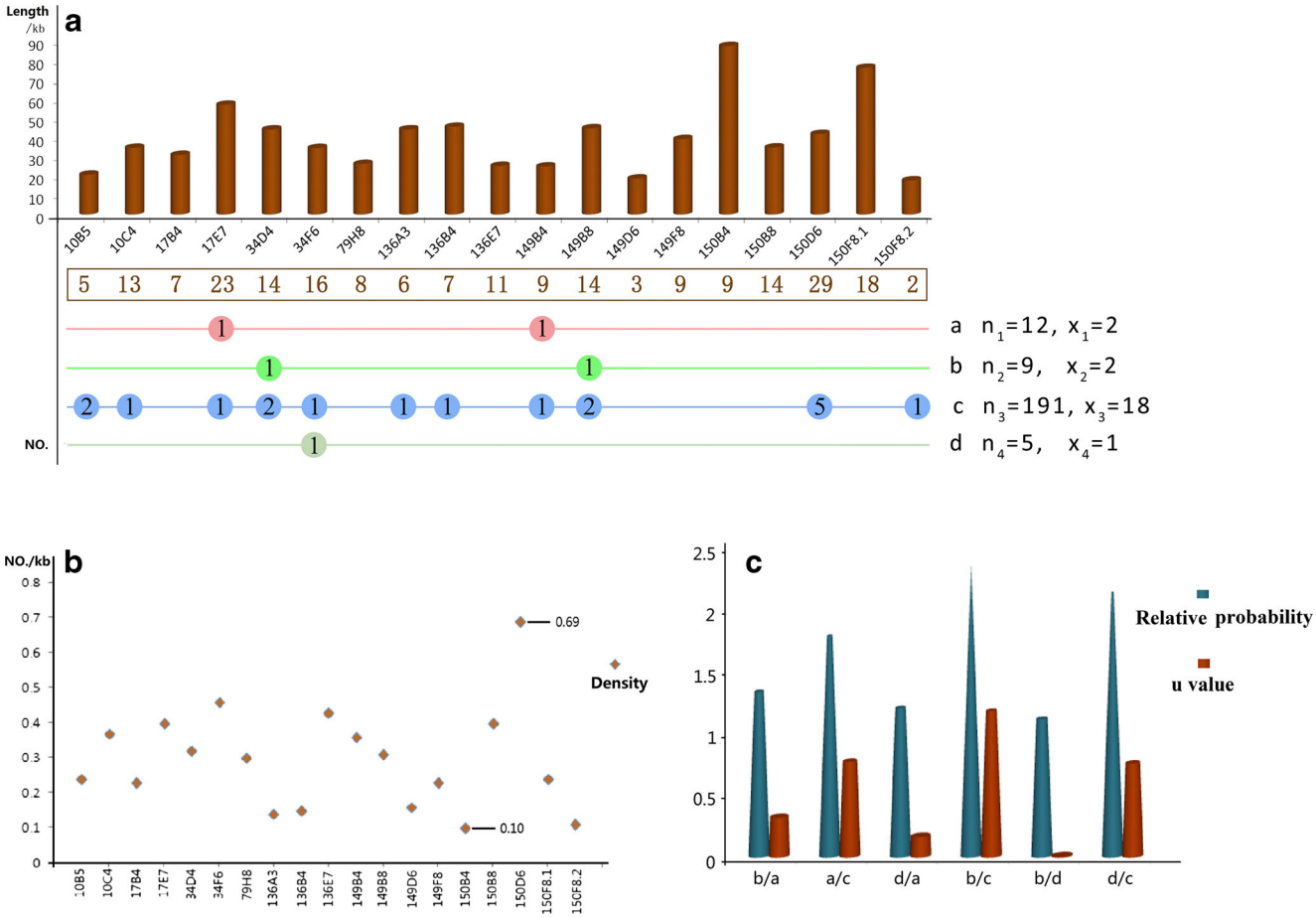

Fig. 3 The statistics of transposable elements. **a** Total length of every BAC sequence (*above* the *horizontal axis*) and the corresponding chimeric DNA numbers linked to TEs in situations a, b, c ,and d (*below* the *horizontal axis*). **b** Density of TEs in the 19 BAC sequences and the highest and lowest TE densities occurred in BACs 4nAT-150D6 (0.69/kb) and 4nAT-150B4 (0.10/kb) respectively. **c** Relative incidence of chimeras and *u* value of significant test among situations a, b, c, and d related with TE loci; all of the *u* values are smaller than $u_{0.05}$ (1.645), which means that the incidence of chimeric genes had no significant differences among the four situations

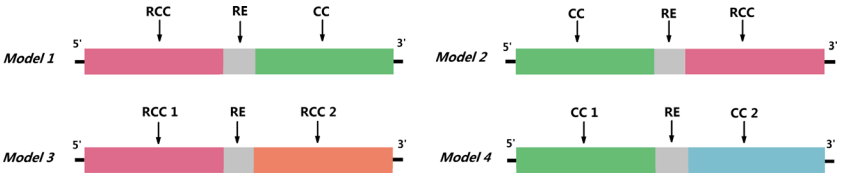

Fig. 4 Recombination models linked to repetitive elements. The above four recombination models are divided according to the distribution of the genomic DNA derived from the parents. The *red* and *orange colors* represent different genomic DNAs derived from red crucian carp (RCC) and the *green* and *blue colors* represent different genomic DNAs derived from common carp (CC). The repetitive elements are represented in *gray*

Insertion and Deletion of TEs

Comparing the BAC sequences with the homologous sequences of both progenitors, insertions/deletions (indels) of TEs are identified in allotetraploid BACs (see Fig. 5). By comparing the 4nAT-150D6 BAC sequence to scaffold CC-1513 of paternal genome, we found that a retrotransposon Gypsy35-I is inserted into 4nAT-150D6 and a transposon, Mariner-1, is deleted in 4nAT-150D6. Meanwhile, by comparing the BAC sequence to scaffold RCC-200893 of maternal genome, a residue of LINE1 is detected to be inserted into 4nAT-150D6 and a LINE1-2 sequence is deleted in 4nAT-150D6.

Chimeras in *FN1* Gene

There is only one chimera detected at nonrepetitive element loci among the 19 BAC clone sequences of tetraploid hybrids, which is the *FN1* gene in BAC 4nAT-10C4. The *FN* gene family encodes fibronectin and is highly conserved in evolution. The *FN1* genes of maternal RCC and paternal CC have high sequence similarity, and some of their coding regions are almost identical; therefore, homoeologous recombination could occur at the *FN1* gene of 4nAT driven by homology. *Fibronectin* is composed of two strands, and each strand consists of a chain of repeat unit. So, although *FN1* gene is not repetitive element, it has the similar structure. Figure 6 shows the structure of *FN1* gene in BAC 4nAT-10C4.

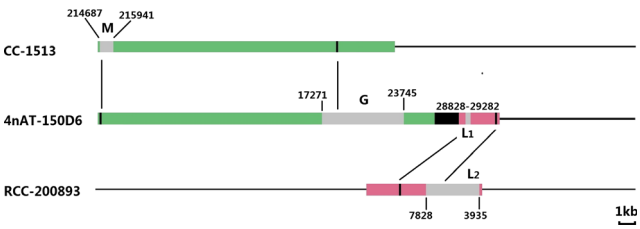

Fig. 5 The alignment among BAC 4nAT-150D6 sequence and the parental genomes. CC-1513 and RCC-200893 are the homologous sequences of paternal species and maternal species, respectively. The *bars* in *gray color* are transposable elements and *M*, *G*, *L1*, and *L2* represent transposon Mariner-1, retrotransposons Gypsy35-I, LINE1, and LINE1-2 respectively. The region in *black color* in 4nAT-150D6 is shared by both of the parental homologous sequences

Discussion

Repetitive Elements in 4nAT Genome

Repetitive elements exist extensively in the genome and have important effect on genomic instability. Therefore, studying repetitive elements can help to understand the evolution of polyploid genomes. The content of repetitive elements is always a subject of major concern for many researchers. In the processes of cloning and sequencing cDNAs of the common bean, no GC TRs are found (Blair et al. 2009). Similarly, G/C and GC TRs are much less than A/T and AT TRs in the allotetraploid BACs. A study shows that there is a positive correlation between the length of TRs and the frequency of variation (Schug et al. 1998). The full BAC sequences also exhibit a plenty of large TRs. Based on these results, it is apparent that lots of the TR loci in tetraploid hybrids are extremely unstable. There are earlier researches that produced similar findings that tandemly repeated DNA sequences can give rise to frequent recombination events within the gene or between the gene and a pseudogene, causing expansion and contraction in the gene size. According to some researchers, the gene size variation can led to phenotypic alterations or even various diseases (Verstrepen et al. 2005). TEs have complicated effects on the recombination of host genomes. Most of the time, they are accompanied with indels and rearrangements of large DNA (Kazazian 2004), which can lead to genomic structural changes (Christiansen et al. 2008). Our analysis of 4nAT BAC sequences produces similar results to these previous studies. There are widespread gene replications in polyploids, and some of the TEs might be activated, resulting in genetic variations such as recombination becoming more common. However, the density of TEs in the genome of *Drosophila* is higher in the lower recombination rate regions (Cridland et al. 2013). We also investigated the density of TEs in the BAC sequences in this study but found no connection between TE

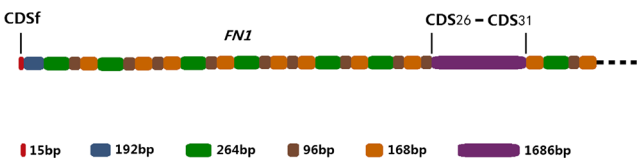

Fig. 6 The structure of *FN1* gene in BAC clone 4nAT-10C4. *FN1* gene contains 35 exons, and is not completely showed here

distribution and incidence of chimeras. This may be due to that the TE distribution also depends on the specific characteristics of chromosomes, the TEs themselves, and the organisms (Rizzon et al. 2002; Cridland et al. 2013). As regulatory elements, retrotransposons have a high tendency of amplification, which can have a large influence on the genome size of plants. For instance, they occupy at least half of the genomes of wheat (Echenique et al. 2002) and maize (Sanmiguel and Bennetzen 1998). By contrast, transposons have relatively less influence on the genome size of plants (Kunze et al. 1997). Intergenomic displacements are common in allopolyploids, which indicate that recombination occurs commonly between homoeologous chromosomes, making the different subgenomes in allopolyploids interdependent in the evolution process (Wendel 2000). Research has shown that some TRs in eukaryotes are derivatives of TEs (Sharma et al. 2013), and furthermore, the evolutions of different genes in polyploid plants are linked to each other (Sang et al. 1995; Wendel et al. 1995). Genomic variation is thus complicated by the interplay among different gene families and repetitive elements.

Mechanism of Recombination

Homologous and homeologous recombinations of allotetraploid fish are observed in this study. Homoeologous recombination that happened in DNA segments of different progenitors (model 1 and model 2 in Fig. 4) that share some degree of homology and homologous recombination occurs in different sequences within the same progenitor genome (model 3 and model 4 in Fig. 4). Research (San et al. 2008) demonstrated that homologous recombinations are the exchange of genetic information between alleles, mediated by the conserved recombinases (such as *Rad51* and *Dmc1*), which are an integral part of mitosis and meiosis and ensure the stability of karyotypes. On the other hand, there was a hypothesis that homeologous recombination can be markedly decreased due to sequence divergence (Li et al. 2006). Both of the homologous and homeologous recombinations can lead to novel gene combinations and generate new phenotypes. At the same time, they can destabilize karyotypes and may even result in aberrant meiotic behavior (Gaeta and Chris 2010). In some polyploids, however, recombination of homoeologous loci is required for stability (Udall et al. 2005). Triggers of recombination include single-strand nicks and DSBs (Szostak et al. 1983); however, given the current data, we do not know whether the recombination hotspots could be changed by different triggers. During the meiotic divisions of plant cells, recombination mainly occurs between allelic sequences of homologous chromosomes (Naranjo and Corredor 2008), but not randomly along the chromosomes. In fact, the recombination hotspots are variable in different species (Anderson et al. 2001). In humans, the sub-telomeres of chromosomes are hotspots of interchromosomal recombination (Linardopoulou et al. 2005); however, in human cancer cells, the ribosomal RNA gene clusters are recombination hotspots (Stults et al. 2009). It has been reported that satellite sequences and TEs could initiate recombination and generate new genes (Yang et al. 2008). Herein, we found similar results except that the satellite sequences initiate recombination. Given that large amounts of repetitive elements have been detected in many species of eukaryote, recombination between ectopic loci seems to be inevitable. The instability of repetitive elements and the intervention of recombinase both can benefit the recombination and produce the chimeras in the hybridization of allotetraploid fish.

TRs and Chimeras

There is no question about the origination and evolution of TRs that involve several complex mechanisms and factors; for example, slipped-strand mispairing (SSM) was proposed as the main mechanism of TR evolution (Levinson and Gutman 1987). Under this mechanism, shorter repeats become longer and the incidence of noncontiguous SSM increases as the repetitive regions become longer and most of the TRs in 4nAT in this study are in line with this trend. According to the studies on yeast, *Drosophila*, and humans, the incidence of SSM of di-nucleotide TRs is the highest and the rate of chimera related to di-nucleotide TRs in the present research is exactly the highest and this suggests that chimeras related to TRs are linked to SSM evolutionary mechanism. There is no objection that the TRs are unstable in the genomes. During the process of DNA replication, DSB (double-strand break) is more likely to occur at TR locus, and then, the chimera is formed through recombination between different sequences in the process of DSB repair. When both of the parental homologous sequences contain the same TR, they are conserved between each other and it is hard to detect the chimera, so the incidence of recombination in the corresponding homologous sequence of 4nAT is the lowest.

TEs and Chimeras

TEs have a sustained impact on genome evolution; for example, they are important ectopic recombination sites (Yang et al. 2008). In some allopolyploid plants, recombinations induced by TEs play an important role in the reconstruction of ribosomal DNA (Pontes et al. 2004). Some studies have claimed that the activities of TEs have an irreplaceable role in the generation of new genes and genome rearrangements of flowering plants (Bennetzen 2005) and are the main drivers of diversity in vertebrate genome (Pontes et al. 2004). Retrotransposons or retrotransposon-like sequences are located in the flanking regions of a series of genes (White and Wessler 1994), and they are demonstrated to mediate

recombinations to form chimeras (Yang et al. 2008). Research shows that even dissimilar sequences could be linked together through TEs (Shapiro 2005), which is analogous to the recombination in the allotetraploid genome. The genome of allotetraploid hybrid fish contains the genomes of both progenitors, which result in a high rate of heterozygosity of the chromosomes. Furthermore, recombination can occur between different LTRs and may lead to the loss of some internal segments (Liu and Wendel 2002) and some of the new LTRs in the allotetraploid hybrid BAC sequences may be generated by these recombinations. Additionally, a group of tyrosine recombinase was reported to have the ability to encode DNA transposons from pathogenic fungi (Goodwin et al. 2003) and it is acceptable to infer that the TE sites benefiting from the recombinases have a role to play. And with the help of recombinases, the TEs can promote genomic rearrangements through ectopic homologous recombination (Belancio et al. 2010), which may be one explanation for why the four models of chimeras at TE loci in 4nAT show no significant difference among composite patterns a, b, c, and d.

Conclusion

The presence of the DNA chimeras found in the polyploid derived from the hybridization is a novel trait, however, little is known about the connection between chimeras and the embedded repetitive elements. Herein, we first report that different models of DNA chimeras in allotetraploid hybrid fish are linked to tandem repeats (TRs) and transposable elements (TEs). In addition, TRs and TEs can also result in indels of DNA segments too. These results show that the repetitive elements can coordinate a balance between the sub-genomes derived from the parental progenitors.

Acknowledgments This work was supported by the National Natural Science Foundation of China (grants 30930071, 91331105, 31360514, 31430088, and 31210103918), the Cooperative Innovation Center of Engineering and New Products for Developmental Biology of Hunan Province (20134486), the Construction Project of Key Discipline of Hunan Province and China, and the National High Technology Research and Development Program of China (Grant No. 2011AA100403).

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

Anderson LK, Hooker KD, Stack SM (2001) The distribution of early recombination nodules on zygotene bivalents from plants. Genetics 159:1259–1269

Belancio VP, Roy-Engel AM, Deininger PL (2010) All y'all need to know 'bout retroelements in cancer. Semin Cancer Biol 20:200–210

Bennetzen JL (2000) Transposable element contributions to plant gene and genome evolution. Plant Mol Biol 42:251–269

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

Bichara M, Wagner J, Lambert IB (2006) Mechanisms of tandem repeat instability in bacteria. Mutat Res Fundam Mol Mech Mutagen 598: 144–163

Blair MW, Torres MM, Giraldo MC, Pedraza F (2009) Development and diversity of Andean-derived, gene-based microsatellites for common bean (*Phaseolus vulgaris L.*) BMC Plant Biol 9:100

Bruggmann R, Bharti AK, Gundlach H, Lai J, Young S et al (2006) Uneven chromosome contraction and expansion in the maize genome. Genome Res 16:1241–1251

Catasti P, Chen X, Mariappan SVS, Bradbury EM, Gupta G (1999) DNA repeats in the human genome. Genetica 106:15–36

Christiansen G, Molitor C, Philmus B, Kurmayer R (2008) Nontoxic strains of *cyanobacteria* are the result of major gene deletion events induced by a transposable element. Mol Biol Evol 25:1695–1704

Cox R, Mirkin SM (1997) Characteristic enrichment of DNA repeats in different genomes. Proc Natl Acad Sci U S A 94:5237–5242

Cridland JM, Macdonald SJ, Long AD, Thornton KR (2013) Abundance and distribution of transposable elements in two *Drosophila* QTL mapping resources. Mol Biol Evol 30:2311–2327

Debrauwere H, Gendrel CG, Lechat S, Dutreix M (1997) Differences and similarities between various tandem repeat sequences: minisatellites and microsatellites. Biochimie 79:577–586

Du RQ (2003) Biostatistics (in Chinese), 2nd edn. Higher Education Press, Beijing, pp 80–81

Echenique VC, Stamova B, Wolters P, Lazo G, Carollo V et al (2002) Frequencies of Ty1-*copia* and Ty3-*gypsy* retroelements within the *Triticeae* EST databases. Theor Appl Genet 104:840–844

Finnegan DJ (1992) Transposable elements. Curr Opin Genet Dev 2(6): 153–184

Gaeta RT, Chris PJ (2010) Homoeologous recombination in allopolyploids: the polyploid ratchet. New Phytol 186:18–28

Gemayel R, Vinces MD, Legendre M, Verstrepen KJ (2010) Variable tandem repeats accelerate evolution of coding and regulatory sequences. Annu Rev Genet 44:445–477

Goodwin TJ, Butler MI, Poulter RT (2003) Cryptons: a group of tyrosine-recombinase-encoding DNA transposons from pathogenic fungi. Microbiology 149:3099–3109

Heyer WD, Ehmsen KT, Jie L (2010) Regulation of homologous recombination in eukaryotes. Annu Rev Genet 44:113–139

Jelesko JG, Carter K, Thompson W, Kinoshita Y, Gruissem W (2004) Meiotic recombination between paralogous *RBCSB* genes on sister chromatids of *Arabidopsis thaliana*. Genetics 166:947–957

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

Kapitonov VV, Jurka J (2006) Self-synthesizing DNA transposons in eukaryotes. Proc Natl Acad Sci U S A 103:4540–4545

Katti MV, Ranjekar PK, Gupta VS (2001) Differential distribution of simple sequence repeats in eukaryotic genome sequences. Mol Biol Evol 18:1161–1167

Kazazian HH (2004) Mobile elements: drivers of genome evolution. Science 303:1626–1632

Kunze R, Saedler H, Lonnig WE (1997) Plant transposable elements. Adv Bot Res 27:331–470

Levinson G, Gutman GA (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol 4:203–221

Li L, Jean M, Belzile F (2006) The impact of sequence divergence and DNA mismatch repair on homeologous recombination in Arabidopsis. Plant J 45:908–916

Lim KY, Kovarik A, Matyasek R, Mark W, Chase MW, James JC et al (2007) Sequence of events leading to near-complete genome turnover in allopolyploid *Nicotiana* within five million years. New Phytol 175:756–763

Linardopoulou EV, Williams EM, Fan Y, Friedman C, Young JM et al (2005) Human subtelomeres are hot spots of interchromosomal recombination and segmental duplication. Nature 437:94–100

Liu B, Wendel JF (2002) Non-Mendelian phenomena in allopolyploid genome evolution. Curr Genomics 3:489–505

Liu S, Liu Y, Zhou G, Zhang X, Luo C et al (2001) The formation of tetraploid stocks of red crucian carp × common carp hybrids as an effect of interspecific hybridization. Aquaculture 192:171–186

Liu S, Luo J, Chai J, Ren L, Zhou Y et al (2016) Genomic incompatibilities in the diploid and tetraploid offspring of the goldfish × common carp cross. Proc Natl Acad Sci U S A 113:1327–1332

Masterson J (1994) Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. Science 264:421–424

Mezard C, Vignard J, Drouaud J, Mercier R (2007) The road to crossovers: plants have their say. Trends Genet 23:91–99

Naranjo T, Corredor E (2008) Nuclear architecture and chromosome dynamics in the search of the pairing partner in meiosis in plants. Cytogenet Genome Res 120:320–330

Oliver KR, Greene WK (2009) Transposable elements: powerful facilitators of evolution. Genes Genomes 31:703–714

Pontes O, Neves N, Silva M, Lewis MS, Madlung A et al (2004) Chromosomal locus rearrangements are a rapid response to formation of the allotetraploid *Arabidopsis suecica* genome. Proc Natl Acad Sci U S A 101:18240–18245

Qi L, Friebe B, Zhang P, Gill BS (2007) Homoeologous recombination, chromosome engineering and crop improvement. Chromosom Res 15:3–19

Rizzon C, Marais G, Gouy M, Biémont C (2002) Recombination rate and the distribution of transposable elements in the *Drosophila melanogaster* genome. Genome Res 12:400–407

Salmon A, Flagel L, Ying B, Udall JA, Wendel JF (2010) Homoeologous nonreciprocal recombination in polyploid cotton. New Phytol 186(1):123–134

San FJ, Sung P, Klein H (2008) Mechanism of eukaryotic homologous recombination. Annu Rev Biochem 77:229–257

Sang T, Crawford DJ, Stuessy TF (1995) Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. Proc Natl Acad Sci U S A 92:6813–6817

Sanmiguel P, Bennetzen JL (1998) Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. Ann Bot 82:37–44

Schug MD, Hutter CM, Wetterstrand KA, Gaudette MS, Mackay TF et al (1998) The mutation rates of di-, tri- and tetranucleotide repeats in *Drosophila melanogaster*. J Neurosci Res 15:1751–1760

Shapiro JA (2005) A 21st century view of evolution: genome system architecture, repetitive DNA, and natural genetic engineering. Gene 345:91–100

Shapiro JA, Sternberg RV (2005) Why repetitive DNA is essential to genome function. Biol Rev 80:227–250

Sharma A, Wolfgruber TK, Presting GG (2013) Tandem repeats derived from centromeric retrotransposons. BMC Genomics 14:1–11

Song C, Liu S, Xiao J, He WG, Zhou Y et al (2012) Polyploid organisms. Sci China Life Sci 55:301–311

Stults DM, Killen MW, Williamson EP, Hourigan JS, Vargas HD et al (2009) Human rRNA gene clusters are recombinational hotspots in cancer. Cancer Res 69:9096–9104

Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW (1983) The double-strand-break repair model for recombination. Cell Sci 33:25–35

Toth G, Gaspari ZJ (2002) Microsatellites in different eukaryotic genomes: survey and analysis. Sociol Rural 46:40–60

Udall JA, Quijada PA, Osborn TC (2005) Detection of chromosomal rearrangements derived from homeologous recombination in four mapping populations of *Brassica napus* L. Genetics 169:967–979

Verstrepen KJ, An J, Lewitter F, Fink GR (2005) Intragenic tandem repeats generate functional variability. Nat Genet 37(9):986

Wang J, Ye LH, Liu QZ, Peng LY, Liu W et al (2015) Rapid genomic DNA changes in allotetraploid fish hybrids. Heredity 114:601–609

Wendel JF (2000) Genome evolution in polyploids. Plant Mol Biol 42: 225–249

Wendel JF, Schnabel T, Seelanan T (1995) Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). Proc Natl Acad Sci U S A 92(1):280–284

White SE, Wessler SR (1994) Retrotransposons in the flanking regions of normal plant genes: a role for *copia*-like elements in the evolution of gene structure and expression. Proc Natl Acad Sci U S A 91:11792–11796

Yang S, Arguello JR, Li X, Ding Y, Zhou Q et al (2008) Repetitive element-mediated recombination as a mechanism for new gene origination in, Drosophila. PLoS Genet 4:63–71

Zwierzykowski Z, Tayyar R, Brunell M, Lukaszewski AJ (1998) Genome recombination in intergeneric hybrids between tetraploid *Festuca pratensis* and *Lolium multiflorum*. J Hered 89:324–328