ORIGINAL PAPER

A radiation hybrid map of chromosome 1D reveals synteny conservation at a wheat speciation locus

**Monika K. Michalak de Jimenez · Filippo M. Bassi · Farhad Ghavami ·
Kristin Simons · Rissa Dizon · Raed I. Seetan · Loai M. Alnemer · Anne M. Denton ·
Münevver Doğramacı · Hana Šimková · Jaroslav Doležel · Kiran Seth ·
Ming-Cheng Luo · Jan Dvorak · Yong Qiang Gu · Shahryar F. Kianian**

Abstract The *species cytoplasm specific* (*scs*) genes affect nuclear-cytoplasmic interactions in interspecific hybrids. A radiation hybrid (RH) mapping population of 188 individuals was employed to refine the location of the *scs*[ae] locus on *Triticum aestivum* chromosome 1D. "Wheat Zapper," a

Monika K. Michalak de Jimènez and Filippo M. Bassi contributed equally to this paper.

M. K. Michalak de Jimenez · F. M. Bassi · F. Ghavami ·
K. Simons · R. Dizon · K. Seth · S. F. Kianian (✉)
Department of Plant Sciences, North Dakota State University,
Fargo, ND 58108-6050, USA
e-mail: S.Kianian@ndsu.edu

R. I. Seetan · A. M. Denton
Department of Computer Science, North Dakota State University,
Fargo, ND 58108-6050, USA

L. M. Alnemer
Computer Information System Department, University of Jordan,
Amman 11942, Jordan

M. Doğramacı
USDA-ARS, BRL, Fargo, ND 58102-2765, USA

H. Šimková · J. Doležel
Center of the Region Hana for Biotechnological and Agricultural
Research, Institute of Experimental Botany, Olomouc CZ-77200,
Czech Republic

M.-C. Luo · J. Dvorak
Department of Plant Sciences, University of California, Davis,
CA 95616, USA

Y. Q. Gu
USDA - Agricultural Research Service, Western Regional
Research Center, Albany,
CA 94710, USA

comparative genomics tool, was used to predict synteny between wheat chromosome 1D, *Oryza sativa*, *Brachypodium distachyon*, and *Sorghum bicolor*. A total of 57 markers were developed based on synteny or literature and genotyped to produce a RH map spanning 205.2 cR. A test-cross methodology was devised for phenotyping of RH progenies, and through forward genetic, the *scs*[ae] locus was pinpointed to a 1.1 Mb-segment containing eight genes. Further, the high resolution provided by RH mapping, combined with chromosome-wise synteny analysis, located the ancestral point of fusion between the telomeric and centromeric repeats of two paleochromosomes that originated chromosome 1D. Also, it indicated that the centromere of this chromosome is likely the result of a neocentromerization event, rather than the conservation of an ancestral centromere as previously believed. Interestingly, location of *scs* locus in the vicinity of paleofusion is not associated with the expected disruption of synteny, but rather with a good degree of conservation across grass species. Indeed, these observations advocate the evolutionary importance of this locus as suggested by "Maan's *scs* hypothesis."

Introduction

The plant cell has tripartite inheritable genetic information unequally distributed between the nuclear, mitochondrial, and chloroplast genomes. In the course of evolution, some plastid genes have migrated to the nuclear genome, preventing any of the three cell components from functioning independently (Chase 2007). As a result the organelles depend on nuclear-encoded genes for their proper functioning and similarly the nucleus relies on the metabolic

activities carried out in the organelles to sustain its molecular activities. This tight cross-talk between the three genomes is referred to as "nuclear-cytoplasmic interactions" (NCI). Whenever the NCI are disrupted, undesirable phenotypes emerge, such as maternally inherited male sterility, female sterility, delayed maturity, reduced vigor, deficiency of chlorophyll, and carotenoids, as well as altered morphology of cotyledons, leaves, and flowers (Maan 1991; Zubko et al. 2001; Bogdanova 2007; Amarnath et al. 2011). This study focused on fine mapping of one of the species cytoplasm-specific (scs) nuclear genes, involved in NCI, utilizing a radiation hybrid (RH) mapping approach.

The effects of scs genes are usually masked in normal plant cells with compatible nucleus and cytoplasm, but when the cytoplasm of wheat (*Triticum turgidum* ssp. *durum* (Desf.) Husnot [AABB; $2n=4x=28$] or *Triticum aestivum* ssp. *aestivum* L. [AABBDD; $2n=6x=42$]) is replaced with the cytoplasm of a wild relative, the importance of scs becomes apparent. This condition, defined as "alloplasmic" (*allo*-alien, *plasmic*-cytoplasm), is typically obtained through substitution backcrossing, where the wild relative is used as the female and wheat is used as the recurrent pollinator (Wu et al. 1998). This technique has been employed to generate cytoplasmic male sterile lines during the early days of hybrid wheat (for review see Cisar and Cooper 2002). However, it is important not to confuse the scs genes with the male fertility restorer genes (*Rf*) because the scs genes are not fertility restorers, but rather maintainers of NCI to provide adequate vigor and viability in alloplasmic plants (Maan 1992a, b, d).

The tetraploid nucleus of durum wheat appears to suffer the alloplasmic condition more dramatically than the hexaploid nucleus of bread wheat. A well-studied example of an incompatible alloplasmic interaction is a line that combines the tetraploid nucleus of wheat with the cytoplasm of *Ae. longissima* (*lo*) [plasmon S^1S^1]. In this case, the alloplasmic line can survive only when one of the 1A chromosomes of the durum nucleus carries a portion of chromosome 1A from *T. timopheevii* ssp. *timopheevii* (AAGG; $2n=4x=28$; Simons et al. 2003) or chromosome 1D from *T. aestivum* (Hossain et al. 2004a). These two chromosomes have been shown to carry scs genes capable of restoring adequate NCI. These genes have been defined as scs^{ti} from *T. timopheevii* and scs^{ae} from *T. aestivum*. The alloplasmic lines, (*lo*) $scs^{ti}−$ and (*lo*) $scs^{ae}−$, that carry either of these genes can be maintained in hemizygous or homozygous condition. Upon test-crossing with genotypes without active scs genes, they produce an equal proportion of shriveled and plump seeds, of which only the latter germinate (Fig. 1; Maan 1992c; Hossain et al. 2004b). These two scs genes have been studied extensively and their general positions on chromosomes 1AL and 1DL are known (Simons et al. 2003; Hossain et al. 2004a). While alternative allelic versions exist at scs^{ti} locus (Gehlhar et al. 2005), the
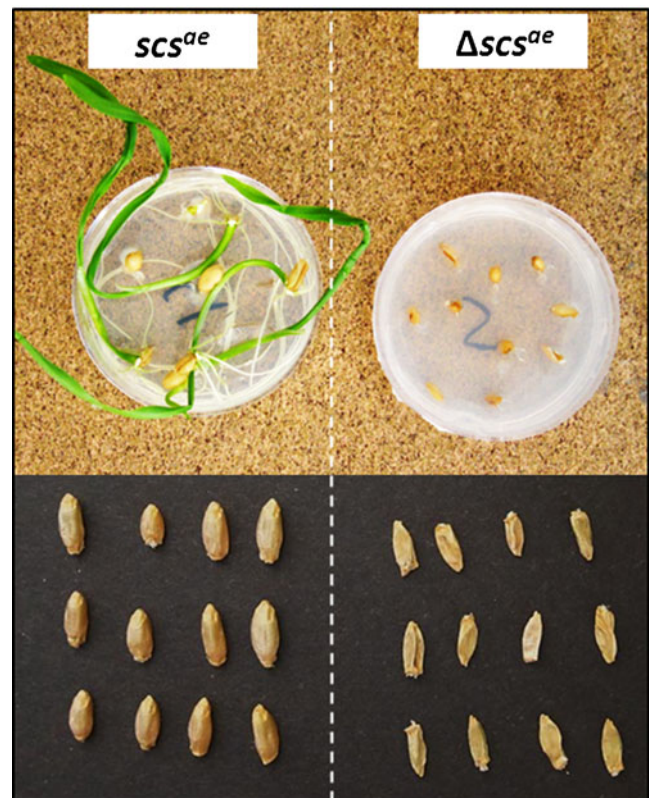

Fig. 1 Alloplasmic durum wheat line with the cytoplasm of *Ae. longissima* can survive only if the scs^{ae} gene is present in its nucleus. Plump seeds (*left*), with the scs^{ae} gene present, germinate properly when cultured on organic media, while shriveled seeds (*right*), without the scs^{ae} gene, do not germinate

scs^{ae} gene on chromosome 1D does not have any known alternative allele. This prevents pinpointing its map location by recombination-based studies. Radiation hybrid mapping is a physical mapping approach based on radiation-induced deletions for mapping markers and phenotypes, and therefore does not rely on allelic polymorphism. In this study the RH methodology is employed to refine the location of the scs^{ae} gene, previously narrowed down to a ~8.3 Mb region in the proximal portion of deletion bin 1DL2-0.41-1.00 (Hossain et al. 2004a; Kalavacharla et al. 2006), by genotyping the RH_1 individuals and phenotyping their RH_2 progeny.

Materials and methods

Plant material

Radiation hybrids

Seeds of durum wheat cultivar "Langdon" (LDN) [AABB; $2n=4x=28$, 13"+1A"] and its aneuploid "Langdon 1D (1A)" (LDN 1D (1A)) [AABB; $2x=4x=28$, 13"+1D"] were obtained from USDA-ARS (Fargo, ND; Joppa 1993). Seeds

of LDN 1D (1A) were equilibrated at approximately 13 % moisture content using a 60 % glycerol solution, and then gamma irradiated at 150 or 350 Greys (Gy) using an Acel Gamma Cell 220 (North Dakota State University, Fargo, ND). The mutagenized M_1 seeds were grown under greenhouse condition, hand emasculated, and used as pollen recipients from non-irradiated LDN plants. Seeds resulting from this artificial pollination belong to the first radiation hybrid generation (RH_1). These are double-monosomic [13" +1A' +1D'^{150} $^{or\ 350\ Gy}$], where any radiation-mediated deletions carried by the chromosomes from LDN 1D (1A) are masked by the non-irradiated LDN chromosomes, with the exception of the monosomic chromosome 1D, where all deletions are fixed. These individuals collectively are referred to as the 1D-RH population. This crossing scheme is explained in detail by Michalak et al. (2009). Also, non-irradiated double-monosomic [13"+1A'+1D'] lines were created and used as experimental controls.

Alloplasmic lines

The wheat alloplasmic stocks have been developed by Shivcharan S. Maan at North Dakota State University (Maan 1973, 1978; Sasakuma and Maan 1978). This study focuses on the durum alloplasmic line that carries the cytoplasm of *Ae. longissima* (*lo*), obtained through over 20 cycles of substitution backcrossing. The resulting plants are male sterile and can be maintained only by artificial crossing. The ((*lo*) *scs*^{ti}− designation identifies a specific durum alloplasmic line, carrying the *scs*^{ti} gene from *T. timopheevii* (Simons et al. 2003). This line is maintained by backcrossing to its B-line *T. turgidum* 56-1 that does not carry any *scs* gene functional for the (*lo*) cytoplasm. The plump seeds resulting from this artificial cross are hemizygous for the *scs*^{ti} gene, and are designated as (*lo*) *scs*^{ti}−. The (*lo*) *scs*^{ae}− line is maintained as a double monosomic (13"+1A'+1D.1AL) with a 1D.1AL translocation, and is hemizygous for the *scs*^{ae} gene (Maan et al. 1999; Hossain et al. 2004b).

Molecular analyses

Marker development

Four strategies were employed for marker development. (a) Retrotransposon junction-based markers (RJ) were designed from the BAC-end sequences of an *Ae. tauschii* [DD; 2*n*= 2*x*=14] library (Li et al. 2004) using the RJ-primer tool available at http://avena.pw.usda.gov/RHmapping/index.html (Wanjugi et al. 2009) and were designated "Ret." Gene-derived markers were designated as NDSU (North Dakota State University) following the indications of the Catalogue of Gene Symbols for Wheat (McIntosh et al. 2008). (b) The markers designated NDSU200 to NDSU212 and NDSU226 to NDSU233 were designed on the known chromosome 1D-specific SNPs (resource available at http://wheat.pw.usda.gov/SNP/project.html; NSF project "Haplotype Polymorphism in Polyploid Wheats and their Diploid Ancestors," PI: J. Dvorak). (c) NDSU213-225 and NDSU290 or higher were developed on the basis of the orthologous information obtained from "Wheat Zapper" (Alnemer et al. 2013 *companion paper* http://wge.ndsu.nodak.edu/wheatzapper/) for *Oryza sativa* L., *Brachypodium distachyon* Beauv., and *Sorghum bicolor* Moench. The "Wheat Zapper" automatic primer design option or Primer Blast (Ye et al. 2012) was used to design primers spanning intron/exon junctions as suggested in Schnurbusch et al. (2007). (d) The NDSU markers designated 3 to 31 were developed following the same procedure as earlier described in (c) with the addition of a sequencing step. Primers spanning the intron/exon boundaries were used to PCR amplify DNA from Phi29-amplified DNA of chromosome arms 1AS, 1AL, 1BS, 1BL, 1DS, and 1DL (Vrána et al. 2000). Preparation of DNA from these particular arms was performed according to the protocol of Wicker et al. (2011). The resulting amplicons were separated by electrophoresis on a 1.5 % agarose gel. The sharpest bands were removed from the gel, eluted following the manufacturer instructions (Gel Elution Kit, Qiagene, Valencia, CA), and sequenced at GeneWiz (South Plainfield, NJ). The sequences from the three homoeologs were then aligned using the ChromasPro software (Technelysium Pty Ltd, Helensvale, QLD, Australia), and homoeologous sequence variations (HSVs) were identified. Primers of 20–23mers were then designed by hand with the chromosome 1D-specific HSV(s) at their 3′ ends.

PCR and gel separation conditions

All markers were tested for chromosome 1D specificity as follows: PCR was carried out in a 20 µl solution containing 1× GoTaq Green Buffer, 2.0 mM of $MgCl_2$, 5 % DMSO, 2.5 mM of each dNTPs, 0.25 µM of each forward and reverse primer, 0.75 units of GoTaq polymerase (Promega, Madison, WI), and 50 ng of DNA from LDN or LDN 1D (1A). The amplification reaction was a standard touch-down method with 10 cycles of 30 s of denaturation at 94ºC, 30 s of annealing with the temperature decreasing 0.5ºC at each cycle, and 45 s of extension at 72ºC, followed by 30 additional cycles with an annealing temperature 5ºC lower than the initial step, and a final extension of 7 min at 72ºC. Four touch-down annealing temperatures were tested for each primer combination: 65ºC, 60ºC, 58ºC, and 55ºC. The resulting amplicons were separated on 9 % non-denaturing polyacrylamide gels and stained with Gel Red (Biotium, Hayward, CA). Only those markers that showed a band present for LDN 1D (1A) and absent in LDN were considered 1D-chromosome specific and used for genotyping.

Genotyping

DNA isolation was carried out as described in Hossain et al. (2004a). Genotyping of RH lines was performed by PCR amplification employing the annealing temperature determined in the 1D-specificity test described earlier. The primer sequences and PCR conditions are available in Table S1. The amplification profile of the RH lines was scored as "1" when the 1D-specific band was present and "0" when the band could not be distinguished. To avoid scoring a failed PCR reaction as a radiation-mediated deletion, the control marker DEASY (Kumar et al. 2012) was multiplexed in each reaction. A score of "0" was given only when the chromosome 1D-specific band was absent but the DEASY band was present. For each marker used for genotyping, the LDN 1D (1A) band was cut from the gel, eluted, and sequenced. The sequence obtained was then aligned against the DNA sequence initially employed for primer design to confirm proper target amplification. The sequences were also searched against the mapped EST database (Peng et al. 2004; Qi et al. 2004; available at http://wheat.pw.usda.gov/GG2/blast.shtml) to identify the deletion-bin they belonged to.
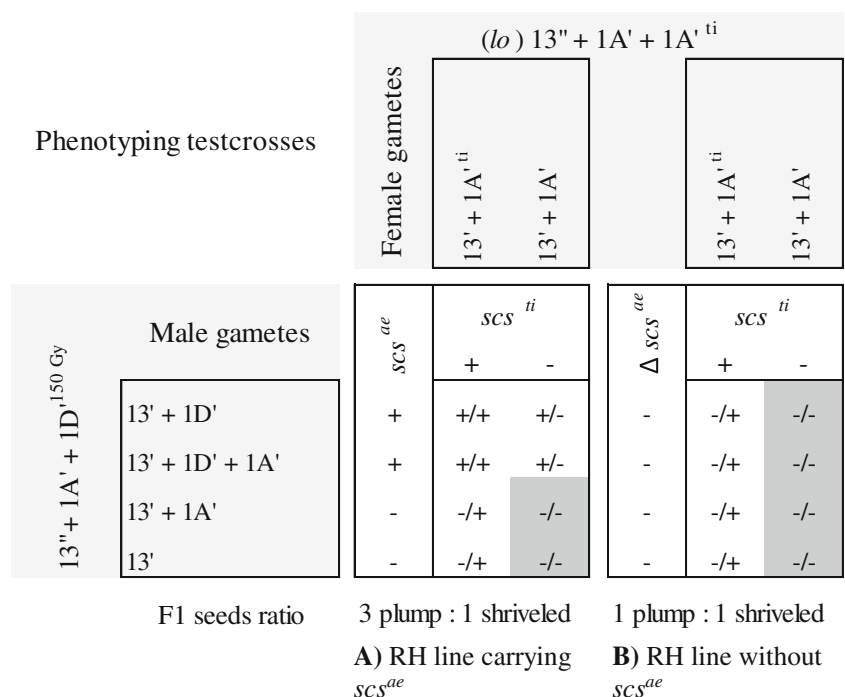
Statistical analysis

The genotyping data was converted into a RH map by employing the Carthagene v1.2.2 software (de Givry et al. 2005) as described by Michalak et al. (2009). The map was then hand curated with respect to the synteny and deletion bin mapping information, but the overall map size was strictly maintained. All other statistical analyses were conducted using the SAS v9.3 package (SAS Institute, Cary, NC). Marker retention frequency is a measure of the frequency at which a given marker locus is deleted within a given RH population, and it is calculated as one minus the ratio between the overall number of genotyped deletions (i.e., "0") and the total number of genotyping data points (Kumar et al. 2012).

Phenotyping

The 1D-RH$_1$ lines are monosomic for chromosome 1D, therefore upon selfing there is theoretically a ~25% chance of losing chromosome 1D (see measured monosomic 1D transfer frequency in Martin et al. 2011). For each RH$_1$ line considered for phenotyping, eight RH$_2$ seeds were planted and screened using NDSU3 and NDSU212. Lines retaining chromosome 1D were considered either monosomic (~75 % chance) or disomic (~25 % chance). Two to three RH$_2$ plants for each RH$_1$ line were selected and used for test-crossing. The presence and number of copies of the scs^{ae} gene in a given genotype can be determined by test-crossing it to a (lo) scs^{ti}− line. Genotypes that do not carry any scs gene in their genome will result in a 1:1 segregation of plump to shriveled seeds (Fig. 2), with the shriveled seeds lacking the scs gene. A genotype hemizygous (i.e. monosomic) for scs will result in a ratio of 3:1 plump to shriveled seeds, while a genotype (disomic) for the scs will produce only plump seeds (i.e., ratio 1:0; Fig. 2). The RH$_2$ lines used in test-crossing are expected to fit the 1:0 or 3:1 ratio when

Fig. 2 Phenotyping of radiation hybrid (RH) lines through test-crossing with a durum alloplasmic line, carrying one copy of scs^{ti} and the cytoplasm of *Ae. longissima* (*lo*). **a** Test-crossed seeds from RH lines with the scs^{ae} gene segregate quasi 3:1 plump to shriveled seeds. **b** Test-crossed seeds from RH lines that have lost the scs^{ae} gene due to radiation-induced deletions segregate 1:1 plump to shriveled

Phenotyping testcrosses

(lo) 13″ + 1A′ + 1A′^{ti}					
Female gametes	13′ + 1A′^{ti}	13′ + 1A′		13′ + 1A′^{ti}	13′ + 1A′

13″ + 1A′ + 1D′$^{150 Gy}$	Male gametes	scs^{ae}	scs^{ti} +	scs^{ti} −	Δscs^{ae}	scs^{ti} +	scs^{ti} −
	13′ + 1D′	+	+/+	+/−	−	−/+	−/−
	13′ + 1D′ + 1A′	+	+/+	+/−	−	−/+	−/−
	13′ + 1A′	−	−/+	−/−	−	−/+	−/−
	13′	−	−/+	−/−	−	−/+	−/−

F1 seeds ratio	3 plump : 1 shriveled	1 plump : 1 shriveled
	A) RH line carrying scs^{ae}	**B)** RH line without scs^{ae}

the scs^{ae} locus has not been affected by the radiation treatment or the 1:1 ratio in cases where a deletion occurred on the locus. The seed ratio of plump to shriveled was statistically tested using the χ^2 analysis. For each phenotyping, at least three independent test-crosses were produced and statistically tested for both shriveled vs. plump and plump vs. shriveled. The degrees of freedom for the analysis was the number of test-crosses times two (plump and shriveled) minus one. Both 1:1 and 3:1 ratios were measured, and the ratio with p of at least 0.05 was considered as the correct ratio and used to predict the state (deleted or retained) of the scs^{ae} locus.

Results

Development of chromosome 1D-specific molecular markers

Four methodologies were followed to enrich chromosome 1D for molecular markers, one based on retroelement junction sites, one employing known gene homoeologous sequence variants (HSVs), and two based on gene colinearity between three sequenced species. A total of 26 retrojunction (RJ) markers were designed on the basis of BAC end sequences originating from the library of *Ae. tauschii* accession AL8/78 (Li et al. 2004); among these, 8 (30.8 %) were confirmed as chromosome 1D specific. All 73 HSVs available for chromosome 1D from the "Haplotype Polymorphism in Polyploid Wheats and their Diploid Ancestors" (PI: J. Dvorak) project were used to design PCR primers and subsequently tested for chromosome 1D specificity. Using the conditions employed in this study, only 21 (28.8 %) of primer pairs could be confirmed as chromosome 1D specific. A total of 94 additional primers were designed spanning intron/exon junctions of predicted orthologous genes of rice, *Brachypodium* and *Sorghum*. Seventeen (~18 %) were confirmed as chromosome 1D specific. Of the remaining primer combinations, 72 were used to amplify amplicons on the homoeologous group 1 flow sorted chromosomes, of which 23 (32 %) amplified amplicons from chromosome 1D and were used for sequencing. For 18 of the sequenced amplicons, it was possible to identify at least one HSV, and 7 were converted into chromosome 1D-specific primers, equal to 10 % of the genes investigated and 39 % of the HSVs identified. Overall, 53 chromosome 1D-specific PCR markers were developed, to which three BAC sequence-based markers (RIOCs) from *Ae. tauschii* accession AL8/78 BAC library (Seth 2009) as well as the UMN25 marker were added (Liu et al. 2008), for a total of 46 gene-based and 11 sequence-based markers.

A comprehensive radiation hybrid map for chromosome 1D

The population of RH_1 lines was generated by gamma radiation followed by artificial pollination. Ninety-four lines were produced at a radiation dosage of 150 Gy, and 94 lines were obtained at 350 Gy. All lines were genotyped with the 57 PCR-based markers described earlier. Surprisingly, the 94 RH_1 lines produced with the lower radiation dosage were the most informative, with a retention frequency of 0.71, as compared to a frequency of 0.98 in the 350 Gy treated lines. The 150 Gy-panel provided a total of 32 lines with breakage (analogous to a "recombination event" in genetic mapping studies) along chromosome 1D, 50 lines with no apparent deletions, and 12 lines missing the entire chromosome 1D (Fig. 3). The informative lines were used to generate a comprehensive map of chromosome 1D at LOD of 10. This map spans 205.2 cR as determined by 36 unique map loci (Fig. 4). For 22 of the markers their deletion bin location is known, and the chromosomal location of the remaining 25 markers can be derived on the basis of their relative map location. Six markers span the telomeric portion of the short arm of chromosome 1D (1DS5-0.70-1.00), none map to the middle portion of 1DS (1DS3-0.70-0.41), eight map within the peri-centromeric and centromeric region (C-1DS3-0.48 or C-1DL2-0.41), and all the remaining markers map on the long arm of chromosome 1D (1DL2-0.41-1.00). The map was strategically enriched with markers spanning the pericentromeric, middle, and distal portion of 1DL for this was the known location of the scs^{ae} gene (Hossain et al. 2004a). The deletion bin corresponding to this region (1DL2-0.41-1.00) represents approximately 37 % of the whole chromosome which is equal to 225 Mb (Dolezel et al. 2009).The forty markers that map within this region span 149 cR with an average resolution of 1.5 Mb per cR.

Colinearity between wheat chromosome 1D, rice, *Brachypodium*, and *Sorghum*

Forty-seven gene-based markers were mapped using the 1D-RH population on the basis of the colinearity predicted by "Wheat Zapper" (Alnemer et al. 2013 *companion paper*). In particular, the re-sequencing of these 47 gene-based markers confirmed a total of 15 orthologous genes on rice chromosome 5 (Os5), *Brachypodium* chromosome 2 (Bd2),
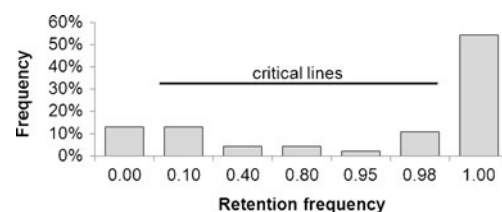

Fig. 3 Retention frequency of 94 radiation hybrid lines for chromosome 1D treated with 150 Gy of gamma rays
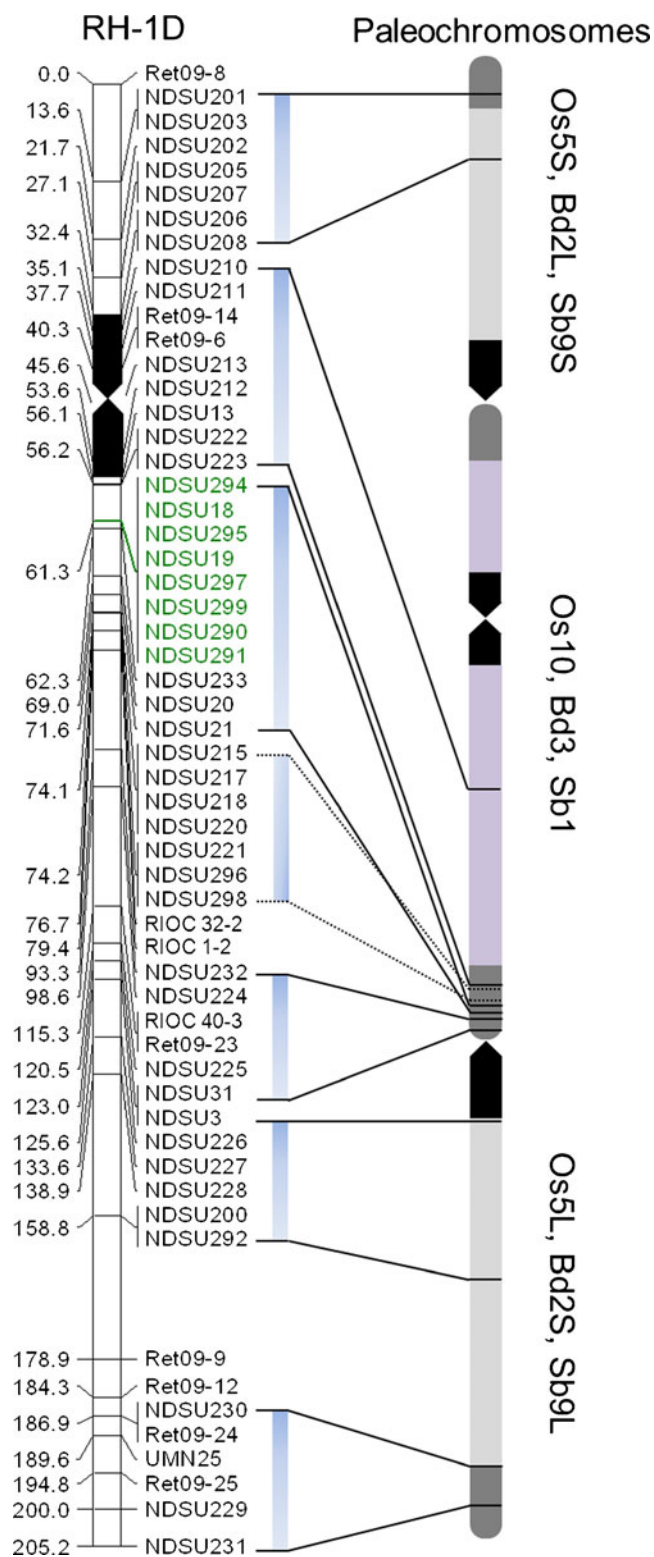
Fig. 4 Radiation hybrid map of chromosome 1D (RH-1D) and its ancestral relationship with rice (Os), *Brachypodium* (Bd), and *Sorghum* (Sb). Locations of the peri- and centromeres are identified with *black boxes*, while the telomeres are indicated in *dark gray*. The blocks of synteny are shown with *vertical blue bars* and their corresponding collinear regions are pointed with *black lines*. The marker loci labeled in green cosegregate with the *scs^ae^* locus

and *Sorghum* chromosome 9 (Sb9) (Fig. 4), and 28 genes on rice chromosome 10 (Os10), 25 genes on *Brachypodium* chromosome 3 (Bd3), and 22 genes on *Sorghum* chromosome 1 (Sb1). Interestingly, the markers collinear to Os5, Bd2, and Sb9 mapped to the distal regions of both arms of chromosome 1D, while Os10, Bd3, and Sb1 show synteny to the proximal regions. In particular, locus *Xndsu201* is orthologous to a gene located 0.2 Mb from the telomeric repeat of Os5S. Locus *Xndsu208* is orthologous to a gene located on the short arm of Os5, 7 Mb distal of the rice centromere (International Rice genome Sequencing Project 2005), while *Xndsu210* corresponds to a gene on the long arm of Os10, 1.5 Mb distal of the rice centromere. The two marker loci are separated by just 2.7 cR. The pericentromeric region in the RH map of chromosome 1D is represented by a 21.1-cR interval, *Xndsu210-Xndsu223*. All loci mapped within this region are orthologous to the middle or distal portion of Os10, Bd3, and Sb1, approximately 1.5 to 10.5 Mb distal of the centromere of Os10. The cosegregating loci *Xndsu31* and *Xndsu3* are orthologous to a gene located 1.7 Mb proximal of the telomere repeats of Os10 and a gene 4.7 Mb distal of the Os5 centromere, respectively. Finally, *Xndsu231* is orthologous to a rice gene placed 2.2 Mb distal of the telomere repeats of Os5L. The distances from the chromosome landmarks described for rice Os5–10 are also valid for Sb1–9 (Paterson et al. 2009) and Bd2–3 (The International Brachypodium Initiative 2010), with the necessary corrections due to their different genome size. The gene order is well conserved among all four species considered; however, microsynteny disruptions occur in the region corresponding to the peri-telomere of Os10, Bd3, and Sb1 (Fig. 4, Table 1). Two syntenic blocks can be identified corresponding to *Xndsu294–Xndsu21* and *Xndsu215–Xndsu298* intervals. These marker loci tend to cosegregate into two groups 12.8 cR apart. Eleven genes belong to the most proximal of these two syntenic groups and seven to the most distal. Based on colinearity, five marker loci (*Xndsu215-221*) of the distal syntenic group should map above the proximal syntenic group, while the two remaining loci (*Xndsu296* and *Xndsu298*) should map within the proximal group based on their location on the genome of rice (Table 1). The ten genes of the proximal syntenic group are all in the proper orthologous position, without any observable order disruption.

A refined location for the *scs^ae^* locus by phenotyping RH₂ lines

To confirm that the activity of *scs^ae^* is not affected by the presence of its *T. timopheevii* counterpart *scs^ti^*, six LDN 1D (1A) genotypes homozygous for *scs^ae^* were test-crossed to the alloplasmic line (*lo*) *scs^ti^−* to generate 134 exclusively plump seeds (shriveled seeds were not observed). This confirmed the ability of LDN 1D (1A) to restore proper seed

Table 1 Fine mapping of the scs^{ae} locus and its syntenic relationship with *Brachypodium* and rice

	cR	3057	3058	3017	3029	3027	3079	3064	*Brachypodium* gene	(Mb)	Rice gene	(Kb)
Plump:shriveled	26:30	26:24	25:25	29:31	24:28	16:17	54:57					
Xndsu211	37.7	0	0	1	1	1	1	1	Bd3g27640	28.6	Os10g30600	15.8
Xndsu213	45.6	0	0	0	1	1	1	1	Bd3g29870	31.9	Os10g35220	18.7
Xndsu212	53.6	0	0	0	0	0	1	1	Bd3g29840	31.9	Os10g35150	18.7
Xndsu13	56.1	0	0	0	0	0	1	1	Bd3g30850	33.1	Os10g37330	19.9
Xndsu222	56.1	0	0	0	0	0	1	1	Bd3g30890	33.2	Os10g37420	20.0
Xndsu223	56.1	0	0	0	0	0	1	1	Bd3g31110	33.3	Os10g37480	20.0
Xndsu294	61.3	0	0	0	0	0	0	0	Bd3g31370	33.5	Os10g37630	20.1
Xndsu18	61.3	0	0	0	0	0	0	0	Bd3g31410	33.5	Os10g37660	20.1
Xndsu295	61.3	0	0	0	0	0	0	0	Bd3g31430	33.6	Os10g37690	20.1
Xndsu19	61.3	0	0	0	0	0	0	0	Bd3g31460	33.6	Os10g37730	20.1
Xscs^{ae}	61.3	0	0	0	0	0	0	0				
Xndsu297	61.3	0	0	0	0	0	0	0	–	–	Os10g37760	20.1
Xndsu299	61.3	0	0	0	0	0	0	0	–	–	Os10g37899	20.2
Xndsu290	61.3	0	0	0	0	0	0	0	Bd3g31630	38.9	Os10g38050	20.3
Xndsu291	61.3	0	0	0	0	0	0	0	Bd3g32000	34.2	Os10g38950	20.7
Xndsu233	62.3	0	0	0	0	0	0	0	Bd3g33240	35.6	Os10g41030	22.0
Xndsu20	69.0	0	0	0	0	1	0	1	Bd3g31520	33.8	Os10g37830	20.1
Xndsu21	71.6	0	0	0	0	1	1	1	Bd3g31530	33.8	Os10g37840	20.2

plumpness to (*lo*) $scs^{ti}-$ genotypes. However, the primary scope of the work presented here is to determine the presence or absence of the scs^{ae} locus in RH_2 lines, which are mainly (~75 %) hemizygous or nullizygous for scs^{ae}, rather than homozygous (~25 %). Six double-monosomic F_1 lines [13"+1A'+1D'] with just one copy of scs^{ae}, and six LDN lines [13"+1A"] nullisomic for scs^{ae} were test-crossed to (*lo*) $scs^{ti}-$ to obtain ratios of 102:35 and 81:85 plump to shriveled seeds, respectively. The χ^2 analysis confirmed ($p=0.05$) that the ratios fit the expected 3:1 and 1:1 pattern, respectively (Fig. 2). Therefore, given that enough test-cross seeds are produced, the double-monosomic progenies of the RH_1 lines can be phenotyped for the presence or absence of one copy of the scs^{ae} gene.

To our advantage, the scs^{ae} locus was previously localized to the proximal portion of the deletion bin 1DL2-0.41-1.00 (Hossain et al. 2004a; Kalavacharla et al. 2006). Seven RH_1 lines were identified as having a breakage within this region. Selfed seeds of these selected lines (RH_2) were used for test-crossing with (*lo*) $scs^{ti}-$, between 33 and 111 seeds were obtained for each line (Table 1). In all seven cases the plump to shriveled ratios fit a 1:1 segregation ($p=0.05$), consistent with the nullizygous state of the scs^{ae} gene (Fig. 2). Subsequently, the RH_2 phenotyping data was converted into the "deletion type" (a.k.a. "haplotype" in genetic mapping studies) of scs^{ae} to map the gene on the RH-map obtained by genotyping RH_1 lines (Table 1). In particular, three RH_1 lines (3027, 3079, and 3064) narrowed the gene location to a region distal of *Xndsu223* and proximal to *Xndsu233*. Even if

none of the seven lines discussed here (Table 1) exclude *Xndsu233* from the scs^{ae} interval, previous studies showed that this marker-locus segregates away from the phenotype (Kalavacharla et al. 2006). The reduced interval spans 6.1 cR equal to approximately 9.1 Mb as derived from the cR to Mb conversion ratio calculated earlier. Based on colinearity (Table 2), the refined scs^{ae} region *Xndsu294–Xndsu299* corresponds to 0.1 Mb of Os10, 0.3 Mb of Bd3, and 0.2 Mb of Sb1. Chromosome 1D of wheat is 604 Mb in size (Dolezel et al. 2009), which is 11 times larger than Os5 and Os10 combined (International Rice genome Sequencing Project 2005), 4 times Bd2 plus Bd3 (The International Brachypodium Initiative 2010), and 5 times Sb1 plus Sb9 (Paterson et al. 2009). Using genome size conversions, it can be estimated that the scs^{ae} region in wheat corresponds to 1.1±1 Mb. This segment is expected to contain 16 orthologous genes, as annotated in one or more of the three sequenced grass model species (Table 2). Among the 16 genes, 10 genes have been mapped on the RH_1 lines, but 4 have been shown to segregate away from the scs^{ae} region.

Discussion

Marker development

The genotype scoring of RH lines is not based on allele polymorphism like in genetic mapping studies, but rather on the presence vs. absence of marker loci. There are two steps

Table 2 Candidate genes within the scs^{ae} interval (adapted from the 'Wheat Zapper' output)

Mapped loci	Distance to scs^{ae} (cR)	Wheat EST	Rice gene (Mb)		Brachypodium gene (Mb)		Sorghum gene (Mb)		Pfam function
Xndsu294	0.0	TC407174	Os10g37630	20.1	Bd3g31370	33.5	Sb1g031330	53.9	UBX-domain
		CJ540389	Os10g37640	20.1	Bd3g31380	33.5	Sb1g031320	53.9	HIT-zinc finger
Xndsu18	0.0	CD885860	Os10g37660	20.1	Bd3g31410	33.5	Sb1g031280	53.8	Trehalase precursor
Xndsu295	0.0	BJ254861	Os10g37690	20.1	Bd3g31430	33.6	Sb1g031260	53.8	*Oxa1* inner membrane
Xndsu296	12.9	CV768721	Os10g37700	20.1	Bd3g31440	33.6	Sb1g031250	53.8	Hydrolase
		TC440753	Os10g37720	20.1	Bd3g31450	33.6	Sb1g031240	53.8	Alpha-beta hydrolase
Xndsu19	0.0	TC388432	Os10g37730	20.1	Bd3g31460	33.6	Sb1g031230	53.8	Pollen ankyrin
		TC406567	Os10g37740	20.1	Bd3g31480	33.6	Sb1g031220	53.8	CGMC_GSK.9
Xndsu297	0.0	TC395424	Os10g37760	20.1	–	–	Sb1g031210	53.8	Rhomboid
		TC412626	Os10g37770	20.1	–	–	–	–	Dehydration response
Xndsu20	6.7	TC370675	Os10g37830	20.2	Bd3g31520	33.8	–	–	OsFBX391
Xndsu21	10.3	TC413544	Os10g37840	20.2	Bd3g31530	33.8	Sb1g031180	53.8	CCB2
		TC397464	Os10g37850	20.2	Bd3g31550	33.8	Sb1g031170	53.7	Armadillo
		TC431553	Os10g37860	20.2	Bd3g31560	33.8	–	–	Expressed protein
Xndsu298	12.9	TC444904	Os10g37880	20.2	Bd3g31570	33.8	–	–	Oxidoreductase
Xndsu299	0.0	TC444904	Os10g37899	20.2	–	–	–	–	Transposon

involved in generating a molecular marker capable of discriminating allelic polymorphism in hexaploid wheat: first the marker needs to uniquely tag only one of the homoeologous chromosomes, and secondly it needs to discriminate the allelic diversity at that locus. While the first step provides discrimination between the homoeologs, which takes advantage of an average frequency of 1 HSV per 24 bp in the expressed portion of the wheat genome, the second step is not as trivial since on average only one SNP per 540 bp discriminates alleles among wheat varieties (Somers et al. 2003). The great advantage of an RH approach is that only the first of the two steps is required to map markers using deletions.

The efficiency of developing RJ markers was lower (31 %) in this study than what was calculated by Paux et al. (2006), who reported an efficiency of 67.2 %. The lower conversion efficiency presented in this study can be due to the smaller number of primer pairs screened here (26) relatively to their study (58). Additionally, in order to develop *T. aestivum* chromosome 3B specific RJ markers Paux et al. (2006) used sequences from a *T. aestivum* BAC library, while here the markers were designed for chromosome 1D of *T. aestivum* from the BAC library of *Ae. tauschii*. The sequence polymorphism that exist between chromosome 1D of *Ae. tauschii* and *T. aestivum* could explain the decreased efficiency observed for this marker system. Still, RJ markers provided the highest success rate among the systems tested for marker development. Also, RJ markers required the least amount of lab work and had the lowest experimental cost between initial design and final genotyping. Thus, RJ markers present the most expedient approach for developing markers in wheat.

For the purpose of fine mapping genes, it is usually best to use EST-based markers, primarily because they provide an entry point into syntenic comparisons with model species, which in turn helps identify possible candidate genes within a region. In this regard, the mere success of scoring does not fully explain the best marker development system. Rather, the best methodology is the one that permits conversion of most genes within targeted region into score-ready markers. This was the case for methodology that employed chromosome flow sorting and sequencing. Even though this approach is more time consuming and relatively more expensive, 39 % of the genes that presented HSVs could be effectively converted into markers. The marker development strategy based on mere design of primers spanning intron/exon junctions followed by screening on high-resolution gels revealed all its limits, with the lowest success ratio at just 18 %. Conversely, it is worth mentioning that regardless of the methodology employed, it is extremely simple and more time effective to develop HSV-based markers for RH, rather than generating polymorphic markers on the basis of SNPs for genetic mapping. Since the flow sorted survey sequences for each chromosome of wheat are currently available, it is now very simple to identify HSVs and design genome-specific primers.

A segregating RH population for chromosome 1D

A radiation hybrid mapping population segregating for chromosome 1D was generated in just one season. A subset of 188 RH$_1$ lines was screened for deletions and the lower

radiation dosage (150 Gy) was found to be the most effective in generating random genomic deletions. It has been shown that a higher dosage of radiation typically causes a higher frequency of genome aberrations (Hossain et al. 2004a). However, in the specific strategy employed here to generate *in vivo* RH$_1$ lines, the radiation mutants are not the final product. Instead, the mutagenized lines are grown under normal greenhouse conditions and their filial generations are used for genotyping. Therefore, the plants need to produce viable progenies. Higher dosages of radiation inevitably result in weaker mutants with lower fitness, and lower likelihood of producing healthy progenies. Hence, the observed higher yield of critical lines from the 150-Gy treatment could be explained as a consequence of healthier mutant lines, which still carried a sufficient number of genomic deletions. The kill curve analyses are typically used to determine the optimal radiation dosage for a study (Wardrop et al. 2002; Hossain et al. 2004a; Al-Kaff et al. 2008; for review Riera-Lizarazu et al. 2008; Kumar et al. 2012). This optimal determination needs to be slightly modified for *in vivo* RH projects, selecting not only on the basis of radiation survival, but also for the good vigor of surviving plants.

The RH$_1$ panel at 150 Gy revealed 34 % of the lines had breakage along chromosome 1D (Fig. 4). This shows the advantage of RH mapping in that even a population of reduced size, as the one employed in this study, still provides sufficient information for mapping. Another advantage of RH mapping over genetic mapping studies is that a population segregating for a trait can be quickly generated within one greenhouse season.

A comprehensive gene-based RH map of chromosome 1D reveals a neo-centromerization event

A subset of 32 segregating RH$_1$ lines provided sufficient information to subdivide chromosome 1D into 36 unique mapping loci, with an overall resolution of 1.5 Mb cR^{-1}. Unfortunately, this resolution is lower than the resolution of the previous RH map of chromosome 1D (Hossain et al. 2004a; Kalavacharla et al. 2006), but similar to what was observed in certain regions of chromosome 3B RH map (Kumar et al. 2012). This lower resolution can be explained by the specific localization of the markers employed in this study. In fact, even though the RH map spans the entire length of chromosome 1D, the markers are clearly not evenly distributed, with an obvious overabundance in the region corresponding to the *scs*^{ae} locus, and a complete lack of markers in the middle portion of chromosome 1DS (Fig. 4). Unfortunately, with the exclusion of the centromere, the short arm of chromosome 1D is represented on the RH map by just eight markers. By comparing to the three sequenced grass species (rice, *Brachypodium*, and *Sorghum*) and to a previously reported map (Luo et al. 2009), it has been determined that a

large segment of chromosome 1D is not represented on this RH map. Since the short arm of chromosome 1D is approximately 37 % of the whole chromosome (Dolezel et al. 2009) with the peri-centromere accounting for ~4 % of the RH map (7.9 cR), it can be estimated that nearly 17 % of the chromosome is missing from this map. The under-representation of markers on the short arm of chromosome 1D might then negatively affect the RH map resolution within this region.

Luo et al. (2009) proposed that chromosome 1D of wheat has evolved from an ancestral fusion between two paleochromosomes corresponding to the modern Os5-10 and Sb9-1 chromosomes. The data gathered here supports this hypothesis, and extends it to Bd2-3, even though a large inversion appears to have occurred between the long and short arm of *Brachypodium* Bd2 (Fig. 4). The precise position and mechanism of chromosomal fusion are still under discussion. However, Luo and colleagues (2009) in their study postulated that the paleochromosome corresponding to Os5, Sb9, and Bd2 likely split at the centromere. The repeated sequences of this region then became "sticky" and fused with the telomeric repeats of a second paleochromosome, corresponding to Os10, Sb1, and Bd3. Finally, the centromere of this second paleochromosome maintained its function and became the centromere of chromosome 1D. Our low marker saturation on the short arm of chromosome 1D does not confirm or contradict the hypothesized centromeric to telomeric fusion point on the short arm. However, our data does help validate the proposed centromeric to telomeric fusion point on the long arm of chromosome 1D. The telomeric region of chromosome 1D also matches the telomeres of *Sorghum*, rice and *Brachypodium*, with the only difference being that chromosome 1DS corresponds to Bd2L and 1DL to Bd2S.

The second part of their hypothesis (Luo et al. 2009) indicated that the modern functional centromere of chromosome 1D corresponds to the ancestral centromere. The centromere position was determined by recombination-based mapping, a methodology that lacks sufficient resolution in this region (Erayman et al. 2004). In fact, the RH map contrasts with this conclusion: the six gene-based markers that map on both sides of the chromosome 1D centromere (as inferred from the deletion bin information) are all colinear to the long arm of Os10, Sb1, and Bd3, rather than to their centromeres (International Rice genome Sequencing Project 2005; The International Brachypodium Initiative 2010; Paterson et al. 2009). Since RH mapping is only indirectly influenced by the low recombination frequency at the centromere (Kumar et al. 2012), it provides sufficient resolution to point out that the centromere of chromosome 1D is likely the result of a neo-centromerization, a fairly common event documented in many species (Ferreri et al. 2005), rather than the functional conservation of an ancestral centromere. Qi et al. (2010) indicated a third hypothesis for the location of the centromere of chromosome 1D,

suggesting its origin from the ancestral location of the Os5 and Bd2 centromeres. The precise origin and location of the 1D centromere still remains controversial and probably only the complete sequencing of this wheat chromosome will provide an answer on its origin.

Fine mapping of the scs^{ae} gene via phenotyping of RH$_2$ lines

Use of radiation hybrids for mapping is a well-established strategy. Despite their advantages, map resolution, speed of population development, and use of monomorphic markers, they have rarely been used in functional studies. To date only radiation mutants have been used in functional studies in plants (Rasko et al. 2000; Kynast et al. 2002; Al-Kaff et al. 2008; Spielmeyer et al. 2008). We have taken radiation hybrid use to the next level, phenotyping and use in a forward genetic study. From a Mendelian genetic perspective, RH lines are ideally suited to study the segregation of a single gene. In fact, the use of near isogenic parents, such as LDN and LDN 1D (1A) to generate the RH$_1$s prevents segregation of any chromosomes, except for the one chromosome under study. Moreover, the homozygosity found in the monosomic state is fixed in just one crossing season compared to the five to six seasons needed for the development of recombinant inbred lines. Finally, genotyping is done on the basis of presence or absence of a gene, simplifying visual scoring.

After an RH map was generated from the RH$_1$ population (Table 1), seven informative RH$_2$ families were phenotyped. These families were selected based on prior knowledge of the scs^{ae} location (Kalavacharla et al. 2006). The scs^{ae} region was then narrowed down to 6.2 cR, calculated to be 1.1 ± 1 Mb based on genome size conversion from the genomes of rice, *Brachypodium*, and *Sorghum*. The scs^{ae} interval, located in close proximity to the centromere, is expected to be rich in transposable elements (TEs) (Neumann et al. 2011). However, the peculiar peri-telomeric origin of this portion of chromosome 1D causes the scs^{ae} interval to be poor in TEs, but abundant in genes. Devos (2010) advocated that TE-poor regions have high gene conservation but Murat et al. (2010) suggested that the gene-rich regions are more prone to rearrangements. Indeed, macro-colinearity is conserved throughout chromosome 1D; however, at the micro-synteny level only small erosions occur in the proximity of the scs^{ae} interval. It is possible that the evolutionary importance of the *scs* gene acts as a stabilizing factor which minimizes variation within this region in different species.

The *scs* genes as likely orchestrators of the evolution of the *Triticeae* tribe

"Maan's *scs* hypothesis" (Maan 1991) postulates that compatible NCIs are pre-established in diploids, and that these interactions guide mating combinations between species. In wheat compatibility between the nucleus and cytoplasm is mediated by the *scs* genes located on the long arm of group 1 chromosomes (Maan 1991). The hexaploid wheat cytoplasm (plasmon B) and the B genome originated from *Ae. speltoides* ssp. *ligustica* (SS; plasmon S; Kilian et al. 2007; Tsunewaki 2009; Haider 2012) suggesting the *scs* gene on chromosome 1B is responsible for compatible NCI. Further evidence of an *scs* gene on chromosome 1B is provided by Asakura et al. (2000) who identified a homoeoallele of scs^{ti} on chromosome 1G. The G genome originated from *Ae. speltoides* (ssp. *speltoides*), which shares similarities with the B genome of wheat (Kilian et al. 2007).

The *scs* genes on chromosomes 1A and 1D ensure compatibility with a cytoplasm of *T. urartu* (AA; plasmon A) and *Ae. tauschii* (plasmon D), respectively (Tsunewaki 1980). Interestingly, scs^{ti} on chromosome 1A and its homoeoallele on chromosome 1G of *T. timopheevii* provide compatibility with the cytoplasm of *Ae. longissima* (plasmon S^1) and *Ae. tauschii*, while the 1A or 1B *scs* genes of *T. turgidum* (plasmon B) do not. It was then determined that two different subspecies of *Ae. speltoides* donated their genomes and plasmons to the two novel tetraploid species. It can also be concluded that two different subspecies or lineages of *T. urartu* donated the A genome to *T. turgidum* and *T. timopheevii*, each harboring a different *scs* allele (Asakura et al. 2000).

These observations advocate for the existence of an ancestral *scs* gene which underwent mutation during speciation and co-evolved with the various cytoplasms. The *scs* genes have a gender-specific preference and can be considered as evolutionary unfair mating referees that assure freedom to the male gametes, while restricting the options for the female gametes on the basis of NCI. Interestingly, the evolutionary drive exercised by the *scs* genes depends on the ploidy level of the nucleus in which they reside (Fig. 5). In diploid progeny of an interspecific hybrid (Fig. 5a), the maternal *scs* gene must be maintained and protected from mutation to assure compatible interaction with the maternal cytoplasm. When the interspecific hybrid produces a polyploid progeny (Fig. 5b) only the maternal copy of the *scs* gene is placed under cytoplasmic pressure, while the "spare" paternal version is left free to accumulate mutations. Hence, the evolution of the *scs* gene happens primarily in polyploids, while diploids act as stable carriers of fixed cytoplasm–allele combinations.

Candidate genes for the scs^{ae} locus

The "Wheat Zapper" tool using rice, *Sorghum* and *Brachypodium*, predicted sixteen orthologous genes within the scs^{ae} interval (Table 2). Four of these genes map just outside the scs^{ae} interval. Six genes have been mapped and found to cosegregate with the scs^{ae} locus. The remaining six
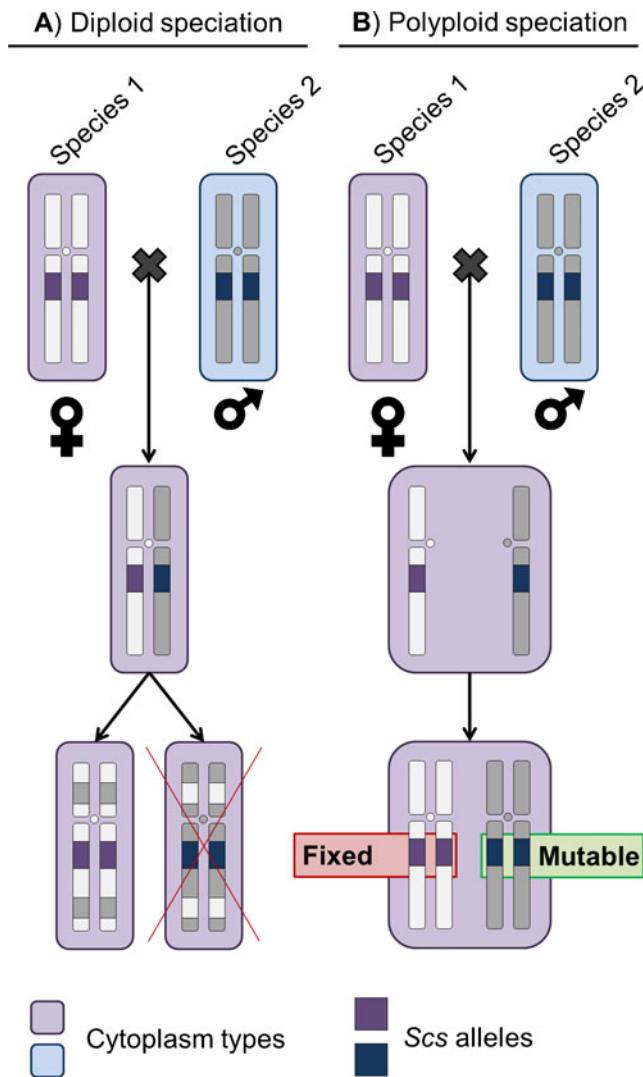
Fig. 5 *Species cytoplasm-specific* genes control the speciation of diploids and polyploids by determining compatible interactions between nucleus and cytoplasm. **a** In interspecific crosses of diploids that result in a heterozygous hybrid, only the progenies that carry the female version of the *scs* allele will survive and evolve into a new species. **b** In interspecific crosses of diploids that result in new polyploid species, all progenies carry the female version of the *scs* allele which is fixed to guarantee survival of the hybrid. The male version of the *scs* allele is free of biological pressure and can accumulate mutations, which can eventually lead to novel versions of the gene

genes have yet to be mapped. Even though small microsynteny erosions occur in this region, the orthologous information has proven very beneficial with 60 % of genes mapping exactly where predicted, and 100 % of the genes within a maximum of 13 cR from their expected positions. There is a possibility that the scs^{ae} gene does not appear in the collinear region of rice, *Brachypodium* or *Sorghum*; however, the proposed extension of the "Maan's *scs* hypothesis" to the whole *Poaceae* family would indicate otherwise. "Wheat Zapper" outputs the Pfam function for all orthologous genes predicted. It is then interesting to discuss the biological functions of these

genes in light of known ability of the *scs* loci to restore compatibility between the nucleus and a specific set of cytoplasms. Since majority of the organellar function is directly encoded by the nucleus (Nott et al. 2006), only those organelles that possess genetic information independent from the nuclear genome appear as the causative agents in differentiating the plasmon types. In plants, only two organelles have genomes not encoded by the nucleus: the mitochondria and the chloroplast. However, the chloroplasts show significantly less structural diversity as compared to the mitochondria (Burger et al. 2003), and are thus not expected to change extensively between plasmons of relatives. Further, the most studied example of incompatible NCI is "cytoplasmic male sterility" (Hanson 1991), and in all cases to date the gene responsible for causing sterility was of mitochondrial origin (Liu et al. 2001; Heazlewood et al. 2003; Akagi et al. 2004; Klein et al. 2005; Saha et al. 2007). Therefore, genes with an activity related to the mitochondria have been suggested as responsible in determining NCI. It can then be hypothesized that the *scs* genes are directly or indirectly involved in mitochondrial activities, with specific *scs* genes interacting with specific mitochondrial types. All of the 12 genes cosegregating or potentially cosegregating with the scs^{ae} locus (6 mapped and 6 unmapped) target the mitochondria directly or indirectly, and at least eight have a predicted function that fits their role as candidates for the scs^{ae} locus.

Among the mapped loci, *Xndsu294* contains the UBX-domain that targets proteins for degradation through the Cdc48/VCP/p97 chaperon ubiquitination machinery (Romisch 2006). This complex dynamically localizes to the mitochondria and endoplasmic reticulum (ER). The UBX-containing gene family participates in an array of cellular functions, but their known involvement in ER-associated degradation (Wang and Lee 2012) could potentially explain the occurrence of shriveled and non-viable seeds (Fig. 1).

Another locus, *Xndsu295*, is derived from a wheat gene homolog of yeast *oxidoreductase1* (*oxa1*), which is known to be involved in the assembly of the cytochrome oxidase (COX) complex and in its interaction with mitochondrial ribosomes (Jia et al. 2003). Yeast knockout mutants for *oxa1* show decreased enzymatic activity of the ATP synthase mitochondria complex, and typically do not survive outside specific growth media (Jia et al. 2007). The known *oxa1* biological functions fit well with what is observed in the scs^{ae} deficient alloplasmic lines, making this gene a good candidate.

The *Xndsu19* locus is homologous to a pollen ankyrin gene. Ankyrins are characterized by a widely occurring repeat motif of 33 amino acids. This motif is thought to be involved in protein-protein interactions (Garcion et al. 2006). More interestingly, these genes have been demonstrated to be involved in embryo development through the recognition between female and male gametes (Yu et al. 2010). Further, this gamete recognition is mediated through the interaction

between ankyrins and the mitochondrial transcription initiator factor SIG5. The role of "unfair mating referee" matching the scs^{ae} gene would be well described by this ankyrin-containing gene.

The *Xndsu297* locus maps within the scs^{ae} interval and is homologous to a *Drosophila rhomboid protease*. This large gene family is comprised of intramembrane serine peptidases and is involved in many biological activities such as cell signaling, programmed cell death, and mitochondria integrity (for review see Knopf and Adam 2012). Their established function in signaling between the nucleus and the mitochondria in patients affected by Parkinson's disease (Whitworth et al. 2008), points out a role for this gene in controlling correct NCI.

In addition to these four mapped candidate genes, another four unmapped orthologous genes within this interval should also be considered as candidate genes. The first contains a HIT-zinc finger domain which is a functional component of many proteins. In particular, *aprataxin* is a DNA strand-break repair protein containing the HIT-zinc finger domain where it was shown to preserve mitochondrial functions in humans (Sykora et al. 2011). The ability of this nuclear gene to regulate mitochondrial activities supports its candidacy for the scs^{ae} function.

Protein kinases promote a variety of activities in plants through peptide phosphorylation. In particular, CGMC_GSK.9 localizes to the mitochondria where it is involved in signal transduction and metabolic processes (Gene Ontology: http://www.gabipd.org/database/cgi-bin/GreenCards.pl.cgi?Mode=ShowBioObject&BioObjectId=2989295). While the information available for this gene does not support a specific activity that would well represent the unique scs^{ae} functionality, the wide array of effects associated with protein kinases and their specific targeting to the mitochondria certainly do not allow discarding this gene as a valid candidate for the scs^{ae} locus.

A similar conclusion can be reached for the dehydration response gene, since evidence indicates seed dehydration is a nuclear directed and mitochondrial driven activity (Song et al. 2009). Also, shriveled seeds (Fig. 1) can be seen physiologically as over-dehydrated endosperms. Hence the scs^{ae} gene might actually be involved in the dehydration process.

Finally, SARM is a TLR adaptor that contains an armadillo motif (Panneerselvam et al. 2012). This nuclear protein targets the mitochondria where the armadillo domain helps control cell apoptosis and acts as an organellar chaperon for other peptides. The specific role of this protein in the interaction between nucleus and cytoplasm is unclear, but its ability to translocate nuclear encoded peptides to the mitochondrial lumen can be hypothesized to cause some of the biological activities associated with the scs^{ae} gene.

Considering the biological functions of the eight orthologous genes described above none can be discarded as a valid candidate for the scs^{ae} locus. Since all genes within this

segment target the mitochondria, the entire interval is probably derived from an ancestral fusion of the mitochondrial genome into the nucleus, a fairly common event in eukaryotic genomes (Leister 2005). Also, the abundance of good candidate genes in close proximity to each other raises the question of whether the scs^{ae} locus is indeed a single gene, or rather a group of genes that exist in very close proximity. Unfortunately, the current 1D-RH map resolution is not sufficient to discard either hypothesis or any of the candidate genes.

References

Akagi H, Nakamura A, Yokozeki-Misono Y et al (2004) Positional cloning of the rice Rf-1 gene, a restorer of BT-type cytoplasmic male sterility that encodes a mitochondria-targeting PPR protein. Theor Appl Genet 108:1449–1457

Al-Kaff N, Knight E, Bertin I et al (2008) Detailed dissection of the chromosomal region containing the Ph1 locus in wheat Triticum aestivum: with deletion mutants and expression profiling. Ann Bot 101:863–872

Alnemer LM, Seetan RI, Bassi FM, et al. (2013) Wheat Zapper: a flexible online tool for colinearity studies in plants. Functional Integrative Genomics (in press)

Amarnath D, Choi I, Moawad AR, Wakayama T, Campbell KH (2011) Nuclear-cytoplasmic incompatibility and inefficient development of pig-mouse cytoplasmic hybrid embryos. Reproduction 142:295–307

Asakura N, Nakamura C, Ohtsuka I (2000) Homoeoallelic gene Ncc-tmp of Triticum timopheevii conferring compatibility with the cytoplasm of Aegilops tauschii in the tetraploid wheat nuclear background. Genome 43:503–511

Bogdanova VS (2007) Inheritance of organelle DNA markers in a pea cross associated with nuclear-cytoplasmic incompatibility. Theor Appl Genet 114:333–339

Burger G, Gray MW, Lang BF (2003) Mitochondrial genomes: anything goes. Trends Genet 19(12):709–716

Chase CD (2007) Cytoplasmic male sterility: a window to the world of plant mitochondrial–nuclear interactions. Trends Genet 23:81–90

Cisar G, Cooper DB (2002) Hybrid wheat. In: Curtis BC, Rajaram S, Macpherson GH (eds) Bread wheat: improvement and production, plant production and protection series no. 30. FAO, Rome, pp 317–330

de Givry S, Bouchez M, Chabrier P, Milan D, Schiex T (2005) Carthagene: multipopulation intergrated genetic and radiated hybrid mapping. Bioinformatics 21:1703–1704

Devos KM (2010) Grass genomes organization and evolution. Curr Opin Plant Biol 13:139–145

Dolezel J, Simkova H, Kubalakova M et al (2009) Chromosome genomics in the Triticeae. In: Feuillet C, Muehlbauer GJ (eds) Genetics and Genomics of the Triticeae. Springer, New York, pp 285–316

Erayman M, Sandhu D, Sidhu D et al (2004) Demarcating the gene-rich regions of the wheat genome. Nucleic Acids Res 32:3546–3565

Ferreri GC, Liscinsky DM, Mack JA, Eldridge MD, O'Neill RJ (2005) Retention of latent centromeres in the Mammalian genome. J Hered 96(3):217–224

Garcion C, Guilleminot J, Kroj T et al (2006) AKRP and EMB506 are two ankyrin repeat proteins essential for plastid differentiation and plant development in Arabidopsis. Plant J 48:895–906

Gehlhar SB, Simons KJ, Maan SS, Kianian SF (2005) Genetic analysis of the species cytoplasm specific gene (scs[d]) derived from durum wheat. J Hered 96(4):404–409

Haider N (2012) Evidence for the origin of the B genome of bread wheat based on chloroplast DNA. Turk J Agric For 36:13–25

Hanson MR (1991) Plant mitochondrial mutations and male sterility. Annu Rev Genet 25:461–486

Heazlewood JL, Whelan J, Millar AH (2003) The products of the mitochondrial orf25 and orfB genes are FO components in the plant F1FO ATP synthase. FEBS Lett 540:201–205

Hossain KG, Riera-Lizarazu O, Kalavacharla V et al (2004a) Radiation hybrid mapping of the species cytoplasm-specific (scs[ae]) gene in wheat. Genetics 168:415–423

Hossain KG, Riera-Lizarazu O, Kalavacharla V et al (2004b) Molecular cytogenetic characterization of an alloplasmic durum wheat line with a portion of chromosome 1D of Triticum aestivum carrying the scs[ae] gene. Genome 47:206–214

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

Jia L, Dienhart M, Schramp M et al (2003) Yeast Oxa1 interacts with mitochondrial ribosomes: the importance of the C-terminal region of Oxa1. EMBO J 22:6438–6447

Jia L, Dienhart MK, Stuart RA (2007) Oxa1 directly interacts with Atp9 and mediates its assembly into the mitochondrial F1Fo-ATP synthase complex. Mol Biol Cell 18:1897–1908

Joppa LR (1993) Chromosome engineering in tetraploid wheat. Crop Sci 33:908–913

Kalavacharla V, Hossain K, Gu YQ et al (2006) High-resolution radiation hybrid map of wheat chromosome 1D. Genetics 173:1089–1099

Kilian B, Özkan H, Deusch O et al (2007) Independent wheat B and G genome origins in outcrossing Ae. progenitor haplotypes. Mol Biol Evol 24:217–227

Klein RR, Klein PE, Mullet JE et al (2005) Fertility restorer locus Rf1 of sorghum (Sorghum bicolor L.) encodes a pentatricopeptide repeat protein not present in the colinear region of rice chromosome 12. Theor Appl Genet 111:994–1012

Knopf RR, Adam Z (2012) Rhomboid proteases in plants – still in square one? Physiol Plant 145:41–51

Kumar A, Bassi FM, Paux E et al (2012) DNA repair and crossing over favor similar chromosome regions as discovered in radiation hybrid of Triticum. BMC Genomics 13:339

Kynast R, Okagaki R, Rines H, Phillips R (2002) Maize individualized chromosome and derived radiation hybrid lines and their use in functional genomics. Funct Integr Genom 2:60–69

Leister D (2005) Origin, evolution and genetic effects of nuclear insertions of organelle DNA. Trends Genet 21:655–663

Li W, Zhang P, Fellers JP, Friebe B, Gill BS (2004) Sequence composition, organization, and evolution of the core Triticeae genome. Plant J 40:500–511

Liu F, Cui S, Horner HT, Weiner H, Schnable PS (2001) Mitochondrial aldehyde dehydrogenase activity is required for male fertility in maize. Plant Cell 13:1063–1078

Liu S, Chao S, Anderson JA (2008) New DNA markers for high molecular weight glutenin subunits in wheat. Theor Appl Genet 118:177–183

Luo MC, Deal KR, Akhunov ED et al (2009) Genome comparisons reveal a dominant mechanism of chromosome number reduction in grasses and accelerated genome evolution in Triticeae. Proc Natl Acad Sci USA 106:15780–15785

Maan SS (1973) Cytoplasmic variability in Triticinae. In: E.R. Sears and. L.M.S. Sears (ed.) Proc IV Int Wheat Genet Symp (pp. 367–373). Columbia, MO

Maan SS (1978) Cytoplasmic relationships among the D- and M-genome Ae. species. In: Ramanujam (ed.) Proc V Int Wheat Genet Symp (pp. 232–250). New Delhi, India

Maan SS (1991) Nucleo-cytoplasmic genetics of wheat. In: Sasakuma T, Kinoshita T (eds) Proceedings of the international symposium on nuclear and organeller genetics of wheat. Hokkaido University, Sapporo, Japan, pp 75–l94

Maan SS (1992a) Genetic analyses of male fertility restoration in wheat: V anomalous results of a monosomic analysis. Crop Sci 32:28–35

Maan SS (1992b) A gene for embryo-endosperm compatibility and seed viability in alloplasmic Triticum turgidum. Genome 35:772–779

Maan SS (1992c) Transfer of the species specific cytoplasm (scs) from Triticum timopheevii to Triticum turgidum. Genome 35:238–243

Maan SS (1992d) The scs and Vi genes correct a syndrome of cytoplasmic effects in alloplasmic durum wheat. Genome 35:780–787

Maan SS, Joppa LR, Kianian SF (1999) Linkage between the centromere and a gene producing nucleocytoplasmic compatibility in durum wheat. Crop Sci 39:1044–1048

Martin A, Simpfendorfer S, Hare RA, Eberhard FS, Sutherland MW (2011) Retention of D genome chromosomes in pentaploid wheat crosses. Heredity 107:315–319

McIntosh RA, Yamazaki Y, Dubcovsky J, et al. (2008) Catalogue of gene symbols for wheat. In: 11th International Wheat Genetics Symposium 24–29 August 2008 Brisbane, Qld Australia

Michalak MK, Ghavami F, Lazo GR, Gu YQ, Kianian SF (2009) Evolutionary relationship of nuclear genes encoding mitochondrial proteins across four grass species and Arabidopsis thaliana. Maydica 54:471–483

Murat F, Xu J-H, Tannier E et al (2010) Ancestral grass karyotype reconstruction unravels new mechanisms of genome shuffling as a source of plant evolution. Genome Res 20:1545–1557

Neumann P, Navrátilová A, Koblížková A et al (2011) Plant centromeric retrotransposons: a structural and cytogenetic perspective. Mob DNA 2:4

Nott A, Jung H-S, Koussevitzky S, Chory J (2006) Plastid-to-nucleus retrograde signaling. Annu Rev Plant Biol 57:739–759

Panneerselvam P, Singh LP, Ho B, Chen J, Ding JL (2012) Targeting of pro-apoptotic TLR adaptor SARM to mitochondria: definition of the critical region and residues in the signal sequence. Biochem J 442:263–271

Paterson AH, Bowers JE, Bruggmann R et al (2009) The Sorghum bicolor genome and the diversification of grasses. Nature 457:551–556

Paux E, Roger D, Badaeva E et al (2006) Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. Plant J 48:463–474

Peng JH, Zadeh H, Lazo GR et al (2004) Chromosome bin map of expressed sequence tags in homoeologous group 1 of hexaploid wheat and homoeology with rice and Arabidopsis. Genetics 168:609–623

Qi LL, Echalier B, Chao S et al (2004) A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genome of polyploidy wheat. Genetics 168:701–712

Qi L, Friebe B, Wu J, et al. (2010) The compact Brachypodium genome conserves centromeric regions of a common ancestor with wheat and rice. Funct and Integr Genomics

Rasko JEJ, Battini J-L, Kruglyak L, Cox DR, Miller AD (2000) Precise gene localization by phenotypic assay of radiation hybrid cells. PNAS 97:7388–7392

Riera-Lizarazu O, Vales MI, Kianian SF (2008) Radiation hybrid (RH) and HAPPY mapping in plants. Cytogenet Genome Res 120:3–4

Romisch K (2006) Cdc48p is UBX-linked to ER ubiquitin ligases. Trends Biochem Sci 31:24–25

Saha D, Prasad AM, Srinivasan R (2007) Pentatricopeptide repeat proteins and their emerging roles in plants. Plant Phys Biochem 45:521–534

Sasakuma T, Maan SS (1978) EMS-induced male-sterile mutants in euplasmic and alloplasmic common wheat. Crop Sci 18:850–853

Schnurbusch T, Collins NC, Eastwood RF et al (2007) Fine mapping and targeted SNP survey using rice-wheat gene colinearity in the region of the Bo1 boron toxicity tolerance locus of bread wheat. Theor Appl Genet 115:451–461

Seth K (2009) High resolution mapping of the scs^{ti} gene in durum wheat and conserved colinearity across three grass genomes: wheat, rice and *Brachypodium*. PhD Thesis Dissertation, North Dakota State University, Fargo

Simons KJ, Gelhar SB, Maan SS, Kianian SF (2003) Detailed mapping of the species cytoplasm-specific (scs) gene in durum wheat. Genetics 165:2129–2136

Somers DJ, Kirkpatrick R, Moniwa M, Walsh W (2003) Mining single-nucleotide polymorphisms from hexaploid wheat ESTs. Genome 49:431–437

Song SQ, Tian MH, Kan J, Cheng HY (2009) The response difference of mitochondria in recalcitrant Antiaris toxicaria axes and orthodox Zea mays embryos to dehydration injury. J Integr Plant Biol 51:646–653

Spielmeyer W, Singh RP, McFadden H et al (2008) Fine scale genetic and physical mapping using interstitial deletion mutants of Lr34/Yr18: a disease resistance locus effective against multiple pathogens in wheat. Theor Appl Genet 116:481–490

Sykora P, Croteau DL, Bohr VA, Wilson DM III (2011) Aprataxin localizes to mitochondria and preserves mitochondrial function. PNAS 108:7437–7442

The International Brachypodium Initiative (2010) Genome sequencing and analysis of the model grass Brachypodium distachyon. Nature 463:763–768

Tsunewaki K (1980) Genetic diversity of the cytoplasm in Triticum and Ae.. pp 49–100, Japan Society for the promotion of science 5-3-1 Kojimachi, Chiyodaku, Tokyo, Japan

Tsunewaki K (2009) Plasmon analysis in the Triticum-Ae. complex. Breed Sci 59:455–470

Vrána J, Kubaláková M, Simková H et al (2000) Flow sorting of mitotic chromosomes in common wheat (Triticum aestivum L.). Genetics 156:2033–2041

Wang CW, Lee SC (2012) The ubiquitin-like (UBX)-domain-containing protein Ubx2/Ubxd8 regulates lipid droplet homeostasis. J Cell Sci 125:2930–2939

Wanjugi H, Coleman-Derr D, Huo N et al (2009) Rapid development of PCR-based genome-specific repetitive DNA junction markers in wheat. Genome 52:576–587

Wardrop J, Snape J, Powell W, Machray GC (2002) Constructing plant radiation hybrid panels. Plant J 31:223–228

Whitworth AJ, Lee JR, Ho VM et al (2008) Rhomboid-7 and HtrA2/Omi act in a common pathway with the Parkinson's disease factor Pink1 and Parkin. Dis Model Mech 1:168–174

Wicker T, Mayer KFX, Gundlach H et al (2011) Frequent gene movement and pseudogene evolution is common to the large and complex genomes of wheat, barley, and their relatives. Plant Cell 23:1706–1718

Wu Y, Zhang C, Liu C, Shuxin R, Yan Z (1998) Breeding technology of alloplasmic wheat. Sci China C Life Sci 41:449–458

Ye J, Coulouris G, Zaretskaya I et al (2012) Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. BMC Bioinforma 13:134

Yu F, Shi J, Zhou J et al (2010) ANK6, a mitochondrial ankyrin repeat protein, is required for male–female gamete recognition in Arabidopsis thaliana. Proc Natl Acad Sci USA 107:22332–22337

Zubko MK, Zubko EI, Ruban AV et al (2001) Extensive developmental and metabolic alterations in cybrids Nicotiana tabacum (Hyoscyamus niger) are caused by complex nucleo-cytoplasmic incompatibility. Plant J 25:627–639