Extensive tRNA gene changes in synthetic *Brassica napus*

Lijuan Wei[1,a, b], Zeshan An[1,c], Annaliese S. Mason[d], Meili Xiao[a], Ying Guo[e], Jiaming Yin[a], Jiana Li[a], Donghui Fu[*,b]

[a] Chongqing Engineering Research Center for Rapeseed, College of Agronomy and Biotechnology, Southwest University, Chongqing 400716, China.

[b] Key Laboratory of Crop Physiology, Ecology and Genetic Breeding, Ministry of Education, Agronomy College, Jiangxi Agricultural University, Nanchang 330045, China.

[c] Crops Research Institute, Gansu Academy Of Agricultural Sciences, GanSu Lanzhou, 730070, China.

[d] School of Agriculture and Food Sciences and Centre for Integrative Legume Research, The University of Queensland, The University of Queensland, Brisbane 4072, Australia.

[e] Wheat Research Institute, Gansu Academy Of Agricultural Sciences, GanSu Lanzhou, 730070, China.

[1]These authors contributed equally to this work.

[*]Corresponding author

Email: fudhui@163.com

Telephone: +086-0791-83813142

Fax: +086-0791-83813185

Abstract Allopolyploidization, where two species come together to form a new species, plays a major role in speciation and genome evolution. Transfer RNAs (abbreviated tRNA) are typically 73 to 94 nucleotides in length, and are indispensable in protein synthesis, transferring amino acids to the cell protein synthesis machinery (ribosome). To date, the regularity and function of tRNA gene sequence variation during the process of allopolyploidization has not been well understood. In this study, the inter-tRNA gene corresponding to tRNA amplification polymorphism (ITAP) method was used to detect changes in tRNA gene sequences in the progeny of interspecific hybrids between *Brassica rapa* and *B. oleracea*, mimicking the original *B. napus* (canola) species formation event. Cluster analysis showed that tRNA gene variation during allopolyploidization did not appear to have a genotypic basis. Significant variation occurred in the early generations of synthetic *Brassica napus* (F_1 and F_2 generations), but fewer alterations were observed in the later generation (F_3). The variation-prone tRNA genes tended to be located in AT-rich regions. BlastN analysis of novel tRNA gene variants against a *Brassica* genome sequence database showed that the variation of these tRNA-gene-associated sequences in allopolyploidization might result in variation of gene structure and function, e.g. metabolic process and transport.

Key words: *Brassica napus*, tRNA gene, allopolyploidization, sequence variation

Introduction

Transfer RNA (tRNA) is usually 75-95 bp in length, and is classed as non-coding RNA. It is ubiquitous in all organisms, and accounts for about 15% of all RNAs in the cell, in between rRNA (approximately 75%) and mRNA (about 5%) in abundance. The roles and functions of tRNA have been widely studied. tRNA plays a critical role in protein synthesis, carrying an amino acid molecule to a peptide chain at the ribosome, through use of a three-letter anticodon corresponding to the three-letter codon of mRNA. tRNA molecules with specific anticodons are known as isoacceptors, and read different codons to produce the same amino acid. There are 21 isoacceptor families, 20 for standard amino acids and one for selenocysteine (Goodenbour and Pan 2006). The secondary structure of tRNA consists of two structural and functional domains: the top half domain, also called the "minihelix", consists of an acceptor stem and pseudouridine (TψC) stem-loop, and the bottom half domain is composed of a dihydrouridine (DHU) stem-loop and an anticodon (AC) stem-loop.

tRNA molecules are highly conserved during evolution. Variation of mitochondrial serine tRNA in 13 Euptychiina butterfly species was mainly observed on the TψC and DHU loop-arm, and variable loop through sequence alignment and structure alignment (Marin et al. 2012). Modern tRNAs are derived from tRNA halves mediated by repeat elements (Zuo et al. 2013). The bottom half is more recent than the top half, as deduced from specific codon recognition in mRNA (Sun and Caetano-Anolles 2008b; Sun and Caetano-Anolles 2008c). Furthermore, the tRNA cloverleaf structure can generally fold into the L-shaped 3D structure. Up until now, two types of tRNA genes have been discovered: continuous tRNA genes, which do not contain introns, and disrupted tRNA genes, which contain permuted, split and intron-containing tRNAs (Randau and Soll 2008).

Allopolyplodization merges two or more divergent sets of chromosomes (genomes). This process, which encompasses interspecific hybridization and subsequent chromosome doubling, is a vital process in the evolution, diversification and speciation of plants (Leitch and Leitch 2008). All angiosperms have experienced polyploidization once or more in their evolution history (Jiao et al. 2011). The genome shock produced by allopolyploid formation was first described by McClintock (McClintock 1984). This genome shock includes extensive genomic changes as a result of homologous and homeologous pairing, emergence of diverse gene expression patterns and gene redundancy, and transposon activation (Kenan-Eichler et al. 2011). A range of rapid genetic and epigenetic alterations accompanying allopolyplodization have been reported, such as DNA arrangement (Pires et al. 2004; Pontes et al. 2004), changes in chromosome number (Negron-Ortiz 2007) and structure (Lim et al. 2008), gain or loss of DNA (Song et al. 1995), transposon activation (Zhang et al. 2013; Zou et al. 2011), transcriptome shock (Hegarty et al. 2006), and epigenetic alterations such as methylation (Lukens et al. 2006; Zhang et al. 2013), histone modifications and variants (Madlung and Wendel 2013), and small RNA change (Ha et al. 2009; Kenan-Eichler et al. 2011).

According to U's triangle of *Brassica* (U 1935), the *Brassica* genus contains three diploid and three amphidiploid species. Genomes A, B, and C shared a common allohexaploid *Brassica* ancestor (Cheng et al. 2013; Röbbelen 1960). *Brassica napus* L. (AACC, $2n = 4x = 38$) is an important oil crop derived from interspecific hybridization and chromosome doubling between *Brassica rapa* (AA, $2n = 2x = 20$) and *Brassica oleracea* (CC, $2n = 2x = 18$). *B. napus* is also an important allopolyploid plant model, and has been used to investigate chromosome rearrangement and compensation so as to maintain genome stability (Szadkowski et al. 2011; Xiong et al. 2011) as well as transcriptome alteration and DNA methylation (Xu et al. 2009) in resynthesized allopolyploids. Little genetic change has been observed in the S_0 generation of resynthesized *B. napus*, but extensive alteration in DNA methylation has been detected (Lukens et al. 2006). Gene expression was found to vary greatly from parent levels in a non-additive fashion in the early stages of synthetic *B. napus* formation (Albertin et al. 2006), but with overall conservation of gene networks and pathways and protein functionality (Albertin et al. 2007). The differential regulation of protein expression in synthetic allopolyploids has been attributed to post-transcriptional modification rather than transcriptional changes, perhaps through the small RNA pathway (Marmagne et al. 2010). In addition, *de novo* variation in phenotypic traits in response to different growth conditions has been found in *B. napus* polyploid lines: in one study, 30% of lines were similar to one of the parents, 50% had intermediate values, and 20% were trangressive (Schranz and Osborn 2004).

Variation in gene expression and phenotype appear to be mainly caused by the exchange of homeologous chromosomes (Gaeta et al. 2007), particularly during the first meiosis of resynthesized lines, which acts as a “genome blender” (Szadkowski et al. 2010). Allopolyploidization also has a moderate effect on transposable element (TE) activation: highly specific TE activation events deriving from the *B. oleracea* genome have been identified (Sarilar et al. 2013). Despite the bulk of evidence for other kinds of genomic change in synthetic polyploids, how tRNA gene sequences change and function during the process of polyplodization has yet to be investigated in detail.

Is variation in tRNA gene sequences during the process of allopolyploidization consistent with variation observed for other types of sequences? How regular is tRNA gene sequence variation, and what is the function of this variation during allopolyploidization? To answer these questions, we investigated tRNA in F_1 hybrids derived from a cross between *B. rapa* and *B. oleracea* and subsequent chromosome doubling, and in their F_2 and F_3 progeny. Specific primers were designed for tRNA genes predicted by a tRNA-scan program, and were used to detect tRNA gene variation induced by allopolyploidization over three generations and to explore the potential effect of tRNA gene variation on the genome. Our results showed that tRNA gene variation during allopolyploidization did not appear to have a genotypic basis and tended to be located in AT-rich regions, displaying a discrete pattern when compared to other types of sequence variation. This study is beneficial to understanding the function of tRNA genes in allopolyploidization.

Results

Primer design

A total of 1157 putative tRNA genes were predicted from *B. rapa* genome sequences using tRNA predicting program “tRNAscan-SE”. These genes were classified into 23 groups according to the type of tRNA isoacceptor (see Fig. 1 for detailed information). There were 67 intron-containing tRNA genes, belonging to four types and making up 5.8% of the total number of tRNA genes. All of them had only one intron. Of the tRNA types with an intron, $tRNA^{Tyr}$ (26, 53.7%) and $tRNA^{Met}$ (36, 38.8%) made up the majority, with an average intron length of 11 bp. The longest intron in the tRNA Lys (CTT) gene was 192 bp, and the remaining four intron-containing tRNA genes were pseudogenes. In addition, only one selenocysteine tRNA gene and three suppressor tRNA genes were discovered, and all of them were discarded due to their low frequency. A total of 31 pseudo-tRNA genes (putatively non-functional) were also eliminated. Finally, 1122 tRNA sequences comprising isoacceptors for the 20 common amino acids were used for subsequent analysis. The largest number of tRNA genes were $tRNA^{Pro}$ type (101, 8.7%), followed by $tRNA^{Gly}$ (94, 8.1%) and $tRNA^{Leu}$ (83, 7.2%), and they were significantly more than the fewest number of tRNA gene type, $tRNA^{Trp}$ (30, 2.6%) ($P<0.01$).

Local all-by-all BlastN with E value set to $1.0E^{-6}$ was performed for these 20 isoacceptors, and the results were imported into the Cytoscape software to display the sequence relationships. ClustalX was performed for each class to find conserved regions for primer design. A total of 20 forward primers and 20 reverse primers were designed, such that every tRNA gene transferring a specific amino acid had one corresponding forward primer and one corresponding reverse primer (Table 1). A total of 400 primer pairs were combined randomly and used to amplify polymorphismic fragments using DNA templates from the two parents (*B. rapa* and *B. oleracea*) and 33 of their offspring lines (including three F_1 plants, 10 F_2 plants and 20 F_3 plants). Finally, 113 primer pairs were selected and a total of 486 polymorphic loci were produced, with an average of 4.3 loci per primer combination. Each pair of polymorphic primers amplified tRNA genes transferring one type or two types of amino acids. The PIC values for these loci ranged from 0.03 to 0.5 with an average of 0.32.

Frequency of genetic variants generated by allopolyploidization

In order to detect variation of tRNA genes during the process of allopolyploidization, "abnormal bands" different from parental bands were scored: novel bands and eliminated bands. Eliminated bands denoted bands present in both parents but absent in the progeny, while novel bands represented bands that did not exist in the parents but were present in the offspring. The number and frequency of abnormal bands over the three generations was scored in Table 2. Novel bands were significantly more frequent than eliminated bands ($P<0.01$) in both the F_1 (16.2% novel bands compared to 8.5% eliminated bands) and F_2 generations (29.8% novel bands and 22.2% eliminated bands). However, there was no difference between the frequency of novel bands (5.8%) and eliminated bands (6.6%) in the F_3 generation ($P>0.05$). Furthermore, the frequency of tRNA gene sequence variation increased from the F_1 generation (24.7%) to reach a maximum in the F_2 generation (52%), before reducing in the F_3

generation (12.4%). In summary, tRNA-gene-associated sequences varied most from the F_1 to F_2 generation, and more novel/duplicated sequences than eliminated sequences were observed during the *Brassica* allopolyploidization process.

In addition, bands could be divided into 20 groups according to the type of amino acids that the tRNA carried. The frequency of variants for tRNA carrying each of the 20 common amino acids (Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr and Val) is displayed in Fig. 2. Variation in the F_3 generation was the smallest of the three generations in each class of tRNA genes. Most classes of tRNA genes showed the greatest changes from the F_1 to the F_2 generation, but some tRNA genes were an exception: those transferring Ala (44.6%), Asp (51.5%), His (37.2%), Leu (45.8%), Met (45.9%) and Thr (43.4%) had the largest variation in F_1 hybrids.

Cluster analysis

To judge whether the phylogenetic relationship based on tRNA loci was consistent with the pedigree of the materials over the three generations, cluster analysis of 33 selfing lines (F_1, F_2, and F_3) was performed with 199 parental polymorphic loci (excluding loci showing variation) using UPGMA (Fig. 3a). Line A30 in the F_3 generation, line A20 in the F_2 generation and line A33 in the F_3 generation clustered into one group respectively, and all other lines clustered into another group. This failed to correspond to the known pedigree relationships of these materials.

In a complementary analysis, in order to detect whether variation in tRNA loci have a genotypic basis during allopolyploidization, all 486 polymorphic loci including eliminated loci and novel loci were clustered across the three generations (Fig. 3b). Results were similar to the previous cluster analysis (Fig. 3a), with the 33 offspring lines clustered into three groups: line A20 in the F_2 generation and line A33 in the F_3 generation were assigned into two separate groups, and the remaining materials clustered into another big group. tRNA loci variation was so extensive that phylogenetic relationships between individuals could not be deduced from tRNA loci after a generation.

Sequencing analysis

In order to investigate the nature of the tRNA variants, 24 bands which were eliminated in the progeny were extracted from the parents, 80 novel bands were extracted from the progeny lines, and positive clones were sequenced. After removing clone vectors, 94 non-repetitive sequences were obtained. To confirm whether the variation of these sequences was related to tRNA genes, we performed local BlastN analysis (E≤e-10) against the set of tRNA genes in the *B. rapa* genome which were used to design the starting primers. Of the 94 sequences, 70 matched well with predicted tRNA genes: 24 eliminated bands and 46 novel bands. The remaining 24 sequences did not match and were discarded. Matched sequences ranged from 88 to 1131 bp in length. Details of the tRNA-gene-associated sequences are shown in Supplementary table 1.

To explore the characteristics of these 70 tRNA-gene-associated sequence variants, homologous sequences were obtained through BlastN analysis against the A and C genomes of the BrBo.chr database, using the lowest E value and higher scores to select the most homologous sequence match. The two 100 bp sequences flanking each homologous sequence were also extracted to produce the final homologous sequences (Supplementary table 2), as the amplified fragments might contain parts of tRNA genes rather than entire tRNA genes. From *in silico* mapping, these sequences were located across 9/10 chromosomes of *B. rapa* (all but A6) and all 9 chromosomes of *B. oleracea*. Using these sequences, 46 tRNA genes formed by 16 tRNA types in 45 homologous sequences (64.3%) were predicted by the tRNA-scan program. The location of the tRNA genes, the type of tRNA and the detailed anticode and intron information are shown in Table 3. The most frequent tRNA type was $tRNA^{Gly}$ (GCC) (11, 24.4%), followed by $tRNA^{Pro}$ (TGG) (7, 15.5%), and all other tRNA types had only one or two sequences. This was consistent with the tRNA prediction results from before in the *B. rapa* genome, and it is likely that these two tRNA types had more primers than other types.

In addition, in order to understand the genomic environment of the tRNA gene variants, the GC content of the 50 bp flanking sequences for the predicted tRNA genes was calculated. The average AT content of the 5′ end flanking sequences varied from 46% to 86% with an average of 71%, while the AT content of the 3′ end flanking sequences ranged from 56% to 86% with an average of 73%, suggesting that the tRNA-gene-associated sequence variants tended to be located in AT-rich regions.

Gene function

All tRNA-gene-associated sequences were used to perform BlastN analysis against the *B. rapa* and *B. oleracea* coding genes in the BRAD database with default values. According to the minimum E values and the maximum scores, 49 tRNA-gene-associated sequences out of 70 sequences (71%) showed high homology with these coding genes, suggesting that these tRNA-gene-associated sequences could result in variation of gene structure and putatively gene function. The detailed alignment information is shown in Supplementary table 3.

The functional annotation of coding sequences based on cellular composition, molecular function and biological process is shown in Fig. 4. For cellular composition, the genes were mainly operative in the organelles (34.4%), including in the chloroplast (15.6%), ribosome (9.4%) and vacuole (9.4%), followed by the cell membrane (34.4%), nucleolus (15.6%), cytoplasm (6.3%), macromolecular complex (6.3%) and cytoskeleton (3.1%). In the annotation of molecular function, binding function occupied the largest proportion (52.8%), including DNA binding, rRNA binding, protein binding, lipid binding, metal ion binding and carbohydrate binding. The remaining molecular functions included catalytic activity (19.4%), transporter activity (11.1%), structural composition (8.3%), transcription activity (2.8%), metallochaperone activity (2.8%) and receptor activity (2.8%). The biological processes were mainly metabolic processes (26.2%), transport (19%), response to stimulus (19%) and translation (11.9%). Overall, the key function of these genes was binding, gene expression mainly occurred in organelles and in the cell membrane, and these genes participated in metabolic process and transport.

Discussion

Hybridization and allopolyploidization, where two species come together to form a new species, are major processes in speciation and genome evolution. Extensive genomic changes are induced by hybridization and genomic introgression, including microsatellite mutation, retrotransposon mobilization, and epigenetic changes (Dong et al. 2013; Zhang et al. 2013; Zou et al. 2011), and these changes often facilitate the establishment of new species (Feldman and Levy 2005; Feldman and Levy 2009). tRNA plays a major role in protein synthesis, transferring amino acids to the protein synthesis machinery. However, how tRNA genes are affected by and vary as a result of allopolyploidization is unknown. In this study, we employed special ITAP markers (inter-tDNA markers corresponding to tRNA amplification polymorphism) to determine tRNA gene alterations over three generations of synthetic *B. napus*, to explore the characteristics of tRNA gene variants, and to investigate tRNA gene function.

The ITAP method was used to detect tRNA gene variants in this study, which could distinguish novel tRNA sequence variation from chromosome rearrangements due to homeologous recombination. Chromosome rearrangement caused by homeologous chromosome replacement and rearrangements was common in the early generations of resynthesized *B. napus* and it was responsible for the production of novel bands and phenotypic variation (Gaeta et al. 2007; Xiong et al. 2011). However, the detected fragments between one or two tRNA genes ranged from 88 to 1131 bp in length, and chromosome breakage was less likely to occur in the fragments. The occurrence of novel bands was due to tRNA gene variants. For eliminated bands, they might be resulted from the lost of chromosome fragments including tRNA genes or chromosome rearrangement during the allopolyploidization.

The tRNA gene variation occurred in early stages of allopolyploidization

Actually, many studies about chromosome rearrangements and homeologous pairing in synthetic *B. napus* have been demonstrated (Nicolas et al. 2007; Song et al. 1995; Udall et al. 2005; Xiong et al. 2011), but little is known about the sequence variation of tRNA genes. In this study, 131 ITAP primers were used to detect tRNA gene changes over three generations (33 offspring) after crossing between *B. rapa* and *B. oleracea* to generate resynthesized *B. napus*. Appearance of tRNA gene variants in the early generations (F_1 or F_2) was the highest, and appearance of tRNA gene variants was the lowest in the F_3 generation for all tRNA loci and for each of the 20 classes of tRNA loci carrying common amino acids. The results were consistent with previous studies of genetic and genomic variation after hybridization: Zou et al. (2011) found that the reactivation of most LTR retrotransposons, which was not due to chromosome rearrangement, occurred in the early generations after interspecific hybridization in *Brassica*. Likewise, Szadkowski et al. (2010) described the first meiosis of resynthesized *B. napus* as a “genome blender”. Most microsatellite variation was also found to occur in the early stages after alien DNA introgression in rice (Dong et al.

2013).

Genetic incompatibility between the two divergent genomes that are merged together in the new allopolyploid is thought to rapidly give way under selection pressure for stable genome inheritance and gene expression (Chen 2007; Leitch and Leitch 2008; Osborn et al. 2003; Otto and Whitton 2000). Significant F_2 fitness breakdown was observed in copepod (*Tigriopus californicus*) between the parental generation and early generation hybrids, but hybrid superiority in later generations preceded hybrid inferiority in the early stage (Edmands et al. 2005). In legume *Chamaecrista fasciculate*, the fitness of a recombinant F_6 population, including the ability to survive and reproduce, demonstrated a strong recovery (Erickson and Fenster 2006).

In our study, the frequency of novel bands was significantly more than that of eliminated bands in the F_1 and F_2 generations, indicating that tRNA-gene-associated sequences altered continuously after polyploidization to adaptive differentiation. Extensive variation of tRNA genes mainly occurred in the early generations of allopolyploidization, which might result in chromosomal difference contributing to hybrid breakdown. The small alterations in the later generations, e.g. F_3 generation, might be beneficial to hybrid superiority. This result was consistent with previous evidence that hybrid breakdown might be temporarily followed by hybrid superiority (Hwang et al. 2011). Novel tRNA-gene-associated sequences may contribute to novel genetic functionality that can aid in allopolyploid establishment (Mallet 2007), whereas loss of tRNA-gene-associated sequences may lead to loss-of-function mutations and hence deleterious phenotypic effects.

tRNA gene evolution

Two cluster analyses were performed: one with only ITAP alleles not derived from abnormal bands, and one with all alleles including novel and eliminated variants. For the first cluster analysis the results did not correspond to the pedigree of the lines, which showed that tRNA gene differentiation was highly prevalent. The clustering results for all loci (including loci containing variants) were similar to the clustering results excluding loci containing variants. The failure of tRNA genes from the 199 tRNA loci to group in pedigrees in the phylogenetic tree in our study suggests that tRNA gene diversification is extremely rapid. This may support the idea that tRNA gene diversification preceded the divergence of species through all tRNA sequences analysis (Sun and Caetano-Anolles 2008b). tRNA and organisms did not co-evolve, and tRNA was created from a single hairpin duplication (Widmann et al. 2005). It is commonly accepted that initially the acceptor stem sequences only attached a few amino acid sequences, and not all 20 amino acids were encoded (Saks and Sampson 1995; Sun and Caetano-Anolles 2008a). If tRNA loci are prone to rapid genetic differentiation over short time frames, this may explain the evolution of tRNA loci to carry all 20 common amino acids in most organisms.

In our previous results assessing transposon-induced sequence variation in synthetic *B. napus* (An et al., in press), synthetic lines were clustered by generations, with each generation more similar to itself than to its parents or progeny. However, tRNA-gene-associated sequences may evolve and differentiate far more irregularly than transposon-associated sequences, possibly having great effect in the allopolyploidization process. Hence, tRNA gene alterations may comprise a major force to drive the differentiation of lines and accelerate the process of polyploidization.

Characteristics and gene functionality of the tRNA-gene-associated sequences

The homologous sequences plus 100 bp flanking sequences of 70 tRNA loci were found through BlastN analysis in the *Brassica* database: 46 tRNA genes formed by 16 tRNA types in 45 homologous sequences (64.3%) were successfully predicted. The two main types of tRNA genes, $tRNA^{Gly}$ and $tRNA^{Pro}$, had high variation frequency, but also accounted for relatively large proportion of tRNA loci identified in the parent *B. rapa* genome. In addition, these tRNA genes were usually located in AT-rich regions. Similarly, the 50 bases upstream of tRNA gene sequences are known to be rich in A and T residues in flowering plants *Arabidopsis thaliana*, *Medicago truncatula*, *Populus trichocarpa*, *Oryza sativa* and *Brachypodium distachyon* (Michaud et al. 2011).

To sum up, extensive tRNA gene alternations were induced in the early stages of allopolyploidization. tRNA variation did not correspond to either generations or pedigrees. The 50 nucleotides upstream and downstream of tRNA genes were in AT-enriched sequences. This region is generally associated with non-coding regions which may be under less selective pressure and the variation is more easily accumulated. tRNA gene variation may have a substantial impact on coding gene functionality, particularly with respect to metabolic processes and transport in the organelles (chloroplast), and cell membrane, since the chloroplast genome was important in encoding all tRNAs for translation (Michaud et al. 2011).

Materials and methods

Materials

Variation of tRNA genes during the process of polyploidization was detected in the parents *B. rapa* (A1) and *B. oleracea* (A2) and in their synthetic *B. napus* offspring ($2n=4x=A^rA^rC^oC^o$). Three F_1 plants (A3, A4 and A5), 10 F_2 plants (A8, A10, A11, A12 and A13 from A3 selfings, A16 and A20 from A4, and A7, A21 and A23 from A5) and 20 F_3 plants from eight F_2 plants whose chromosomal number was 38 by cytological observation (Li and Heneen 1999) were investigated. Detailed characteristics and pedigrees of the lines used in this study are described in An et al. (in press).

Prediction of tRNA genes in the *B. rapa* genome

Approximately 290 Mb of *B. rapa* genome sequence (v 1.2) was downloaded from the BRAD database (http://brassicadb.org/brad/) (Cheng et al. 2011). tRNA gene sequences were predicted using tRNAscan-SE v.1.23 software (http://lowelab.ucsc.edu/tRNAscan-SE/) (Schattner et al. 2005) with default values.

Primer design

Predicted tRNA gene sequences were used to perform local all-by-all BlastN by Bioedit, using a cut-off E value of < $1.0E^{-6}$ (Hall 1999). Resulting gene sequences were classified into different classes and displayed using Cytoscape 2.0 (Smoot et al. 2011) with default values. For each class of tRNA gene, the conserved sequences were obtained by sequence alignment using ClustalX software (Thompson et al. 1997), and were then used to design primers using Primer 3 (Rozen and Skaletsky 2000). The parameters for primer design were set as follows: the average annealing temperature was 60°C (range from 55°C to 65°C), GC content was 40-60% and primer length was 18-27 bp.

Cluster analysis

Total genomic DNA for all materials was extracted using the CTAB method (Rogers and Bendich 1985). The forward and reverse primers designed from tRNA gene sequences were combined randomly and screened between the parents and their offspring. Inter-tRNA amplification polymorphism (ITAP) was detected using 6% denaturing polyacrylamide gel electrophoresis to visualise DNA fragments. A sketch of PCR products amplified with ITAP primers is shown in Fig. 5, and these sequences are also referred to as amplified tRNA-gene-associated sequences. Polymorphic primers were then used to amplify sequences in the parents, F_1, F_2 and F_3 plants. Clear bands were scored as 1 (present band) or 0 (absent band). The polymorphic information content (PIC) and cluster analyses were performed using the statistical software package NTSYS-pc 2.02 (Rohlf 1998) using the unweighted pair group method of arithmetic means (UPGMA).

Variation frequency

In order to understand the variation of tRNA-gene-associated sequences during the process of polyploidization, abnormal DNA fragments exhibiting novel variation (eliminated and novel bands) were assessed. An eliminated band refers to a band that is present in both parents but that disappeared in some F_1, F_2, or F_3 lines, interpreted as the disappearance of tRNA-gene-associated sequences. Novel bands were defined as bands that were absent in both parents but present in the F_1, F_2, or F_3 lines, denoting the generation of new tRNA-gene-associated sequences and the variation of tRNA-gene-associated sequences. DNA from both parents was mixed in a 1:1 ratio as a template control to avoid experimental error. The variation frequency (F) was calculated according the formula: $F = I/K*100\%$, where I denoted the number of eliminated bands (or novel bands) and K denoted the total number of all bands appearing in the offspring.

Acquisition and sequencing of abnormal bands

Abnormal fragments including novel bands and eliminated bands in the F_1, F_2, or F_3 lines were extracted from denaturing polyacrylamide gel, dissolved in 20 μl of double-distilled water and heated at 95°C for 20 minutes before finally being cooled to room temperature. The eluted DNA was used as the template for PCR amplification, and then PCR products were purified on a 1.5% agarose gel. After ligating into the pMD19-T plasmid vector (TaKaRa, Japan), the recombinant plasmid was transformed into *Escherichia coli* and cultivated overnight in a lysogeny broth (LB) medium at 16 °C. Two positive clones per band were selected and sequenced.

The validation of tRNA sequences

The vector sequences of sequenced fragments were removed using VecScreen (NCBI website (http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html). These sequences were aligned with the putative tRNA genes in the *B. rapa* genome which were used to design primers in order to confirm the existence of the tRNA genes in the sequenced fragments and to assess if the tRNA genes were the origin of the genetic variation in the sequenced samples.

Sequence analysis

The validated sequences were used to perform BlastN analysis against the *B. rapa* and *B. oleracea* genomes from the “BrBo.chr” database (http://www.oilcrops.info/), with an E value cutoff of < $1.0E^{-10}$. The homologous sequences were selected and extracted according to the maximum E value, and if the E values of more than one pair of alignments were equal, then the hit with the highest score was chosen. The hit sequences as well as the corresponding flanking 100 bp sequences were also extracted as the final homologous sequences. To identify the type and characteristics of variation-prone tRNA gene sequences, the homologous sequences were used to perform tRNA gene prediction using tRNAscan-SE v.1.23 with default values. The 50 bp flanking sequences of these hits were also extracted, and their GC contents were calculated using the software program Bioedit (Hall 1999).

Functional annotation

To determine the effect of tRNA-gene-associated sequences on gene function, BlastN analysis with default values was performed against coding genes of *B. rapa* and *B. oleracea* in the BRAD database. The target coding genes were selected on the basis of the smallest E values, and if the E values of several pairs of alignments were the same, the hit with the greatest score was chosen. The coding gene sequences chosen were obtained through the BRAD database (http://brassicadb.org/brad/). Putative gene functionality, including molecular function, biological processes, and cellular composition, was predicted using Blast2Go (www.hindawi.com/journals/ijpg/2008/619832.html) (Conesa and Gotz 2008) with default values.

Acknowledgements

This work was supported financially by National Natural Science Foundation of China (code: 31260335), and Jiangxi Science and Technology Support Program (code: 20132BBF60013). A. Mason is supported by an Australian Research Council Discovery Early Career Researcher Award (DE120100668).

Figure Captions

Fig. 1 The number of each of 23 classes of tRNA sequences carrying different amino acids

Fig. 2 Frequency of tRNA gene variants for tRNAs carrying each of the 20 common amino acids

Fig. 3 Dendrograms of cluster analysis based on inter-tRNA amplification polymorphism (ITAP) markers using UPGMA. Fig. 3a was made by 199 parental polymorphic loci (excluding loci showing novel or eliminated bands) in 33 self-pollinated lines derived from interspecific hybridization between *B. rapa* and *B. oleracea*. Fig. 3b was made by incorporating all 486 polymorphic loci in *B. rapa*, *B. oleracea* and their offspring (F_1, F_2, and F_3)

Fig. 4 Functional annotation of *Brassica* genes associated with tRNA-gene-associated sequences

Fig. 5 The sketch of primer design and amplified products of an inter-tRNA amplification polymorphism (ITAP) marker. Gray areas denote a tRNA gene sequence. Black arrows indicate primer-binding regions and direction of amplification. Black lines indicate sequence between the two tRNA genes. The PCR amplification product is indicated by the black line between the two arrows, and is referred to as a tRNA-gene-associated sequence

Table 1 Detailed information for tRNA primers in *Brassica*

Table 2 The number and frequency of abnormal band variation for tRNA loci in synthetic *Brassica napus*

Table 3 The basic characteristics of predicted tRNAs in 45 homologous *Brassica* sequences

Supplementary table 1 tRNA-gene-associated sequence information

Supplementary table 2 Homologous sequence and 100 bp flanking sequences of tRNA genes

Supplementary table 3 Alignments between tRNA-gene-associated sequences and genes in *Brassica rapa* and *Brassica oleracea*

References

Albertin W, Alix K, Balliau T, Brabant P, Davanture M, Malosse C, Valot B, Thiellement H (2007) Differential regulation of gene products in newly synthesized *Brassica napus* allotetraploids is not related to protein function nor subcellular localization. BMC Genomics 8:56

Albertin W, Balliau T, Brabant P, Chevre AM, Eber F, Malosse C, Thiellement H (2006) Numerous and rapid nonstochastic modifications of gene products in newly synthesized *Brassica napus* allotetraploids. Genetics 173:1101-1113

Chen ZJ (2007) Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. Annu Rev Plant Biol 58:377-406

Cheng F, Liu S, Wu J, Fang L, Sun S, Liu B, Li P, Hua W, Wang X (2011) BRAD, the genetics and genomics database for *Brassica* plants. BMC Plant Biol 11:136

Cheng F, Mandakova T, Wu J, Xie Q, Lysak MA, Wang X (2013) Deciphering the diploid ancestral genome of the Mesohexaploid *Brassica rapa*. Plant Cell 25:1541-1554

Conesa A, Gotz S (2008) Blast2GO: A comprehensive suite for functional analysis in plant genomics. Int J Plant Genomics 2008:619832

Dong ZY, Wang HY, Dong YZ, Wang YM, Liu W, Miao GJ, Lin XY, Wang DQ, Liu B (2013) Extensive microsatellite variation in rice induced by introgression from wild rice (*Zizania latifolia* Griseb.). Plos One 8:e62317

Edmands S, Feaman HV, Harrison JS, Timmerman CC (2005) Genetic consequences of many generations of hybridization between divergent copepod populations. Journal of Heredity 96:114-123

Erickson DL, Fenster CB (2006) Intraspecific hybridization and the recovery of fitness in the native legume *Chamaecrista fasciculata*. Evolution 60:225-233

Feldman M, Levy AA (2005) Allopolyploidy - a shaping force in the evolution of wheat genomes. Cytogenetic and Genome Research 109:250-258

Feldman M, Levy AA (2009) Genome evolution in allopolyploid wheat-a revolutionary reprogramming followed by gradual changes. Journal of Genetics and Genomics 36:511-518

Gaeta RT, Pires JC, Iniguez-Luy F, Leon E, Osborn TC (2007) Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. Plant Cell 19:3403-3417

Goodenbour JM, Pan T (2006) Diversity of tRNA genes in eukaryotes. Nucleic Acids Research 34:6137-6146

Ha M, Lu J, Tian L, Ramachandran V, Kasschau KD, Chapman EJ, Carrington JC, Chen XM, Wang XJ, Chen ZJ (2009) Small RNAs serve as a genetic buffer against genomic shock in Arabidopsis interspecific hybrids and allopolyploids. Proceedings of the National Academy of Sciences of the United States of America 106:17835-17840

Hall TA (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser: 95-98

Hegarty MJ, Barker GL, Wilson ID, Abbott RJ, Edwards KJ, Hiscock SJ (2006) Transcriptome shock after interspecific hybridization in senecio is ameliorated by genome duplication. Curr Biol 16:1652-1659

Hwang AS, Northrup SL, Alexander JK, Vo KT, Edmands S (2011) Long-term experimental hybrid swarms between moderately incompatible Tigriopus californicus populations: hybrid inferiority in early generations yields to hybrid superiority in later generations. Conservation Genetics 12:895-909

Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS, Soltis DE, Clifton SW, Schlarbaum SE, Schuster SC, Ma H, Leebens-Mack J, dePamphilis CW (2011) Ancestral polyploidy in seed plants and angiosperms. Nature 473:97-100

Kenan-Eichler M, Leshkowitz D, Tal L, Noor E, Melamed-Bessudo C, Feldman M, Levy AA (2011) Wheat hybridization and polyploidization results in deregulation of small RNAs. Genetics 188:263-272

Leitch AR, Leitch IJ (2008) Genomic plasticity and the diversity of polyploid plants. Science 320:481-483

Li Z, Heneen WK (1999) Production and cytogenetics of intergeneric hybrids between the three cultivated *Brassica* diploids and *Orychophragmus violaceus*. Theor Appl Genet 99:694-704

Lim KY, Soltis DE, Soltis PS, Tate J, Matyasek R, Srubarova H, Kovarik A, Pires JC, Xiong Z, Leitch AR (2008) Rapid chromosome evolution in recently formed polyploids in *Tragopogon* (*Asteraceae*). PLoS One 3:e3353

Lukens LN, Pires JC, Leon E, Vogelzang R, Oslach L, Osborn T (2006) Patterns of sequence loss and cytosine methylation within a population of newly resynthesized *Brassica napus* allopolyploids. Plant Physiol 140:336-348

Madlung A, Wendel JF (2013) Genetic and epigenetic aspects of polyploid evolution in plants. Cytogenet Genome Res 140:270-285

Mallet J (2007) Hybrid speciation. Nature 446:279-283

Marin MA, Lopez A, Uribe SI (2012) Interspecific variation in mitochondrial serine transfer RNA (UCN) in Euptychiina butterflies (Lepidoptera: Satyrinae): Structure and alignment. Mitochondrial DNA 23:208-215

Marmagne A, Brabant P, Thiellement H, Alix K (2010) Analysis of gene expression in resynthesized *Brassica napus* allotetraploids: transcriptional changes do not explain differential protein regulation. New Phytol 186:216-227

McClintock B (1984) The significance of responses of the genome to challenge. Science 226:792-801

Michaud M, Cognat V, Duchene AM, Marechal-Drouard L (2011) A global picture of tRNA genes in plant genomes. Plant Journal 66:80-93

Negron-Ortiz V (2007) Chromosome numbers, nuclear DNA content, and polyploidy in *Consolea* (*Cactaceae*), an endemic cactus of the Caribbean Islands. Am J Bot 94:1360-1370

Nicolas SD, Le Mignon G, Eber F, Coriton O, Monod H, Clouet V, Huteau V, Lostanlen A, Delourme R, Chalhoub B, Ryder CD, Chevre AM, Jenczewski E (2007) Homeologous recombination plays a major role in chromosome rearrangements that occur during meiosis of *Brassica napus* haploids. Genetics 175:487-503

Osborn TC, Pires JC, Birchler JA, Auger DL, Chen ZJ, Lee HS, Comai L, Madlung A, Doerge RW, Colot V, Martienssen RA (2003) Understanding mechanisms of novel gene expression in polyploids. Trends in Genetics 19:141-147

Otto SP, Whitton J (2000) Polyploid incidence and evolution. Annu Rev Genet 34:401-437

Pires JC, Zhao JW, Schranz ME, Leon EJ, Quijada PA, Lukens LN, Osborn TC (2004) Flowering time divergence and genomic rearrangements in resynthesized *Brassica* polyploids (*Brassicaceae*). Biological Journal of the Linnean Society 82:675-688

Pontes O, Neves N, Silva M, Lewis MS, Madlung A, Comai L, Viegas W, Pikaard CS (2004) Chromosomal locus rearrangements are a rapid response to formation of the allotetraploid *Arabidopsis suecica* genome. Proc Natl Acad Sci USA 101:18240-18245

Randau L, Soll D (2008) Transfer RNA genes in pieces. Embo Reports 9:623-628

Röbbelen G (1960) Contributions to the analysis of the *Brassica*-genome. Chromosoma 11:205-228

Rogers SO, Bendich AJ (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Plant Molecular Biology 5:69-76

Rohlf FJ (1998) NTSYSpc: Numerical taxonomy and multivariate analysis system, version 2.01. Department of ecology and evolution, State University of New York

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132:365-386

Saks ME, Sampson JR (1995) Evolution of tRNA recognition systems and tRNA gene sequences. J Mol Evol 40:509-518

Sarilar V, Palacios PM, Rousselet A, Ridel C, Falque M, Eber F, Chevre AM, Joets J, Brabant P, Alix K (2013) Allopolyploidy has a moderate impact on restructuring at three contrasting transposable element insertion sites in resynthesized *Brassica napus* allotetraploids. New Phytologist 198:593-604

Schattner P, Brooks AN, Lowe TM (2005) The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. Nucleic Acids Research 33:W686-W689

Schranz ME, Osborn TC (2004) De novo variation in life-history traits and responses to growth conditions of resynthesized polyploid *Brassica napus* (*Brassicaceae*). American Journal of Botany 91:174-183

Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T (2011) Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics 27:431-432

Song K, Lu P, Tang K, Osborn TC (1995) Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. Proc Natl Acad Sci USA 92:7719-7723

Sun FJ, Caetano-Anolles G (2008a) Evolutionary patterns in the sequence and structure of transfer RNA: A window into early translation and the genetic code. Plos One 3:e2799

Sun FJ, Caetano-Anolles G (2008b) The origin and evolution of tRNA inferred from phylogenetic analysis of structure. J Mol Evol 66:21-35

Sun FJ, Caetano-Anolles G (2008c) Transfer RNA and the origins of diversified life. Sci Prog 91:265-284

Szadkowski E, Eber F, Huteau V, Lode M, Coriton O, Jenczewski E, Chevre AM (2011) Polyploid formation pathways have an impact on genetic rearrangements in resynthesized *Brassica napus*. New Phytologist 191:884-894

Szadkowski E, Eber F, Huteau V, Lode M, Huneau C, Belcram H, Coriton O, Manzanares-Dauleux MJ, Delourme R, King GJ, Chalhoub B, Jenczewski E, Chevre AM (2010) The first meiosis of resynthesized *Brassica napus*, a genome blender. New Phytologist 186:102-112

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876-4882

U N (1935) Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. Japanese Journal of Botany 7:389-452

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

Widmann J, Di Giulio M, Yarus M, Knight R (2005) tRNA creation by hairpin duplication. J Mol Evol 61:524-530

Xiong Z, Gaeta RT, Pires JC (2011) Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid *Brassica napus*. Proc Natl Acad Sci USA 108:7908-7913

Xu Y, Zhong L, Wu X, Fang X, Wang J (2009) Rapid alterations of gene expression and cytosine methylation in newly synthesized *Brassica napus* allopolyploids. Planta 229:471-483

Zhang X, Ge X, Shao Y, Sun G, Li Z (2013) Genomic change, retrotransposon mobilization and extensive cytosine methylation alteration in *Brassica napus* introgressions from two intertribal hybridizations. PLoS One 8:e56346

Zou J, Fu D, Gong H, Qian W, Xia W, Pires JC, Li R, Long Y, Mason AS, Yang TJ, Lim YP, Park BS, Meng J (2011) De novo genetic variation associated with retrotransposon activation, genomic rearrangements and trait variation in a recombinant inbred line population of *Brassica napus* derived from interspecific hybridization with *Brassica rapa*. Plant J 68:212-224

Zuo Z, Peng D, Yin X, Zhou X, Cheng H, Zhou R (2013) Genome-wide analysis reveals origin of transfer RNA genes from tRNA halves. Mol Biol Evol doi: 10.1093/molbev/mst107

Table 1 Detailed information for tRNA primers in *Brassica*

Primer Code	tRNA type	Primer	Length	Tm	GC%	Sequence
tR1	Ala	Forward primer	20	60.21	50	TTGGTCTTATGAGCCGAAGG
tR2	Ala	Reverse primer	21	60.07	52.4	CAACGCTCTAACCAACTGAGC
tR3	Arg	Forward primer	20	59.87	50	ATAAGGCGCTGGTCTACGAA
tR4	Arg	Reverse primer	20	59.87	50	TTCGTAGACCAGCGCCTTAT
tR9	Asn	Forward primer	20	60.07	55	AGGTTCGACCCCTCCTTCTA
tR10	Asn	Reverse primer	20	60.32	55	GGTTAACAGCCACACGCTCT
tR11	Asp	Forward primer	20	59.45	50	ACAGGCGGGAATACTTACCA
tR12	Asp	Reverse primer	20	59.45	50	TGGTAAGTATTCCCGCCTGT
tR13	Cys	Forward primer	20	59.81	50	ATCAATAGGTCACCGGTTCG
tR14	Cys	Reverse primer	20	59.81	50	CGAACCGGTGACCTATTGAT
tR15	Gln	Forward primer	21	59.72	47.6	TCCAATGTCCTAACCGCTAGA
tR16	Gln	Reverse primer	20	59.56	55	GCGGTTAGGACATTGGACTC
tR17	Glu	Forward primer	20	59.83	60	CAGGATACTCGGCTCTCACC
tR18	Glu	Reverse primer	20	59.83	60	GGTGAGAGCCGAGTATCCTG
tR19	Gly	Forward primer	20	59.08	55	CCGTGGCAGGGTACTATTCT
tR20	Gly	Reverse primer	20	61.13	55	GTACAGACCCGGGTTCGATT
tR23	His	Forward primer	20	58.98	50	CAACGTGGAATTCTCACCAC
tR24	His	Reverse primer	20	58.98	45	TGGTGAGAATTCCACGTTGT
tR25	Ile	Forward primer	20	59.63	50	TAGCACGACGCTCTAACCAA
tR26	Ile	Reverse primer	20	60.28	50	CGTGCTAATAACGCGAAGGT
tR27	Leu	Forward primer	20	59.82	50	ACCAACTCGGCCATATCAAC
tR28	Leu	Reverse primer	20	59.82	50	GTTGATATGGCCGAGTTGGT
tR31	Lys	Forward primer	20	59.75	55	TAAGAGCCTTGCGCTCTACC
tR32	Lys	Reverse primer	20	61.03	55	AGAGCGCAAGGCTCTTAACC
tR35	Met	Forward primer	20	60.28	60	GCTCGAACTCTCGACCTCAG
tR36	Met	Reverse primer	20	60.13	60	CCTGAGGTCGAGAGTTCGAG
tR41	Phe	Forward primer	20	60.16	55	ACGCTCTCCCAACTGAGCTA
tR42	Phe	Reverse primer	20	59.87	50	AAGATCTGAAGGTCGCGTGT
tR43	Pro	Forward primer	20	59.53	50	CCCAAAGCGAGAATCATACC
tR44	Pro	Reverse primer	20	60.6	55	GAGAGGTCCCGAGTTCGATT
tR45	Ser	Forward primer	20	60.08	45	AGAAATCATGTGGGCTTTGC
tR46	Ser	Reverse primer	20	60.08	45	GCAAAGCCCACATGATTTCT
tR49	Thr	Forward primer	20	59.9	55	CGCTTACTAAACGGGTGCTC
tR50	Thr	Reverse primer	20	60.77	55	AAGCGGGAGGTCTTGAGTTC
tR55	Trp	Forward primer	20	59.74	45	AGAAGGTTGCGTGTTCGATT
tR56	Trp	Reverse primer	20	59.74	45	AATCGAACACGCAACCTTCT
tR57	Tyr	Forward primer	21	59.62	52.4	CTCAGTTGGTAGAGCGGAAGA
tR58	Tyr	Reverse primer	20	60.21	50	GAATCGAACCAGCGACCTAA
tR59	Val	Forward primer	20	59.15	55	ACACTGAAGGTCTCCGGTTC
tR60	Val	Reverse primer	20	60.15	55	AACCGGAGACCTTCAGTGTG

Table 2 The number and frequency of abnormal band variation for tRNA loci in synthetic *Brassica napus*

	P to F1		F1 to F2		F2 to F3	
	I	F	I	F	I	F
Eliminated bands	32	8.5%	103	22.2%	28	5.8%
Novel bands	61	16.2%	138	29.8%	32	6.6%
Total bands (K)	376	-	463	-	481	-

I denoted the number of eliminated bands (or novel bands) appearing in the offspring

K denoted the total number of all bands appearing in the offspring

The variation frequency (F) was calculated according the formula: F = I/K*100%

Table 3 The basic characteristics of predicted tRNAs in 45 homologous *Brassica* sequences

Code	tRNA Begin	tRNA End	tRNA Type	Anticode	Intron Begin	Intron End
3	348	278	Gly	GCC	0	0
4	63	133	Gly	GCC	0	0
5	203	133	Gly	GCC	0	0
6	59	129	Gly	GCC	0	0
7	16	86	Gly	GCC	0	0
8	63	133	Gly	GCC	0	0
14	50	123	Asn	GTT	0	0
15	62	133	Pro	TGG	0	0
18	120	36	Tyr	GTA	83	72
19	156	85	Cys	GCA	0	0
20	513	441	Glu	CTC	0	0
21	8	80	Glu	CTC	0	0
22	136	66	Gly	GCC	0	0
23	533	603	Gly	GCC	0	0
24	268	198	Gly	GCC	0	0
25	343	273	Gly	GCC	0	0
27	733	662	Gln	CTG	0	0
28	126	55	Lys	TTT	0	0
29	137	65	Lys	CTT	0	0
30	355	427	Lys	CTT	0	0
34	127	55	Phe	GAA	0	0
38	882	801	Ser	AGA	0	0
41	441	370	His	GTG	0	0
42	91	174	Tyr	GTA	128	138
43	406	479	Val	AAC	0	0
45	74	147	Ile	TAT	0	0
48	63	133	Gly	GCC	0	0
49	62	133	Pro	TGG	0	0
49	233	304	Pro	TGG	0	0
50	135	64	Asp	GTC	0	0
51	135	64	Asp	GTC	0	0
52	135	64	Asp	GTC	0	0
53	411	482	Asp	GTC	0	0
54	391	461	Asp	GTC	0	0
56	673	746	Thr	AGT	0	0
58	257	177	Leu	AAG	0	0
59	128	55	Asn	GTT	0	0
62	127	55	Phe	GAA	0	0
64	127	55	Phe	GAA	0	0
65	499	570	Pro	TGG	0	0
66	132	61	Pro	AGG	0	0
67	402	475	Asn	GTT	0	0
68	402	475	Asn	GTT	0	0
69	375	446	Pro	TGG	0	0
70	333	404	Cys	GCA	0	0

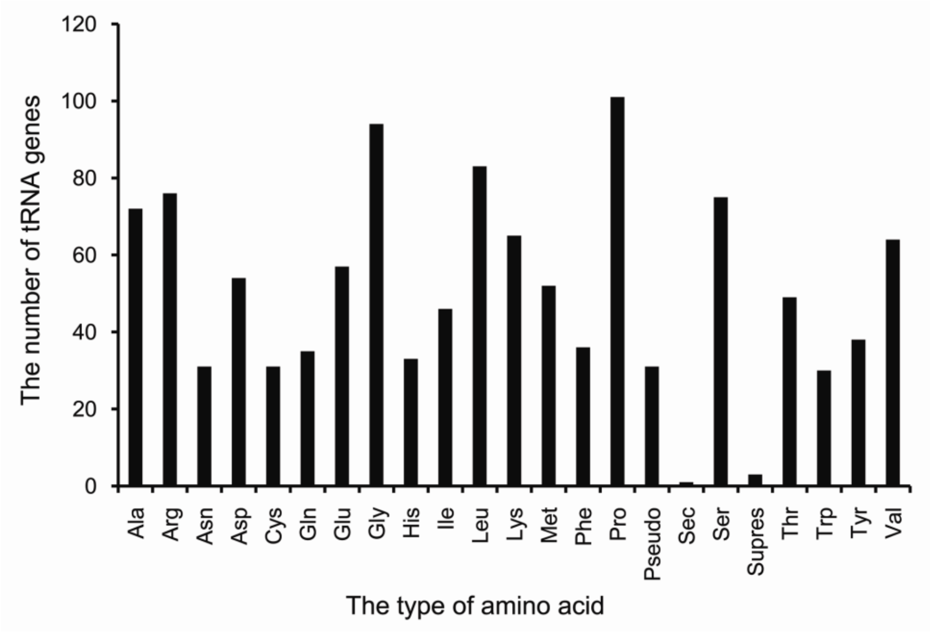


Fig. 1 The number of each of 23 classes of tRNA sequences carrying different amino acids

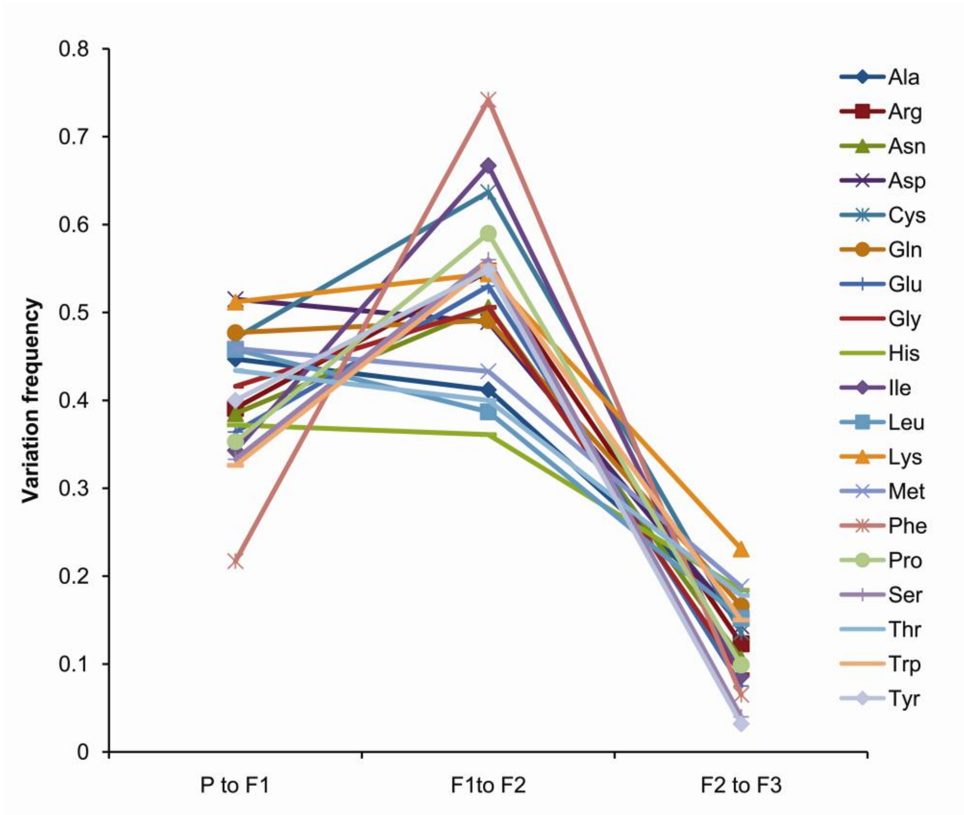
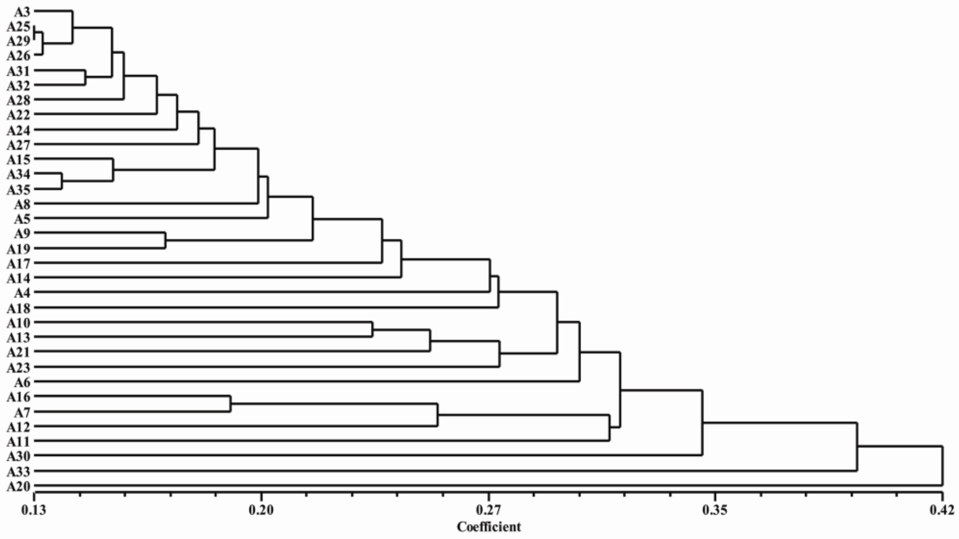
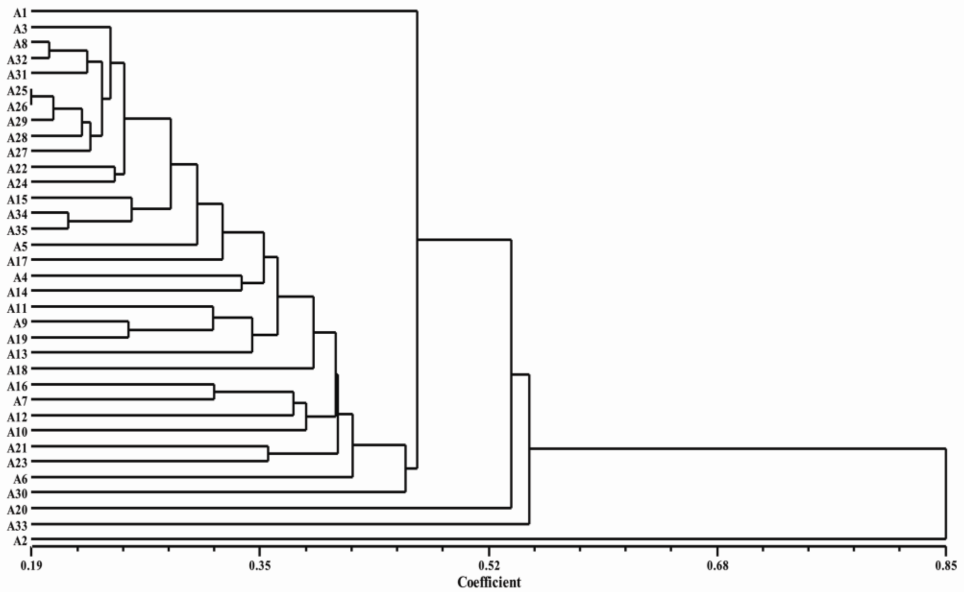


Fig. 2 Frequency of tRNA gene variants for tRNAs carrying each of the 20 common amino acids



a



b

Fig. 3 Dendrograms of cluster analysis based on inter-tRNA amplification polymorphism (ITAP) markers using UPGMA. Fig. 3a was made by 199 parental polymorphic loci (excluding loci showing novel or eliminated bands) in 33 self-pollinated lines derived from interspecific hybridization between *B. rapa* and *B. oleracea*. Fig. 3b was made by incorporating all 486 polymorphic loci in *B. rapa*, *B. oleracea* and their offspring (F_1, F_2, and F_3)

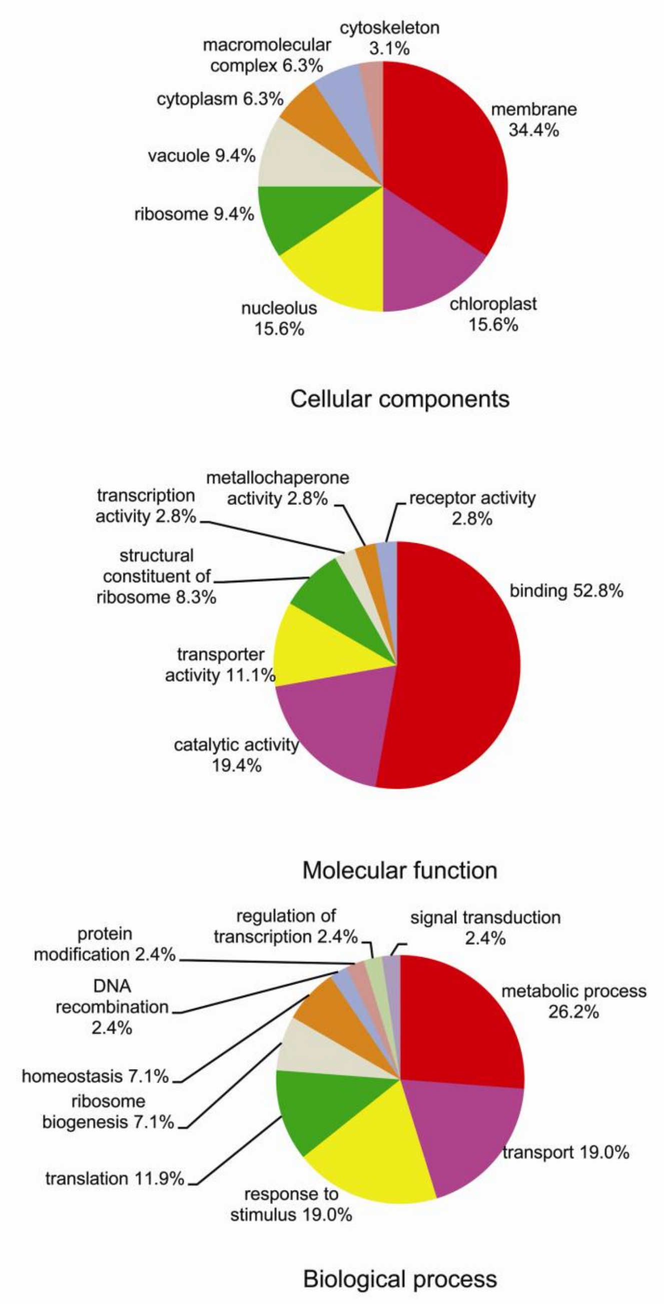


Fig. 4 Functional annotation of *Brassica* genes associated with tRNA-gene-associated sequences

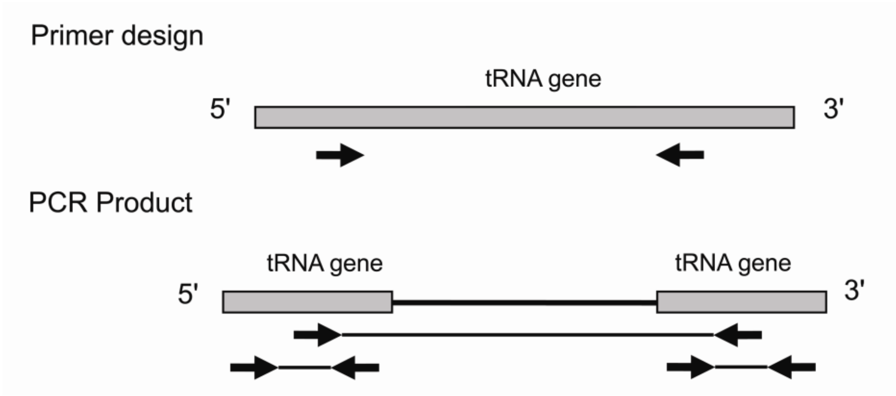


Fig. 5 The sketch of primer design and amplified products of an inter-tRNA amplification polymorphism (ITAP) marker. Gray areas denote a tRNA gene sequence. Black arrows indicate primer-binding regions and direction of amplification. Black lines indicate sequence between the two tRNA genes. The PCR amplification product is indicated by the black line between the two arrows, and is referred to as a tRNA-gene-associated sequence