Genetic and Epigenetic Changes in Somatic Hybrid Introgression Lines between Wheat and Tall Wheatgrass

Shuwei Liu*1, Fei Li*1,2, Lina Kong*, Yang Sun*, Lumin Qin*, Suiyun Chen*3, Haifeng Cui*, Yinghua Huang† and Guangmin Xia*4

*The Key Laboratory of Plant Cell Engineering and Germplasm Innovation, Ministry of Education, School of Life Sciences, Shandong University, Jinan 250100, P. R. China
†USDA/ARS Plant Science Research Laboratory, 1301 N. Western Rd, Stillwater, OK, 74075, USA

1These authors contributed equally to this work.
2Present Address: Cotton Research Center, Shandong Academy of Agricultural Sciences, Jinan 250100, P. R. China
3Present Address: Plant Science Institute, School of Life Sciences, Yunnan University, Kunming, 650091, P. R. China.
4Corresponding author: School of Life Sciences, Shandong University, Jinan 250100, P. R. China. E-mail: Xiagm@sdu.edu.cn

Copyright 2015.
Running title: Genetic and epigenetic changes in somatic hybrids

Key words: bread wheat; asymmetric somatic hybridization; introgression line; genomic shock, genetic and epigenetic alteration.

Corresponding author: Guangmin Xia, School of Life Sciences, Shandong University, Jinan 250100, P. R. China. E-mail: Xiagm@sdu.edu.cn
ABSTRACT

Broad phenotypic variations were induced in derivatives of an asymmetric somatic hybridization of bread wheat (*Triticum aestivum*) and tall wheatgrass (*Thinopyrum ponticum* Podp), however, how did these variations happened was unknown. We explored the nature of these variations by cytogenetic assays and DNA profiling techniques to characterize six genetically stable somatic introgression lines. Karyotyping results show the six lines similar to their wheat parent, but GISH analysis identified the presence of a number of short introgressed tall wheatgrass chromatin segments. DNA profiling revealed many genetic and epigenetic differences, including sequences deletions, altered regulation of gene expression, changed patterns of cytosine methylation and the reactivation of retrotransposons. Phenotypic variations appear to result from altered repetitive sequences combined with the epigenetic regulation of gene expression and/or retrotransposon transposition. The extent of genetic and epigenetic variation due to the maintenance of parent wheat cells in tissue culture was assessed and shown to be considerably lower than had been induced in the introgression lines. Asymmetric somatic hybridization provides appropriate material to explore the nature of the genetic and epigenetic variations induced by genomic shock.
INTRODUCTION

With the world's population continuing to increase, achieving a sustainable mode of food production represents an ever-growing challenge. Plant breeding has narrowed the genetic base of many crop species, but as yet has had little impact on the genetic diversity present in their wild relatives. In principle, this diversity can be introgressed into crops via sexual hybridization and subsequent backcrossing. However, in practice, wild crop manipulation has been severely restricted by difficulties in creating the initial sexual hybrid, and by sterility issues in the early backcross generations (Xia 2009). Asymmetric somatic hybridization is a viable alternative to introgression, especially where wide crosses are not feasible. It has been successfully exploited in bread wheat to transfer chromosomal segments from a number of related species (Xia et al. 2003; Xiang et al. 2003, 2004; Cheng et al. 2004; Zhou and Xia 2005; Xia 2009). More importantly, asymmetric somatic hybridization offers smaller alien chromatin introgression, thereby overcoming a significant problem in wheat sexual hybrids where the Ph1 gene prevents homoeologous recombination (Griffiths et al. 2006).

Newly synthesized allopolyploids have provided an opportunity to explore the nature of the genetic and epigenetic changes triggered by polyploidization (Song et al. 1995; Comai et al. 2000; Ozkan et al. 2001; Shaked et al. 2001; Madlung et al. 2002; Han et al. 2003; Ma et al. 2004; Wang et al. 2004a; Salmon et al. 2005; Tate et al. 2006; Bassene et al. 2010; Xu et al. 2014). Although, a few allopolyploids not accompanied with such changes (Liu et al. 2001). Experimental results indicate the majority of events are highly reproducible
(Bento et al. 2010). In particular, sequence deletion is common (Feldman et al. 1997; Liu et al. 1998a,b; Ozkan et al. 2001; Shaked et al. 2001; Kashkush et al. 2002; Ma et al. 2004; Ma and Gustafson 2006). Epigenetic modifications, such as changes in the pattern of cytosine methylation, have also been shown to induce changes in gene expression and activate transposon transcription (Comai et al. 2000; Shaked et al. 2001; Kashkush et al. 2002, 2003). However, these changes induced by ‘genomic shock’ during polyploid synthesis do not represent the changes in somatic introgressions. Allopolyploids represent a combination of nuclear genomes in a fixed cytoplasmic context, while somatic hybrids combine both the nuclear and cytoplasmic genomes within a single cell. The introgression of chromatin segments by asymmetric somatic hybridization likely occurs via non-homologous end-joining of fragmented genome pieces rather than by homologous recombination, which would show specific genetic and epigenetic changes in these materials. Moreover, the epigenetic state of somatic cells tend to be distinct from gametal cells given that the mutagenesis in gametal cells are more tightly controlled to ensure genetic fidelity (Bird 1997, 2002). Thus, the variations induced by ‘somatic genomic shock’ likely have unique characteristics compared with allopolyploids.

A number of hybrid progenies regenerated from asymmetric somatic hybrids [bread wheat cultivar Jinan 177 (JN177) and tall wheatgrass (Th. ponticum)] have proven phenotypically stable for a number of generations (Xia et al. 2003; Chen et al. 2004a; Wang et al. 2004b; Liu and Xia 2014). The hetero-cytoplasmic nature of these hybrid lines were confirmed, with the chloroplast genomic components are dominated by wheat,
while a few sequence of the chloroplast genome of tall wheatgrass were also detected in these lines (Chen et al. 2004b). DNA comparison of a well characterized set of glutenin proteins among parents and derivatives shows that all novel glutenin genes in hybrid progenies originated from alien genes of tall wheatgrass and allelic variation of parent wheat genes (Liu et al. 2007, 2009). Such de novo alleles do not arise simply as a result of UV-induced mutagenesis, as high frequency of the glutenin alleles were also found in symmetric somatic hybrids without UV pre-treatment (Gao et al. 2010). Moreover, somaclonal variation of parent wheat is too rare to account for the observed high frequency of novel glutenin alleles in the somatic hybrids (Feng et al. 2004). It’s more likely they derive from ‘genomic shock’ triggered by introgression of alien chromosome fragments during somatic hybridization. Therefore, the suggestion is that somatic hybridization provides a means of introgression distinct from sexual wide crossing and such introgression induces genomic variations. However, the details of the genomic variations and mechanism of transfer is unknown. Grass genomes are composed primarily of repetitive sequences, especially transposable elements (Feschotte et al. 2002). Whether or not somatic hybridization induces broad variations in repetitive sequences deserves further investigation. Moreover, increasing evidence shows epigenetic modifications, such as DNA methylation, play important roles in a wide range of biological processes including transposon inactivation and regulation of gene expression (Bird 2002; Zhang et al. 2006). Whether somatic hybridization induces epigenetic variations that affect gene expression and/or transposon activation needs further examination. Here, we used DNA profiling techniques to characterize genetic and
epigenetic alterations from ‘somatic genomic shock’ in six derivatives of bread wheat/tall wheatgrass somatic hybrids with different phenotypes.

**MATERIALS AND METHODS**

Plant Materials

Shanrong No. 1 (SR1), Shanrong No. 2 (SR2), Shanrong No. 3 (SR3), Shanrong No. 6 (SR6), Shanrong No. 10 (SR10) and Shanrong No. 12 (SR12) are six representative introgression lines derived from asymmetric somatic hybridization between bread wheat and tall wheatgrass (Xia et al. 2003; Chen et al. 2004a, b). By analogy with terminology applied to generations post-sexual crossing, mutagenesis and transformation, the regenerated plant post-fusion is referred to as R1; segregating progeny obtained by self-fertilization in successive generations are R2, R3, R4 etc (Figure 1). Only 1-2 seeds could produce from R1 and form a few R2 seedlings, but adequate R3 lines can be used for phenotypic and genomic investigation. We chose the R3 and R10 generations of the six introgression lines SR1, SR2, SR3, SR6, SR10 and SR12 for analysis. Four regenerated plants (R1-R4) were derived from protoplasts isolated from cultured embryonic calli of JN177 (R177) by the method previously described by Feng et al. (2004).

Mitotic and meiotic chromosome analyses

Root-tip meristems were squashed in 45% glacial acetic acid for GISH karyotyping according to previously described methods by Xiang et al. (2003). Tall wheatgrass genomic DNA was labeled as probes and genomic DNA from the bread wheat cv. Chinese Spring...
was used for blocking. A probe: block ratio of 1:150 was used. For meiotic chromosome analysis, anthers were fixed, squashed and stained according to previously described methods by Wang *et al.* (2004b).

Amplified fragment length polymorphism (AFLP) and methylation-sensitive amplification polymorphism (MSAP) fingerprinting AFLP and MSAP were performed according to Shaked *et al.* (2001), with minor modifications. A set of 44 AFLP and 13 MSAP primer combinations were used (Supplemental Table S1). Three independent technical replicates were performed for each of the three biological replicates for AFLP and MSAP reaction, and only intensely stained fragments larger than 100 bp were scored.

RT-PCR analysis RT-PCR was performed with cDNA templates prepared from 2 µg RNA extracted from seedling leaves and roots of cv. JN177 and the R₃ of six introgression lines and four R177 plants treated with DNase. The first cDNA strand was synthesized with M-MLV reverse transcriptase (Invitrogen), following the manufacturer's instructions. The PCR amplicons were separated using SSCP (single strand conformation polymorphism) electrophoresis through 6% non-denaturing polyacrylamide gels at 4°C, and visualized by silver staining. The set of 80 primer pairs (Supplemental Table S2) was developed by Bottley *et al.* (2006). Three independent technical replicates were performed for each of the three biological replicates for SSCP analysis.
Inter-retrotransposon amplified polymorphism (IRAP) and retrotransposon microsatellite amplified polymorphism (REMAP) genotyping

IRAP and REMAP PCR were performed in 20 µl reactions, as described by Kalendar et al. (1999). The ten IRAP and 12 REMAP primer pairs were developed by Bento et al. (2008) (Supplemental Table S3). The amplicons were electrophoresed through 2% agarose gels, and visualized by ethidium bromide staining. Three independent technical replicates were performed for each primer combination.

Cloning and sequencing of gel fragments

Selected gel fragments were excised from gels and re-amplified using the primers from which they were generated. The amplicons were separated by electrophoresis through 1% agarose gels, gel-purified, and ligated into the pEASY-T1 plasmid (Transgene), which was transformed into E. coli strain DH105 for sequencing based on universal M13 primers. The sequences were BLAST against the NCBI and/or the TREP (the Triticeae Repeat Sequence Database) database.

Southern Blot

Fragments with polymorphisms from parent wheat cv. JN177 and the introgression lines in AFLP, SSCP, REMAP and IRAP analysis were re-amplified from the corresponding reconstituted plasmid and labeled with DIG High Prime DNA Labeling and Detection Starter Kit I (Roche) and used as hybridization probes for DNA gel blot analysis. Genomic
DNA from the introgression lines, cv. JN177 and R177 plants was digested with restriction enzyme EcoR I and blotted onto nylon membranes through vacuum transfer. Pre-hybridization, hybridization and color detection were performed according to protocols provided by the DIG High Prime DNA Labeling and Detection Starter Kit I (Roche).

Bisulfite genomic sequencing

For bisulfite genomic sequencing, genomic DNA was processed with an EpiTect Bisulfite kit (Qiagen). The amplification primer pairs for each of MSAP-isolated fragments (MIFs) were designed using MethPrimer software (Li and Dahiya 2002). Primer sequences are given in Supplementary Table S4. The PCR products were ligated with the pEASY-T Vector (TransGen), and at least 30 clones per insert processed for sequencing. The ratio of C methylation at each cytosine site was calculated and transformed into a percentage using CyMATE software (www.gmi.oeaw.ac.at/research-groups/cymate/cymate/).

RESULTS

Generation of the introgression lines

Our asymmetric somatic hybridization between bread wheat cv. JN177 and tall wheatgrass produced hundreds of derivatives (Figure 1) classified into five groups according to their phenotypes (Xia et al. 2003; Zhang et al. 2005; Xia 2009). We chose the six introgression lines SR1, SR2, SR3, SR6, SR10 and SR12 to represent the five phenotypic groups (SR2 and SR3 belong to the same group), which has proven to be phenotypically stable over many generations. The first group showed high grain yield which represented by SR1, the
grain yield of SR1 was 35% higher compared with JN177 under non-nutrient limiting
conditions. These data correlate with the fact that SR1 is registered in China as a high
yielding cultivar. The second group showed higher abiotic stress tolerance represented by
SR2 and SR3; SR2 showed high levels of drought tolerance, while SR3 was noticeably
more salinity tolerant than cv. JN177. In addition, SR3 also showed higher grain yield
compared with JN177 under non-nutrient limiting conditions, it is designated as a salinity
tolerant cultivar in Shandong Province. The third group derivatives showed higher disease
resistance represented by SR6; this derivative was highly resistant to powdery mildew and
stripe rust. The fourth group derivatives were all dwarfs which represented by SR10, the
stems of SR10 was 48.20±1.31 cm, but it possessed strong tillering ability and end-use
quality. The last group of derivatives showed large ears and grains. The 1000 grain weight
of one member, SR12, was nearly 60 g. All measured agricultural traits of the six
introgression lines are listed in Table 1.

Chromosomal constitution of the six introgression lines

We investigated the introgression of Th. ponticum chromosomal segments into wheat
chromosomes in the six lines. Most of the root tip mitotic cells in the R₃ generation
contained 42 chromosomes with a stable somatic chromosome number from generation R₃
to R₁₀ (data not shown). Karyotypes were nearly identical from cv. JN177, with some
minor variations in arm length ratio and chromosome length (data not shown). According
to genomic in situ hybridization (GISH) analysis, we found several small tall wheatgrass
chromosomal segments present in each line (Figure 2). We also found over 90% ring
bivalent and over 95% total bivalent frequencies in the six introgression lines at meiosis (Supplemental Table S5). This suggests that introgressed tall wheatgrass chromosomal segments had little effect on chromosome pairing behavior, likely due to the small segment size. Given that the karyotypes and DNA fingerprinting profiles were stable for each introgression line between the R₃ and R₁₀ generations (data not shown), all of the following analysis was restricted to the DNA of R₃ plants.

Frequency of sequence variation as estimated from AFLP profiles

In this study, only a few tall wheatgrass segments were introgressed into the wheat genome. Our expectation was that the AFLP profiles of the introgression lines would be similar to that of cv. JN177, with the addition of some tall wheatgrass derived fragments. We focused on fragments present in JN177 but lacking in one, or more, of the introgression lines, and the presence of fragments in the introgression line(s) but not in JN177. Only fragments larger than 100 bp were scored.

AFLP profiling revealed 904 cv. JN177 and 777 tall wheatgrass fragments; 388 were shared by both templates (Figure 3, A and B). Given that AFLP assay markers are dominant, the level (%) of cv. JN177 fragment loss could only be related to the 516 (904-388) JN177 fragments that did not co-migrate with a tall wheatgrass fragment. The number of fragments lost from the profile of the six introgression lines ranged from 22 to 28, with an overall frequency of 4.7% (Table 2). In addition, the six introgression lines displayed 12 to 18 new fragments not present in the cv. JN177 profile, giving an overall frequency of 2.9% (Table 2). AFLP assay results for cv. JN177 and four R177 lines revealed an average of
only 3.5 deleted and 2.0 new bands compared with JN177 (Table 2).

The AFLP-observed fragment changes may have resulted from several mechanisms including sequence change at restriction sites, methylation alteration, or sequence elimination induced by somatic hybridization. To distinguish between these possibilities, we isolated four bands deleted in some of the introgression lines from acrylamide AFLP gels and used them as probes for Southern blot analysis. The results confirm deletion of one band, with a single copy fragment (Figure 3C). The remaining three fragments, which had low copy DNA, showed no obvious differences between JN177 and the introgression lines (Figure 3D-F). Thus, variations found in AFLP analysis are likely mainly due to sequence changes at restriction sites or from methylation alteration, rather than sequence elimination.

Somatic hybridization induced changes in methylation pattern

The isoschizomers *Msp* I and *Hpa* II share the same recognition sequence, but differ in their sensitivity to C methylation. *Hpa* II activity is suppressed when the internal cytosine at the CCGG recognition site is methylated. Therefore, C methylation differences can be detected by MSAP analysis, comparing profiles generated by *Eco*RI-*Msp* I and *Eco*RI-*Hpa* II digestion (Shaked et al., 2001). The standards applied to scoring MSAP profiles and AFLP analyses are similar. The 13 MSAP primer combinations amplified 347 fragments, of which 210 were present in cv. JN177 but not in the tall wheatgrass parent (Figure 4). Of the 210, the number of hypermethylated loci in different introgression lines ranged from 16 to 27, giving an overall frequency of 10.5% (Table 3). Among the loci methylated in cv. JN177, 35 were hypomethylated in SR1, 33 in SR3, 29 in SR6, 22 in...
SR10 and 23 in both SR2 and SR12, giving an overall frequency of 13.1% (Table 3). The C methylation in cv. JN177 and four R177 lines were also detected by MSAP analysis, the results indicated that only 5.5 loci hypermethylated in the regenerated plant and 8.0 loci hypomethylated (Table 3). The sequencing of five of the fragments hypomethylated in the introgression lines showed that three were high copy number retrotransposons (TREP accession numbers TREP255, TREP1418, TREP3251), while the other two showed no significant similarity to any known sequence. None of the five sequenced fragments that were hypermethylated in the introgression lines showed any significant similarity to DNA sequences of publicly available databases.

To further investigate the C methylation changes induced by somatic hybridization in the introgression lines, C methylation status in several variant MSAP bands was identified by bisulfite sequencing. We found the methylation rates at CG, CHG and CHH sites of MSAP-isolated fragment 1 (MIF1) were all elevated in SR2 and SR12 (Figure 5A). The level of C methylation at CG sites of MIF2 was elevated in SR1, SR2, SR3 and SR10 but reduced in SR6 and SR12, although the CHG and CHH sites were also hypermethylated in SR3 (Figure 5B). Moreover, the level of C methylation at CG and CHG sites of MIF3 was elevated in SR1, and those at CG, CHG and CHH sites were all elevated in R177 lines (Figure 5C). The level of C methylation at CG sites of MIF4 was elevated in SR1, SR2, SR6 and SR10 (Figure 5D).

Gene expression alterations as estimated by SSCP

To survey the homoeologous gene silencing or activation induced by somatic hybridization,
we used SSCP to compare the amplification profiles of the six introgression lines with those of seedling shoot and root RNA extracted from cv. JN177 and R177. Expression variation was noted for 11 of the 80 targeted sequences. Eight of these involved the loss of a specific homoeologous product in at least one introgression line (Figure 6, A-D). In two cases, a specific mRNA was present in some introgression lines but not in cv. JN177 (Figure 6C). In three cases, the introgression line profiles included a variant of one of the cv. JN177 fragments (Figure 6D). To test whether genetic variations were responsible for the changes of the expression, genomic DNA was also included as a template. In four of the eight genes involved in homeologous gene silencing, the corresponding fragment was also absent from the genomic DNA profile (Figure 6, A and D). Southern blot analysis based on genomic DNA showed that one of the homoeologues of BE490384, BE494911 and BE606965 had been deleted in the introgression lines (Figure 6, E-G). But, this was not the case for BE399113 (Figure 6H). No variation was found between cv. JN177 and the four R177 lines in all 80 targeted sequences.

Genomic rearrangements in retrotransposon-adjacent regions

IRAP and REMAP assays were developed to detect variation in sequence flanking retrotransposons (Kalendar et al. 1999). We used these two techniques to identify genome rearrangements from retrotransposition or inter- and intra-element recombination. Here, 219 wheat specific loci were detected. Five were lost in SR1, SR2 and SR3; 8 were lost in SR6 and SR12; 11 lost in SR10; with an overall mean frequency of lost fragments across the six introgression lines at 3.2% (Figure 7, A-D, Table 4). An average of 1.5 fragments was lost in
R177 plants when compared with cv. JN177 (Table 4). Seven novel fragments were found in SR1, SR2 and SR12 that were not amplified in both cv. JN177 and tall wheatgrass. Additionally, 6, 5 and 12 novel fragments also appeared in SR3, SR6 and SR10, respectively. These values give an overall new-band frequency of 3.3% (Figure 7, A-D, Table 4). Only 1.0 novel fragment appeared in R177 plants when compared with cv. JN177 (Table 4).

Southern blot analysis based on genomic DNA using polymorphic fragments isolated from IRAP and REMAP gels as probes show the lost fragments were deleted in some of the introgression lines (Figure 7, E and F). We sequenced a sample of these fragments to identify the genes associated with these genomic rearrangements. A fragment present in cv. JN177, but absent in some of the introgression lines, resembled a Ty3-gypsy subclass retrotransposon (ABA97230.1). A second polymorphic fragment (present in some of the introgression lines, but not in cv. JN177) appeared to be highly homologous to a wheat pore-forming toxin-like protein Hfr-2 (AAW48295.1) involved in plant defense (Puthoff et al. 2005).

**DISCUSSION**

Merging divergent genomes into a single nucleus can trigger a “highly-programmed sequence of events within the cell that serves to cushion the effect of [genomic] shock” (McClintock 1984). However, the mechanisms of genomic shock are not well understood. The introgression of alien chromosome fragments can also lead to severe effects in the accepting genome, which trigger genetic and genomic changes. Exploring genetic and genomic changes induced by ‘somatic genomic shock’ in the somatic introgression lines
will be helpful for understanding the nature of ‘genomic shock’. Here, we demonstrated that the somatic hybridization process induced widespread genetic and epigenetic changes including sequence absence, modified regulation of gene expression, alteration of cytosine methylation patterns and the activation of quiescent retrotransposons. Similar genetic and/or epigenetic alterations occurred following the formation of de novo polyploids involving Arabidopsis spp. (Comai et al. 2000; Madlung et al. 2002; Wang et al. 2004a), Brassica spp. (Song et al. 1995), Spartina spp. (Salmon et al. 2005), Tragopogon spp. (Tate et al. 2006) and Triticeae spp. (Ozkan et al. 2001; Shaked et al. 2001; Han et al. 2003; Ma et al. 2004). However, the ‘somatic genomic shock’ induced by somatic hybridization might be stronger than those in allopolyploids given the former resulted in more genetic and epigenetic instabilities. In somatic hybrids, the requirement for a period of in vitro culture adds an additional potential source of variation. However, the contrast made between the DNA extracted from cv. JN177 and R177 plants implies that somaclonal variation was responsible for only a small proportion of the overall variation induced in the introgression lines. Thus, genome introgression was likely responsible for most genetic and epigenetic variations. Although we cannot exclude the effect of organellar genome introgression, considering the limited amount of alien organellar DNA in the hybrids (Chen et al. 2004b), the introgression of nuclear DNA was likely responsible for most genetic and epigenetic variations.

Sequence absence

One of the most common responses to polyploidization is the rapid elimination of DNA
sequences (Feldman et al. 1997; Liu et al. 1998a, b; Ozkan et al. 2001; Shaked et al. 2001; Kashkush et al. 2002; Ma et al. 2004; Ma and Gustafson 2006). The frequency of elimination is particularly marked in synthetic tetraploid wheat and triticale (Shaked et al. 2001; Ma et al. 2004; Ma and Gustafson 2006). In our study, sequence loss was also observed following somatic hybridization (Figure 3), but at a rather lower rate. Only 4.7% of the cv. JN177 AFLP fragments were not recovered in the somatic introgression lines, while 0.7% were concluded to have been induced by somaclonal variation (Table 2). Moreover, southern blot analysis of the lost AFLP fragments indicated that only a few were eliminated from the genome of the introgression lines. Most of the band absence may be due to sequence changes of restriction sites or from methylation alterations. Sequence elimination during allopolyploidization provides the physical basis for the diploid-like meiotic behavior of newly formed allopolyploids (Ozkan et al. 2001). Unlike allopolyploidization, where the whole parental genomes are combined, the size of the individual donor chromosome fragments introgressed into cv. JN177 via asymmetric somatic hybridization is too small to affect meiotic chromosome pairing (Figure 2). Thus, meiotic pressure is likely to be lower compared with a polyploid. This may partly explain why sequence elimination is not a predominant effect in somatic hybridization, unlike in polyploidization.

Epigenetic alterations

Patterns of gene expression are affected by allopolyploidization, particularly in wheat, where 1-5% of the donor genes have altered expression in a synthetic allotetraploid, largely
via epigenetic changes (Kashkush et al. 2002). Changes in gene expression are mediated at
either the transcriptional and/or the post-transcriptional levels. Both transcriptional and
post-transcriptional gene silencing are associated with alterations to DNA methylation
(Paszkowski and Whitham 2001). In newly synthesized allohexaploid wheat, approximately 13% of the loci experience an alteration in cytosine methylation (Shaked et al. 2001). A commonality between the effect of somatic hybridization and allopolyploidization appears to be that both result in changes to cytosine methylation, which results in the perturbation of gene expression (Comai et al. 2000; Shaked et al. 2001; Kashkush et al. 2002; Han et al. 2003). However, the epigenetic variations induced by somatic hybridization are stronger than those by allopolyploidization. Based on the analysis of our sample of 80 cDNAs, altered expression was induced by the somatic hybridization procedure in 10% of the introgression line transcripts (Figure 6). The notable frequency of cytosine methylation changes observed in the asymmetric somatic hybrids (23.6%) is consistent with large-scale epigenetic regulation (Table 3). Thus, we concluded that somatic hybridization induced a broad spectrum of C methylation changes that perturbed gene expression to a larger extent than allopolyploidization.

Activation and repression of transposon activity

Genomic shock activates quiescent transposons and causes genome restructuring (McClintock 1984). In fact, transcriptional activation of transposable elements has been shown in some newly synthesized allopolyploids (Kashkush et al. 2002, 2003). IRAP and REMAP both represent multilocus PCR-based methods designed to detect retrotransposon
related genome rearrangements, and have been used to uncover sequence rearrangements associated with retrotransposon-rich regions in triticale (Bento et al. 2008, 2010). Here, 6.5% of the IRAP and REMAP fragments were polymorphic between cv. JN177 and the introgression lines (Figure 7, Table 4). The introgression events were associated with novel sequence polymorphisms at the junctions between the donor and the recipient DNA. However, given the overall number of wheatgrass segments present in the introgression lines, frequency of polymorphisms revealed by IRAP and REMAP was too high to merely reflect introgression (Figure 2). Instead, the polymorphisms detected by IRAP and REMAP likely arose from de novo transposition events. These have been detected in the vicinity of genes encoding certain endosperm storage proteins by Liu et al. (2009).

Evidence suggests that C methylation has evolved as a means of repressing transposon activity. Clearly, highly repetitive DNA, much of which is contributed by retrotransposons in the large genome grass species, in plant genomes is heavily methylated (Bennetzen et al. 1994). The activity of retrotransposons needs to be tightly controlled to ensure viability and survival of the host, and a major control step involves the inhibition of transcription. Reactivation of transposons has been observed either as a response to stress or to genomic shock (Bennetzen 2000; Liu and Wendel 2000; Kashkush et al. 2002, 2003; Shan et al. 2005, 2009; Bento et al. 2008, 2010). Our MSAP analysis of the introgression lines indicates that alteration in C methylation involves repetitive sequences. In particular, the hypomethylation of retrotransposons revealed is suggestive of retrotransposon activation in the introgression lines.

Although considerable transposition occurred during the somatic hybridization event,
the introgression lines were highly stable from R₃ to R₁₀ generation. The implied repression of retrotransposon activity closely resembles the behavior of rice introgression lines (Liu and Wendel 2000), although the underlying mechanism is unknown. Wolffe and Matzke (1999) have proposed the involvement of homology-dependent gene silencing (HDGS), an epigenetic phenomenon whereby one gene is suppressed by the action of a homologous gene (Zaratiegui et al. 2007). When transposon activity is activated by hypomethylation, HDGS is triggered and similar sequences become targets for suppression. How important the rapid repression of retrotransposon activity is for the stabilization of the introgression lines following somatic hybridization clearly needs further investigation.

Genetic alterations driven by repetitive sequences

Repetitive sequences are a major source of genome instability as they are susceptible to expansion and contraction (Heidenfelder and Topal 2003). As the most abundant repetitive sequences in wheat genome, the retrotransposon related sequences showed high frequency of alteration (6.5%) in the introgression lines (Table 4). Similarly, high frequency sequence variation through indels of repetitive motifs (14/37, 37.8%) was also observed in the glutenin gene family (Feng et al. 2004; Liu et al. 2007, 2009; Gao et al. 2010). This supports the notion that repetitive sequences are a driving force behind the de novo genetic variation generated in the introgression lines. Double strand breaks (DSBs) appear to be common in regions enriched for repetitive sequences, and these “fragile sites” represent hotspots for recombination (Wicker et al. 2010). The suggestion is therefore that repetitive sequences drive the formation of DSBs or other genetic alterations that occur as a result of
genomic shock and subsequent introgression.

CONCLUSIONS

We generated a set of introgression lines displaying a variety of genetic and epigenetic changes relative to their parents using somatic hybridization between wheat and tall wheatgrass. The somatic hybridization process mimics many of the genetic alterations induced by polyploidization or sexual wide hybridization, in a remarkably short time frame and with stronger extent. Therefore, we suggest the somatic hybridization approach not only provides an effective and time efficient means of achieving introgression into a crop species from its wild relatives, but also provides a means to explore the nature of the genetic and epigenetic events induced by ‘somatic genomic shock’. An understanding of the underlying mechanisms may shed light on the variation released in sexual wild hybrids, which has been broadly described as being due to genomic shock.

SUPPORTING INFORMATION

Supporting information is available online at http://www.genetics.org/

ACKNOWLEDGMENTS

This work was supported by the funds of the Natural Science Foundation of China (No. 30871320; 31000568), Major Program of the Natural Science Foundation of China (No. 31030053), Shandong Province Program (No. Q2006D02). We thank Dr. Austin Cape for careful reading and feedback.
LITERATURE CITED


Griffiths, S., R. Sharp, T. N. Foote, I. Bertin, M. Wanous et al., 2006 Molecular characterization of Ph1 as a major chromosome pairing locus in polyploid wheat. Nature


Liu, B., J. M. Vega, G. Segal, S. Abbo, M. Rodova et al., 1998b Rapid genomic changes in newly synthesized amphiploids of *Triticum* and *Aegilops*. I. Changes in low-copy


Table 1. Phenotypic traits associated with the six introgression lines and JN177.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Growth</th>
<th>Plant height (cm)</th>
<th>Spike length (cm)</th>
<th>Number of spikelet/spike</th>
<th>Number of grain/spike</th>
<th>Effective tillers</th>
<th>1000 grain weight (g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>JN177</td>
<td>Arrest</td>
<td>66.90±3.07</td>
<td>9.30±0.63</td>
<td>19.7±0.95</td>
<td>42.8</td>
<td>4.7±0.95</td>
<td>37.88</td>
</tr>
<tr>
<td>SR1</td>
<td>Creep</td>
<td>75.60±2.27</td>
<td>12.95±0.69</td>
<td>22.9±0.74</td>
<td>77.1</td>
<td>8.4±3.72</td>
<td>46.96</td>
</tr>
<tr>
<td>SR2</td>
<td>Creep</td>
<td>74.30±3.74</td>
<td>11.40±0.94</td>
<td>21.8±1.14</td>
<td>71.0</td>
<td>6.4±1.71</td>
<td>43.17</td>
</tr>
<tr>
<td>SR3</td>
<td>Creep</td>
<td>75.70±3.05</td>
<td>14.00±0.62</td>
<td>24.3±1.33</td>
<td>66.7</td>
<td>5.5±1.18</td>
<td>41.99</td>
</tr>
<tr>
<td>SR6</td>
<td>Arrest</td>
<td>75.20±2.90</td>
<td>8.00±0.33</td>
<td>18.5±0.97</td>
<td>50.8</td>
<td>6.8±1.62</td>
<td>33.03</td>
</tr>
<tr>
<td>SR10</td>
<td>Creep</td>
<td>48.20±1.31</td>
<td>9.30±0.48</td>
<td>16.6±0.97</td>
<td>41.5</td>
<td>13.5±3.54</td>
<td>38.65</td>
</tr>
<tr>
<td>SR12</td>
<td>Creep</td>
<td>74.30±3.27</td>
<td>12.90±0.39</td>
<td>19.8±0.79</td>
<td>54.8</td>
<td>4.9±1.29</td>
<td>57.77</td>
</tr>
</tbody>
</table>
Table 2. AFLP profiles of cv. JN177, the six introgression lines and regenerated JN177 plants (R177).

<table>
<thead>
<tr>
<th>Introgression lines</th>
<th>SR1</th>
<th>SR2</th>
<th>SR3</th>
<th>SR6</th>
<th>SR10</th>
<th>SR12</th>
<th>Average</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deletions in introgression lines</td>
<td>23</td>
<td>22</td>
<td>26</td>
<td>25</td>
<td>28</td>
<td>22</td>
<td>24.3</td>
<td>4.7%</td>
</tr>
<tr>
<td>New bands in introgression lines</td>
<td>16</td>
<td>15</td>
<td>18</td>
<td>16</td>
<td>12</td>
<td>13</td>
<td>15.0</td>
<td>2.9%</td>
</tr>
<tr>
<td>Regenerated plants</td>
<td>R1</td>
<td>R2</td>
<td>R3</td>
<td>R4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deletions in R177</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td>3.5</td>
<td>0.7%</td>
</tr>
<tr>
<td>New bands in R177</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td>2.0</td>
<td>0.4%</td>
</tr>
</tbody>
</table>
Table 3. Variation in C methylation in the introgression lines and regenerated JN177 plants (R177).

<table>
<thead>
<tr>
<th>Introgression lines</th>
<th>SR1</th>
<th>SR2</th>
<th>SR3</th>
<th>SR6</th>
<th>SR10</th>
<th>SR12</th>
<th>Average</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypermethylation in introgression lines</td>
<td>27</td>
<td>23</td>
<td>24</td>
<td>23</td>
<td>16</td>
<td>19</td>
<td>22.0</td>
<td>10.5%</td>
</tr>
<tr>
<td>Hypomethylation in introgression lines</td>
<td>35</td>
<td>23</td>
<td>33</td>
<td>29</td>
<td>22</td>
<td>23</td>
<td>27.5</td>
<td>13.1%</td>
</tr>
<tr>
<td>Regenerated plants</td>
<td>R1</td>
<td>R2</td>
<td>R3</td>
<td>R4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypermethylation in R177</td>
<td>5</td>
<td>9</td>
<td>5</td>
<td>3</td>
<td></td>
<td></td>
<td>5.5</td>
<td>2.6%</td>
</tr>
<tr>
<td>Hypomethylation in R177</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>11</td>
<td></td>
<td></td>
<td>8.0</td>
<td>3.8%</td>
</tr>
</tbody>
</table>
Table 4. Genomic rearrangements in retrotransposon related regions in the introgression lines and regenerated JN177 plants (R177).

<table>
<thead>
<tr>
<th>Introgession lines</th>
<th>SR1</th>
<th>SR2</th>
<th>SR3</th>
<th>SR6</th>
<th>SR10</th>
<th>SR12</th>
<th>Average</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bands lost in</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>8</td>
<td>11</td>
<td>8</td>
<td>7.0</td>
<td>3.2%</td>
</tr>
<tr>
<td>introgression lines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New bands in</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>12</td>
<td>7</td>
<td>7.3</td>
<td>3.3%</td>
</tr>
<tr>
<td>introgression lines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regenerated plants</td>
<td>R1</td>
<td>R2</td>
<td>R3</td>
<td>R4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bands lost in R177</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>1.5</td>
<td>0.7%</td>
</tr>
<tr>
<td>New bands in R177</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td>1.0</td>
<td>0.5%</td>
</tr>
</tbody>
</table>
Figure 1. Generation of somatic hybridization introgression lines. Circles represent nuclei; lines are chromosomes and chromosome segments.
Figure 2. Cytogenetic analysis of the six introgression lines at R₃. GISH karyotypes of (A) tall wheatgrass and (B) bread wheat cv. JN177, with tall wheatgrass genomic DNA as probe; SR1: (C) GISH karyotype, (D) chromosome configuration at meiotic metaphase I; SR2: (E) GISH karyotype, (F) chromosome configuration at meiotic metaphase I; SR3: (G) GISH karyotype, (H) chromosome configuration at meiotic metaphase I; SR6: (I) GISH karyotype, (J) chromosome configuration at meiotic metaphase I; SR10: (K) GISH karyotype, (L) chromosome configuration at meiotic metaphase I; SR12: (M) GISH karyotype, (N) chromosome configuration at meiotic metaphase I.
Figure 3. AFLP profile and southern blot analysis of deleted fragments. (A) AFLP profile based on primer combination $E_{AGC}+M_{CTT}$, (B) primer combination $E_{AAT}+M_{CTT}$. Fragments absent from the introgression lines or regenerated plant from JN177 calli (R) are indicated by triangles, novel fragments indicated by tailed arrows. Southern blots verified the deletion of genomic DNA in SR6 with probes derived from primer combination $E_{ATC}+M_{CTT}$ (C), but not with probes derived from primer combination $E_{AGG}+M_{CTT}$ (D), $E_{AGC}+M_{CTT}$ (E) and $E_{AAA}+M_{CTT}$ (F). 1: SR1; 2: SR2; 3: SR3; 6: SR6; 10: SR10; 12: SR12; 177: parent wheat cv. JN177; R: regenerated JN177 plant R1; Th: tall wheatgrass *Th. ponticum*; L: the 100 bp DNA ladder marker.
Figure 4. MSAP profiling based on primer combinations E\textsubscript{ATC}+HM\textsubscript{TCAA}. M: DNA restricted by EcoR I+Msp I, H: EcoR I+Hpa II. Tailed arrows indicate hypermethylated fragments, triangles indicate hypomethylated fragments. 1: SR1; 2: SR2; 3: SR3; 6: SR6; 10: SR10; 12: SR12; 177: parent wheat cv. JN177; R: regenerated JN177 plant R1; Th: tall wheatgrass Th. ponticum; L: the 100 bp DNA ladder marker.
Figure 5. C methylation status in the MSAP-isolated fragments (MIFs) as identified by bisulfite sequencing. (A) The level of C methylation at CG, CHG and CHH sites of MIF1 was all elevated in SR2 and SR12; (B) The level of C methylation at CG sites of MIF2 was elevated in SR3 and SR10 but reduced in SR6 and SR12 compared to JN177, those at CHG and CHH sites were also elevated in SR3; (C) The level of C methylation at CG and CHG sites of MIF3 was elevated in SR1 and those at CG, CHG and CHH sites were all elevated in regenerated JN177 plant R1; (D) The level of C methylation at CG sites of MIF4 was elevated in SR1, SR2, SR6 and SR10 compared to JN177. 177: parent wheat cv. JN177; R: regenerated JN177 plant R1.
Figure 6. RT-PCR and genomic DNA characterization of differentially expressed genes. (A) One homoeoallele of BE494911 is silenced in SR3 and SR10 while another homoeoallele silenced in SR6 and root of SR12; the silencing of BE494911 in SR3, SR6 and SR10 is caused by a deletion in the coding region; (B) BE202265 silenced in shoot and root of SR1, SR2, SR3 and shoot of SR12; (C) One homoeoallele of BE606719 is activated in root of SR2; while another homoeoallele silenced in SR10 root; (D) For the gene, BE606965, one homoeologous copy differs from that of cv. JN177 in SR1, SR2, SR6 and SR10; while another homoeoallele is deleted in the coding region of SR2, SR6 and SR10. Southern
blots demonstrate the deletion of genomic DNA in (E) BE490384, (F) BE494911 and (G) BE606965, but not in (H) BE399113. Arrows indicate the genomic fragments deleted in the introgression lines. 1: SR1; 2: SR2; 3: SR3; 6: SR6; 10: SR10; 12: SR12; 177: parent wheat cv. JN177; R: regenerated JN177 plant R1; Th: tall wheatgrass *Th. ponticum*.
Figure 7. REMAP and IRAP analysis, based on primer combinations (A) (GA)$_9$C + sukkula, (B) (GA)$_9$C + sabrina, (C) nikita + sukkula, (D) nikita + sabrina. Southern blots demonstrate the retrotransposon-related deletions in the introgression lines with probes derived from (CA)$_9$G + sabrina (E) and nikita + sukkula (F). Triangles indicate retrotransposon-related deletions in the introgression lines induced by asymmetric somatic hybridization, while tailed arrows indicate retrotransposon-related insertions in the introgression lines induced by asymmetric somatic hybridization. 1: SR1; 2: SR2; 3: SR3; 6: SR6; 10: SR10; 12: SR12; 177: parent wheat cv. JN177; R: regenerated JN177 plant R1; Th: tall wheatgrass *Th. ponticum*. 