UC Riverside
UC Riverside Previously Published Works

Title

Mapping translocation breakpoints using a wheat microarray.

Permalink

Journal

ISSN

Authors

Bhat, Prasanna R
Lukaszewski, Adam
Cui, Xinping
et al.

Publication Date

2007

DOI

Copyright Information

Peer reviewed

Mapping translocation breakpoints using a wheat microarray

Prasanna R. Bhat[1], Adam Lukaszewski[1], Xinping Cui[2], Jin Xu[2,3], Jan T. Svensson[1], Steve Wanamaker[1], J. Giles Waines[1] and Timothy J. Close[1,*]

[1]Department of Botany and Plant Sciences, University of California, Riverside, California, USA 92521-0124, [2]Department of Statistics, University of California, Riverside, California, USA 92521-0124 and [3]Department of Statistics, East China Normal University, Shanghai, China, 200062

ABSTRACT

We report mapping of translocation breakpoints using a microarray. We used complex RNA to compare normal hexaploid wheat (17 000 Mb genome) to a ditelosomic stock missing the short arm of chromosome 1B (1BS) and wheat-rye translocations that replace portions of 1BS with rye 1RS. Transcripts detected by a probe set can come from all three Triticeae genomes in ABD hexaploid wheat, and sequences of homoeologous genes on 1AS, 1BS and 1DS often differ from each other. Absence or replacement of 1BS therefore must sometimes result in patterns within a probe set that deviate from hexaploid wheat. We termed these 'high variance probe sets' (HVPs) and examined the extent to which HVPs associated with 1BS aneuploidy are related to rice genes on syntenic rice chromosome 5 short arm (5S). We observed an enrichment of such probe sets to 15–20% of all HVPs, while 1BS represents ∼2% of the total genome. In total 257 HVPs constitute wheat 1BS markers. Two wheat-rye translocations subdivided 1BS HVPs into three groups, allocating translocation breakpoints to narrow intervals defined by rice 5S coordinates. This approach could be extended to the entire wheat genome or any organism with suitable aneuploid or translocation stocks.

INTRODUCTION

Interest in translocations and other chromosomal rearrangements stems in part from their association with novel phenotypes, or their usefulness in genetic mapping (1,2). A number of human diseases arise from rearrangements involving unstable regions of DNA (3). Several methods have been tried to map genomic rearrangements. Comparative genomic hybridization (CGH) to microarrays detects numeral (copy number) changes but provides no information on the structure of abnormalities (4), while multicolor spectral karyotyping detects both structural and numeral aberrations but has low resolution (5). End-sequence profiling (ESP) is well suited to the detection of genomic rearrangements and complements existing cytogenetic techniques by making use of a reference genome sequence to identify BAC clones whose ends span the junctions of rearranged positions (6). The speed of translocation breakpoint mapping also has been improved by array painting (7), a method in which chromosomes are flow-sorted and competitively hybridized onto a large insert clone array to reveal a shift in relative hybridization intensity ratios, thus mapping the breakpoint at a resolution defined by size of the clones on the array. Array painting and PCR to amplify the junctions have facilitated the sequencing of translocation breakpoints in humans (7). More recently a single nucleotide polymorphism (SNP) based CGH approach for the high-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping was described for humans (8).

Translocation breakpoint mapping in plants lags far behind that in humans. *In situ* hybridization with total genomic DNA (GISH) has been useful for the characterization of alien introgressions in plants (9). Translocation breakpoints have been used as physical landmarks in cytologically integrated molecular linkage maps of barley chromosomes (10). With recent simplifications, GISH can now be used in large-scale screening for new chromosome constructs. The level of resolution of GISH has been tested in many plant species, for example recombined wheat-rye chromosomes with genetically determined positions of the translocation breakpoints. Depending on the detection method used, the limits of resolution were at least 9.8 and 3.5 cM, and depending on the translocation configuration and the detection method used, some distally located breakpoints could not be visualized at all. These resolution limits of GISH may hamper isolation of critical translocation breakpoints in

---

some chromosome engineering efforts (11). The use of microarrays in plants that lack a complete genome sequence provides an alternative tactic to improve the resolution of translocation breakpoint mapping.

Wheat is one of the most important agricultural crops globally and has a $1x$ genome size of $\sim$5700 Mb. Hexaploid wheat ($6x$) has seven homoeologous groups of chromosomes, each group composed of a representative of the A, B and D genomes (12) for a total ($3x$) genome size of 17 000 Mb. The availability of aneuploids and translocation lines makes wheat an attractive system for mapping using methods other than those that depend on meiotic recombination frequencies among populations segregating marker alleles. The group 1 homoeologous chromosomes have been the most studied, primarily because these chromosomes have several agronomically important genes. This includes 22 genes or QTLs conferring resistance to stripe rust in wild emmer wheat (13,14). In a single gene-rich region 1S0.8 of the short arm of chromosome 1, 41 RFLP markers were identified (15). One tally of markers for wheat group 1 chromosomes placed the total at $\sim$1500, mainly RFLPs and SSRs (16). In a more recent effort, 2212 EST loci were assigned to group 1 chromosomes including 326 associated with 1BS (17) using wheat nullisomic-tetrasomic, ditelosomic and deletion lines.

Here, we report the use of a novel alternative method to develop markers of genes present in a single chromosome arm of hexaploid wheat and mapping of translocation breakpoints using a wheat genome array. We used RNA from aneuploid stocks, a commercially available wheat genome array, and a statistical method that took into consideration the polyploid nature of hexaploid wheat.

MATERIALS AND METHODS

Plant materials

The set of lines used in this study was developed in hexaploid wheat cultivar 'Pavon 76', here referred to as 'Pavon'. Ditelosomic lines for the group-1 chromosomes were produced by centric misdivision of univalents. Centric translocation 1RS.1BL was transferred into Pavon from cv. Genaro, obtained from the International Center for the Improvement of Maize and Wheat (CIMMYT), Mexico. This translocation originated from the cv. Kavkaz (18–20). It was combined with the *ph1b* mutation and a series of homoeologous recombinants 1BS-1RS were produced (21). For this experiment, two recombinants were selected namely (i) $1B + 5$ has a 1B chromosome of Pavon with a terminal segment of the short arm replaced by a corresponding segment of 1RS. (ii) $T - 1$ is a 1RS.1BL translocation with the terminal segment of 1RS replaced by a corresponding segment of 1BS, as illustrated in Figure 1.

RNA extraction, labeling, hybridization to wheat genome array

Seed germination was done in sterile glass dishes. Seedlings were harvested in a well-hydrated state to provide unstressed tissue. Additional seedlings were
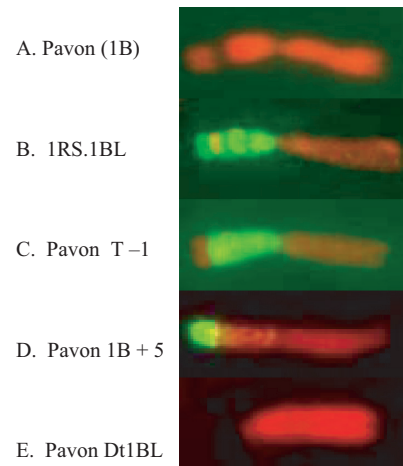

Figure 1. Translocation stocks used in this study. Wheat chromosomes are colored red and introgressed rye segments are green from fluorescence *in situ* hybridization with a probe of the total genomic DNA of rye. Long arm 1BL is present in each case. (A) normal hexaploid Pavon (with normal1BS), (B) Pavon carrying a translocated chromosome in which the entire short arm of rye 1R has replaced the entire short arm of 1B (1RS.1BL), (C) Pavon carrying a translocated chromosome 1B with a proximal segment of rye 1R replacing a proximal region of 1BS ($T-1$), (D) Pavon carrying a translocated chromosome 1B with a terminal segment of rye 1RS replacing a terminal segment of 1BS ($1B + 5$), and (E) Pavon carrying a ditelosomic chromosome 1B missing the short arm (Dt1BL).

exposed to 90% relative humidity in desiccator jars for two days to provide stressed tissue. Total RNA was initially isolated from frozen whole seedlings using TRIzol (Gibco BRL Life Technologies, Rockville, MD, USA) reagent. The RNA was further purified using an RNeasy spin column (Qiagen, Chatsworth, CA, USA) and an on-column DNase treatment. For each genotype, a 50:50 mixture of stressed and unstressed whole seedling RNA was used in each of two replicate samples. The purpose of combining unstressed and stressed RNA from whole seedlings was to maximize the complexity of the RNA population in order to utilize RNA as a surrogate for genomic DNA, as described earlier (22,23). RNA integrity was assessed prior to target preparation using RNA Lab-On-A-Chip (Caliper Technologies Corp., Mountain View, CA, USA) evaluated on an Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA).

Two separate RNA samples per genotype were used to make biotin-tagged cRNAs which were then hybridized to the Affymetrix wheat genome array (a total of 10 chips for five genotypes). Labeling, hybridization, washing and scanning methods were performed as recommended by Affymetrix (Affymetrix GeneChip Expression Analysis Technical Manual; Affymetrix, Inc., Santa Clara, CA, USA) by Sriti Misra at the DNA and Protein Microarray Facility at the University of California, Irvine. The chip images were scanned for visible defects and then extracted using default settings and tabulated as CEL files using Affymetrix GeneChip Operating Software (GCOS 1.2). We used a global scaling factor of 500, a normalization value of 1 (i.e. no normalization procedure is applied to the data), and default parameter settings for the wheat

genome array. The detection calls (present, absent or marginal) for the probe sets were made by GCOS. The present calls ranged from 47.8 to 52.1% (consult Affymetrix manual, 2001 for the details of 'present' calls). For all pairwise comparisons, as long as the genotype containing the 1BS arm yielded a 'present' call for a probe set in duplicate samples, then that probe set was included in the HVP analysis.

Synteny between rice and wheat

The availability of a complete rice genome sequence and the syntenic relationship between rice and wheat provided the means to test the 'high variance probe set' hypothesis. We expected to find an enrichment of probe sets with orthology to genes on rice chromosome 5 short arm known to be syntenic with Triticeae 1S arms (17,24). We searched for rice orthologs of the wheat genes that had high variance probe sets using blastn (e-4 or better) of TIGRs rice version 4 gene models and recorded the physical positions in the rice genome. Graphical displays of the locations of these rice gene models on rice chromosomes were made using GeneSpring (Agilent Technologies, Palo Alto, CA, USA).

STATISTICAL ANALYSIS OF THE ARRAY DATA

The wheat genome array

Sequence information used for development of the Affymetrix wheat genome array includes public content from *Triticum aestivum* UniGene Build #38 (National Center for Biotechnology Information, build date 24 April 2004). Also included are ESTs from the wheat species *T. monococcum*, *T. turgidum*, and *Aegilops tauschii* and GenBank full-length mRNAs from all these species through 18 May 2004. The wheat genome array contains 61 127 probe sets that bind to 55 052 wheat transcripts. For 21 chromosomes in the hexaploid wheat genome, each with two arms, this averages out to ~1310 genes per chromosome arm. The wheat genome array uses probe sets composed of 11 probe pairs to measure the expression of each gene. Each probe pair consists of a perfect match (PM) probe and a mismatch probe.

We extracted the target sequences of all the probe sets from the sequence information file (.sif) for the wheat genome array. The target sequence extends from the 5′ end of the 5′-most probe to the 3′ end of the 3′-most probe. The target sequences were then searched using blastn against the TIGR rice pseudomolecules, release 4 (www.tigr.org/tdb/e2k1/osa1). Annotations for the wheat probe sets are contained in HarvEST:Affymetrix Wheat (http://harvest.ucr.edu/; http://www.harvest-web.org), so each HVP was used to query HarvEST. The corresponding best rice gene models were exported and then plotted on a linear diagram utilizing rice genome coordinates in GeneSpring (Agilent Technologies, Palo Alto, CA, USA).

Method for identifying HVPs

We utilized only the scaled (scaling factor 500) and log2-transformed values from the PM probes. We did not apply more advanced nonlinear normalization methods, such as quantile normalization, which may smooth out the probe level differentiation between two lines. Such differentiation is the primary target of our detection method.

First, we claim that for genes that are present equivalently in both of two genotypes the signal intensities of the probes from the corresponding probe sets shall display a parallel pattern between the two genotypes (see Figure 2A and B), while genes that are not present equivalently shall display two non-parallel patterns. These non-parallel patterns then constitute high variation in probe intensity differences between the two genotypes (see Figure 2C and D). In our case, a probe set for a gene that exists in three homoeologous versions on the A, B and D genomes will give a characteristic pattern of probe intensities that is different from the pattern of probe intensities that would be produced if the B genome contribution is missing. This leads us to use the following algorithm.

(S1) Use only probe sets with 'present' call by GCOS in both replicates of Pavon (complete ABD) samples. This step filters out probe sets with low quality values.

(S2) Let S_{tij} be the log (base 2) scaled observed signal intensity of the jth probe in the ith probe set of genotype t. (e.g. $t = 1, 2$, representing Pavon and Pavon (Dt1BL), respectively.)

(S3) Let $\hat{S}_{tij}$ denote the sample mean of S_{tij} taken over duplicate samples.

(S4) For every probe set i, let the probe difference between two lines be $d_{ij} = \hat{S}_{1ij} - \hat{S}_{2ij}$ and compute the within probe set variances, denoted by v_i, $v_i =$ variance of $\{d_{ij}, j = 1, 2, \ldots, 11\}$

(S5) The candidate probe sets are selected from the right tail of the distribution of v, which corresponds to high within probe sets variance. The actual cutoff may be chosen subjectively by the user based on the distribution of v. In the case of 1BS, it was not difficult to find the value (0.2) with largest descendant rate from the distribution of v (see Supplementary Figure 1).

RESULTS

The 'high variance probe set' hypothesis

Each probe set on an Affymetrix wheat genome array is composed of 11 pairs of 25-mer oligonucleotides, each pair including one perfect match 25-mer and one center-nucleotide-mismatch to the target gene sequence. In the present work, we used only the perfect match probe data (see Materials and methods section), i.e. 11 data points per probe set. Because of the high similarity of genes on homoeologous chromosomes in hexaploid wheat, a given probe set must often measure intensities from all three genomes (A, B and D). But given also that homoeologous genes usually do not have identical sequences, and that 25-mer hybridization is affected by mismatches, then it follows that the relative signal intensities of each individual probe in a probe set must not always be
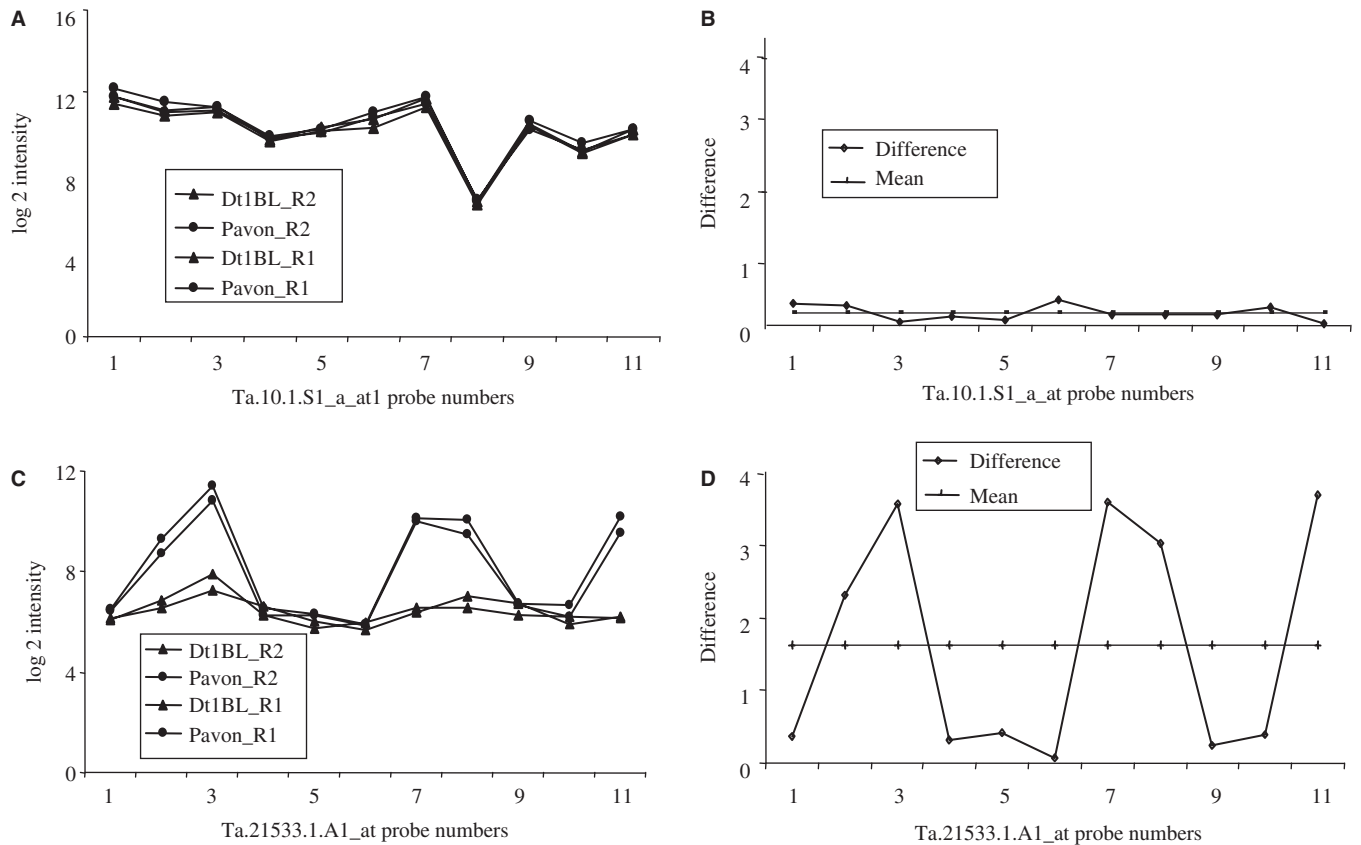
Figure 2. Definition of high variance probe set. (**A**) Signal intensities for low variance probe set Ta.10.1.S1_a_at1 from two replicate samples for each of two genotypes, Pavon and Pavon Dt1BL. Values on *X*-axis indicate probe numbers in the probe set and values on *Y*-axis are log 2 intensities. (**B**) Difference plot generated by subtracting the mean values of Dt1BL from the mean values of Pavon for probe set Ta.10.1.S1_a_at. (**C**) Signal intensities for high-variance probe set Ta.21533.1.A1_at, as in (A). (**D**). Difference plot for probe set Ta.21533.1.A1_at, as in (B).

identical in their contribution to the aggregate signal. Our intention was to exploit this non-identical contribution of individual homoeologous genes to the aggregate signal to assign probe sets to individual chromosome arms using aneuploid stocks of wheat. For this purpose, if we compare the signals that come from all three genomes in normal hexaploid wheat with the signals from an aneuploid missing one chromosome arm, then the differences in signal intensity, when considering each probe in the probe set, will have generally higher variance when a probe set represents a gene that is absent from the aneuploid (resides on the missing arm) than when it is present in the aneuploid (Figure 2C and D). In contrast, when the probe set variance is low (Figure 2A and B) we cannot infer that the gene is located on the missing arm; this could be the result of genes located elsewhere in the genome.

The net consequence of the above consideration is that probe sets which mark genes on the missing chromosome arm should be enriched among 'high variance probe sets' (HVPs). Similarly, we can expect that if a segment of a wheat chromosome arm is replaced by a corresponding segment of a rye homoeologue, then different probe set signal patterns will be observed for the genes involved. Furthermore, it follows that the HVPs, when consistent across a series of recombinant lines, can provide genetic markers that delimit the position of translocation breakpoints. Screening of a set of 1BS-1RS recombinant lines was therefore expected to identify genes present in each bin defined by translocation breakpoints.

Compilation of HVPs

The data from five genotypes were used to compile a list of HVPs by comparing the genotypes in nine of ten possible pairwise combinations (Supplementary Table 1). In order to find an appropriate statistical cutoff point for HVPs (see Materials and methods section), the density of HVPs related to rice gene models on the short arm of rice chromosome 5S was examined as a function of the variance cutoff. The Pavon versus Pavon Dt1BL comparison provided the most extreme case of genetic difference (presence versus absence of the complete 1BS arm) and is shown in Figure 3. At the highest variance cutoff level (1.0), ~32% of HVPs matched rice chromosome 5 genes above a fairly constant background level of 5–10% for each of the remaining 11 rice chromosomes. As the variance cutoff point was reduced, the percentage of rice chromosome 5 genes declined accordingly, indicating that the higher the variance the greater the confidence that HVPs truly match rice chromosome 5 genes. We took
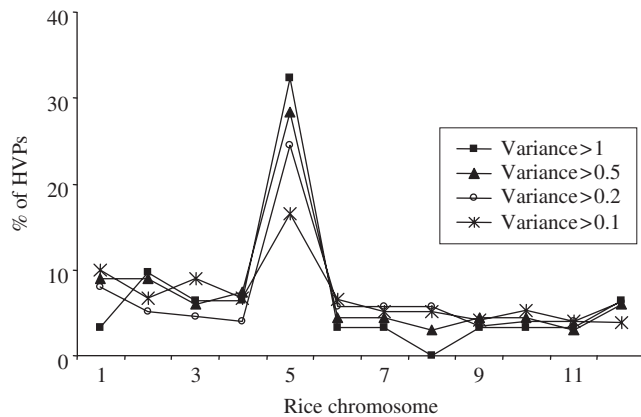
Figure 3. Genome distribution of rice orthologs of wheat genes represented on the wheat genome array. Percent of wheat probe sets (*Y*-axis) with rice gene models on each of 12 rice chromosomes (*X*-axis).

0.1 (17% of HVPs match rice chromosome 5) as an appropriate cutoff point. The range, mean, median and distribution of all HVPs are listed in Supplementary Table 2.

1BS probe sets

Figure 1 illustrates the different cytogenetic stocks used in this study (see Methods section for details). An illustration of some of the genotype comparisons is shown in Figure 4, where wheat chromosomes are colored red and introgressed rye segments are green from fluorescence *in situ* hybridization with a probe of the total genomic DNA of rye labeled with fluorescein isothiocyanate (FITC). In each comparison between normal wheat (containing complete 1B) and wheat with an introgressed 1RS segment replacing a corresponding 1BS segment, the positions of rice orthologs of the enriched wheat probe sets are clustered on the short arm of rice chromosome 5.

The comparison of Pavon versus Pavon Dt1BL targets genes present anywhere on 1BS. This comparison yielded 648 HVPs (variance >0.1), of which 521 had a rice blastn e value of −4 or better. These 521 HVPs matched a total of 432 distinct rice gene models. Among these, 68 gene models (15.7% of the total) were clustered on the terminal portion of the short arm of rice chromosome 5 (5S), which is syntenic to wheat chromosome arm 1BS (17,24). Figure 5 illustrates the physical position of the rice orthologs for these HVPs, notably the high concentration on rice 5S, which constitutes 2.8% of the entire rice genome. These 68 unique rice gene models are represented by 86 wheat probe sets as listed in Supplementary Table 1 and displayed in Figure 5. We also compared Pavon to Pavon 1RS.1BL in which the entire 1BS arm is replaced by the entire rye 1RS arm. This comparison yielded 655 HVPs (variance >0.1), of which 529 had a corresponding rice gene model, 72 (13.6%) of which were clustered on rice 5S (Supplementary online Table 2) as shown in

Figure 4A. This result was very similar to that obtained from the Pavon versus Pavon Dt1BL comparison.

We used translocations at different breakpoints to allocate HVPs to different segments of the wheat 1BS arm. Pavon (1B + 5) has a terminal segment from rye and the proximal segment from wheat. Pavon (T − 1) is a reciprocal configuration. The rye segments in these two recombinants overlap, as explained in Materials and methods section. In a comparison between Pavon (1RS.1BL) and Pavon 1B + 5, the rice gene models corresponding to 30 HVPs are clustered on the proximal region of rice 5S (Figure 4B). In a comparison between Pavon and Pavon 1B + 5, the rice gene models for 54 HVPS are clustered on the terminal portion of rice 5S (Figure 4C). Three segments of the wheat 1B chromosome short arm were defined using these translocations containing two breakpoints.

In total, we identified a collective 390 instances of wheat probe sets corresponding to rice 5S (Supplementary Table 1). Many of these HVPs are common between different comparisons, which adds confidence to the HVPs but reduces the number of unique HVPs to 208. Further redundancy (and confidence) is apparent in the rice 5S gene models, which number only 148 unique genes. The variance scores and rice gene models are listed in Supplementary online Table 1.

Wheat 1BS translocation breakpoint mapping on rice 5S

Rice chromosome 5 short arm is 12.5 Mb in length (http://www.tigr.org/tdb/e2k1/osa1/). To determine the positions of wheat 1BS translocation breakpoints, we made further use of the physical map coordinates of the rice gene models that match the HVPs. We noted the outermost rice gene model within each HVP marker bin derived from the two breakpoints (T − 1 and 1B + 5) (Figure 6). The T − 1 breakpoint falls within the interval of 0.2–0.3 Mb on rice 5S between gene models LOC_Os05g01110.1 and LOC_Os05g01610.1, whereas the 1B + 5 breakpoint was positioned within the 1.6–2.1 Mb interval between rice gene models LOC_Os05g03780.1 and LOC_Os05g04610.1. Therefore, the syntenic region of rice 5S that defines both of these breakpoints is not more than 2.1 Mb from the telomere end.

DISCUSSION

Alternative estimations of HVPs for wheat 1BS

As noted in the Introduction section, a total of 326 ESTs previously were associated with loci on 1BS by DNA hybridization of cDNAs to genomic DNA (17). But these were just a subset of 769 ESTs that were associated with one or more 1S arm. We searched these 769 EST sequences against the content of the Affymetrix wheat GeneChip and then compared the best blastn hits to the HVP list. A total of 2983 HVPs from nine informative pairwise genotype comparisons are compiled in Supplementary Table 1. Of the 769 'mapped' ESTs, 101 had a match to an HVP wheat source sequence. Among these 101, a total of 52 have a rice best hit to a gene model on rice 5S, 42 have a best hit elsewhere in the rice genome,
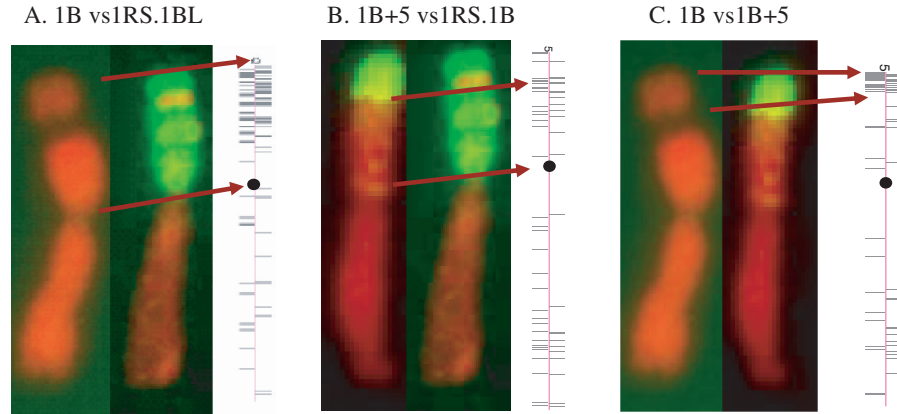
Figure 4. HVP-synteny-based mapping of translocation breakpoints in Pavon (1BS). 1RS.1BL is a translocation of the entire rye short arm of chromosome 1; 1B + 5 is a terminal translocation of rye 1R onto wheat chromosome 1B. Arrows indicate the concentration of rice 5S gene models corresponding to high variance probe sets at >0.1 variance from the three pairwise wheat genotype comparisons shown: (**A**) concentration of HVPs on entire rice 5S in Pavon versus 1RS.1BL, (**B**) concentration of HVPs on proximal rice 5S in 1B + 5 versus 1RS.1BL, (**C**) concentration of HVPs on terminal rice 5S in 1B versus 1B + 5.
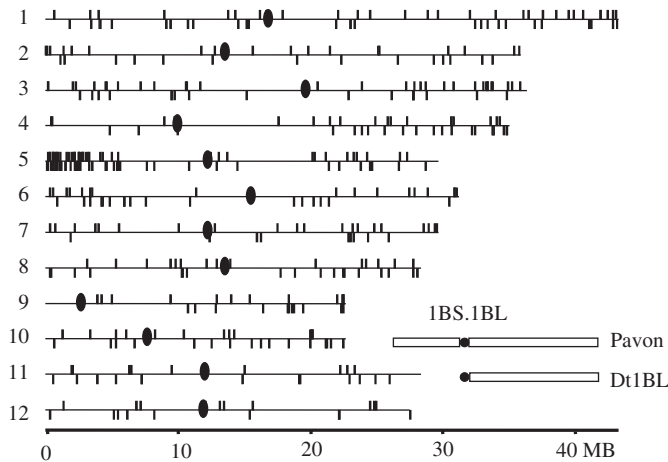

Figure 5. Map position of HVPs between wheat 1BS versus Dt1BL on 12 chromosomes of rice. Inset: Pavon has the 1BS arm while the Pavon Dt1BL is missing 1BS. Rice chromosome 5 short arm is known to be syntenic to wheat chromosome 1BS. Vertical lines indicate rice gene model representing the wheat HVPs and the filled circles are the centromere positions. Numbers indicate which rice chromosomes; values on *X*-axis indicate base pairs. Variance cutoff = 0.1.

and seven have no rice best hit of e − 4 or better. Therefore, we can conclude that the true percentage of HVPs associated with 1BS is approximately double (101/52 × 100% = 194%) the number that results from the more conservative view of only counting those with rice 5S best hits. Among all 2983 HVPs (not just those matching mapped ESTs), 208 HVPs are associated with wheat 1BS based on wheat 1BS-rice 5S synteny. When adding 49 HVPs that do not have rice 5S best hits but match ESTs mapped on wheat 1BS, a total of 257 HVPs could be used preferentially as wheat 1BS markers. Given the homoeologous nature of wheat arms these HVPs must equally mark all three homoeologous arms.

Why do most HVPs not have a blast hit on Rice 5S?

Most of the HVPs (∼80%) in any given pairwise comparison do not have a blast hit on rice 5S. Many might arise from unequal consequences of aneuploidy on gene expression across the entire genome. In other words, deletions of 1BS might cause changes in the relative proportion of A, B and D genome transcripts from genes located anywhere else in the genome. Such changes in relative expression levels would be detected as HVPs. Addition or subtraction of a single chromosome arm altered gene expression patterns extensively in maize (25).

Use of microarrays for genotyping and mapping

The use of microarrays for genotyping by single feature polymorphisms (SFPs) has been reported using DNA in small-genome (140 Mb) Arabidopsis (26). Similarly, RNA was used as a surrogate for DNA to identify SFPs in large-genome (5300 Mb) barley using the Affymetrix barley genome array (22,27). The concept of genotyping with expression microarrays by hybridizing with cRNA instead of genomic DNA for simultaneous genotyping and gene expression analyses was also demonstrated in yeast (28) and subsequently Arabidopsis (29). The possibility to use SFPs for genome-wide association mapping or linkage disequilibrium has also been discussed (30). Here we report another variant of SFP analysis, one that seeks HVPs for assignment of genes to regions of one chromosome arm and translocation breakpoint mapping in polyploid wheat.

The results of this study establish a general approach that could be extended to the entire wheat genome, or to other polyploid organisms for which suitable cytogenetic stocks are available. Specifically, this approach provides a sensitive method to detect deletions, duplications or substitutions within wheat-alien recombinant chromosomes and translocation stocks, including those that carry agronomically desirable genes. Because HVPs are well suited for analyses of cytogenetic stocks, they are
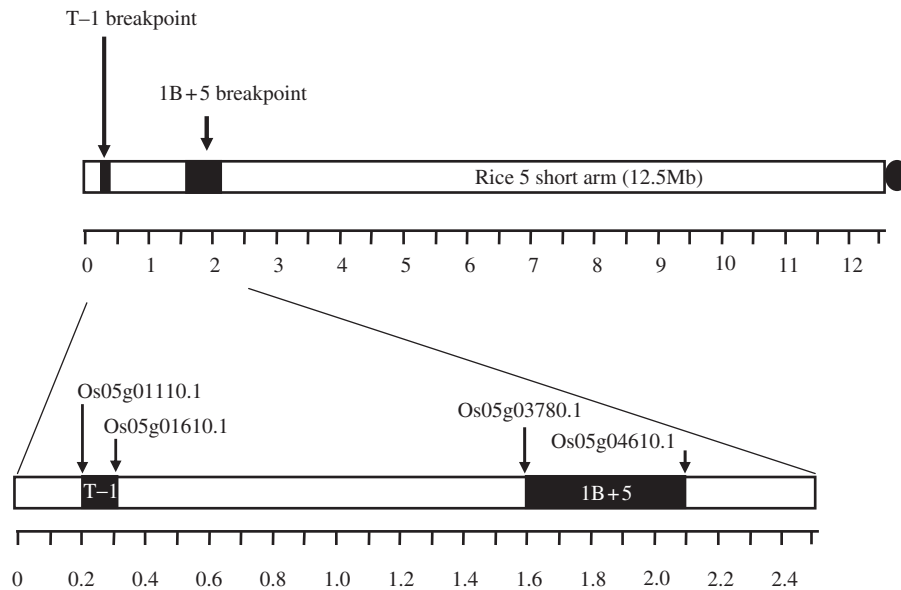
Figure 6. Intervals on rice 5S that correspond to T − 1 and 1B + 5 translocation breakpoints. Rice 5S is 12.5 Mb in length (http://www.tigr.org/tdb/ e2k1/osa1/). The outermost rice gene models from the HVP lists were used to demarcate the outermost possible boundaries of the translocation break points.

attractive for the allocation of genes (represented by probe sets) to chromosome arm bins defined by deletions or translocation junctions, and thus their wider application could accelerate and simplify physical mapping of large genomes such as wheat and other species.

Economical gene chip technology for mapping

Our analysis of the cost of arm-specific genetic marker development comparing this method with conventional PCR amplicon methods indicated that this array-based method is more economical. We designed 91 primer pairs from EST sequences found by others using DNA hybridization to be markers of genes in the wheat 1BS region (http://wheat.pw.usda.gov/cgi-bin/westsql/ map_locus.cgi). We screened the five genotypes used in the current study and detected amplicon presence/absence polymorphisms for only eight primer pairs (results not shown). When weighing the modest up-front costs of the genome array and RNA processing against outlays for PCR reagents and personnel time, it appears that the array-based method is the more economical approach to the development of wheat 1S specific markers.

SUPPLEMENTARY DATA

Supplementary Data is available at NAR online.

ACKNOWLEDGEMENTS

REFERENCES

1. Perkins,D.D. (1986) Determining the order of genes, centromeres, and rearrangement breakpoints in Neurospora by tests of duplication coverage. *J. Genet.*, **65**, 121–144.
2. Antonarakis,S.E. (1993) Human chromosome 21: genome mapping and exploration. *Trends Genet.*, **9**, 142–148.
3. Stankiewicz,P. and Lupski,J.R. (2002) Genome architecture, rearrangements and genomic disorders. *Trends Genet.*, **18**, 74–82.
4. Pinkel,D., Seagraves,R., Sudar,D., Clark,S., Poole,I., Kowbel,D., Collins,C., Kuo,W.L., Chen,C. *et al.* (1998) High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat. Genet.*, **20**, 207–211.
5. Schrock,E., duManoir,S., Veldman,T., Schoell,B., Wienberg,J., FergusonSmith,M.A., Ning,Y., Ledbetter,D.H., BarAm,I. *et al.* (1996) Multicolor spectral karyotyping of human chromosomes. *Science*, **273**, 494–497.
6. Volik,S., Zhao,S.Y., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.Z., Kowbel,D., Huang,G.Q., Lapuk,A. *et al.* (2003) End-sequence profiling: sequence-based analysis of aberrant genomes. *Proc. Natl Acad. Sci. USA*, **100**, 7696–7701.
7. Gribble,S.M., Kalaitzopoulos,D., Burford,D.C., Prigmore,E., Selzer,R.R., Ng,B.L., Matthews,N.S.W., Porter,K.M., Curley,R. *et al.* (2007) Ultra-high resolution array painting facilitates breakpoint sequencing. *J. Med. Genet.*, **44**, 51–58.
8. Peiffer,D.A., Le,J.M., Steemers,F.J., Chang,W.H., Jenniges,T., Garcia,F., Haden,K., Li,J.Z., Shaw,C.A. *et al.* (2006) High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. *Genome Res.*, **16**, 1136–1148.

9. Friebe,B., Jiang,J., Raupp,W.J., McIntosh,R.A. and Gill,B.S. (1996) Characterization of wheat-alien translocations conferring resistance to diseases and pests: current status. *Euphytica*, **91**, 59–87.

10. Kunzel,G., Korzun,L. and Meister,A. (2000) Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. *Genetics*, **154**, 397–412.

11. Lukaszewski,A.J., Lapinski,B. and Rybka,K. (2005) Limitations of *in situ* hybridization with total genomic DNA in routine screening for alien introgressions in wheat, *Cytogenet. Genome Res.*, **109**, 373–377.

12. Sears,E.R. (1969) Wheat cytogenetics. *Annu. Rev. Genet.*, **3**, 451–468.

13. Peng,J.H., (2000) Genomics of wild emmer wheat, *Triticum dicoccoides*: genetic maps, mapping of stripe rust resistance genes and QTLs, for agronomic traits. *Ph.D. Thesis*. University of Haifa, Haifa, Israel.

14. Dilbirligi,M., Erayman,M., Sandhu,D., Sidhu,D. and Gill,K.S. (2004) Identification of wheat chromosomal regions containing expressed resistance genes. *Genetics*, **166**, 461–481.

15. Sandhu,D., Champoux,J.A., Bondareva,S.N. and Gill,K.S. (2001) Identification and physical localization of useful genes and markers to a major gene-rich region on wheat group 1S chromosomes. *Genetics*, **157**, 1735–1747.

16. McIntosh,R.A., Yamazaki,Y., Devos,K.M., Dubcovsky,J., Rogers W.J. and Appels,R. (2003). Catalogue of gene symbols for wheat. In Pogna,N.E., Romano,M., Pogna,E. and Galterio,G. (eds), *Proceedings of the 10th International Wheat Genetics Symposium*, Instituto Sperimentale per la Cerealicotura, Rome. Vol. IV, pp. 1–34.

17. Peng,J.H., Zadeh,H., Lazo,G.R., Gustafson,J.P., Chao,S., Anderson,O.D., Qi,L.L., Echalier,B., Gill,B.S. *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.

18. Rajaram,S., Villareal,R.L. and Mujeeb-Kazi,A. (1990) *The global impact of 1B/1R spring wheat*. Agronomy abstracts. ASA press, Madison, WI, p. 105.

19. Lukaszewski,A.J. (1993) Reconstruction in wheat of complete chromosome-1B and chromosome-1R from the 1Rs.1Bl transloca-tion of Kavkaz origin. *Genome*, **36**, 821–824.

20. Lukaszewski,A.J. and Curtis,C.A. (1993) Physical distribution of recombination in B-genome chromosomes of tetraploid wheat. *Theor. Appl. Genet.*, **86**, 121–127.

21. Lukaszewski,A.J. (2000) Manipulation of the 1RS.1BL transloca-tion in wheat by induced homoeologous recombination. *Crop Sci.*, **40**, 216–225.

22. Cui,X.P., Xu,J., Asghar,R., Condamine,P., Svensson,J.T., Wanamaker,S., Stein,N., Roose,M. and Close,T.J. (2005) Detecting single-feature polymorphisms using oligonucleotide arrays and robustified projection pursuit. *Bioinformatics*, **21**, 3852–3858.

23. Zhang,D., Choi,D.W., Wanamaker,S., Fenton,R.D., Chin,A., Malatrasi,M., Turuspekov,Y., Walia,H., Akhunov,E.D. *et al.* (2004) Construction and evaluation of cDNA libraries for large-scale expressed sequence tag sequencing in wheat (*Triticum aestivum* L.). *Genetics*, **168**, 595–608.

24. Sorrells,M.E., La Rota,M., Bermudez-Kandianis,C.E., Greene,R.A., Kantety,R., Munkvold,J.D., Miftahudin, Mahmoud,A., Ma,X.F., Gustafson,P.J. *et al.* (2003) Comparative DNA sequence analysis of wheat and rice genomes. *Genome Res.*, **13**, 1818–1827.

25. Guo,M., Davis,D. and Birchler,J.A. (1996) Dosage effects on gene expression in a maize ploidy series. *Genetics*, **142**, 1349–1355.

26. Borevitz,J.O., Liang,D., Plouffe,D., Chang,H.S., Zhu,T., Weigel,D., Berry,C.C., Winzeler,E. and Chory,J. (2003) Large-scale identifica-tion of single-feature polymorphisms in complex genomes. *Genome Res.*, **13**, 513–523.

27. Rostoks,N., Borevitz,J.O., Hedley,P.E., Russell,J., Mudie,S., Morris,J., Cardle,L., Marshall,D.F. and Waugh,R. (2005) Single-feature polymorphism discovery in the barley transcriptome. *Genome Biol.*, **6**, R54.

28. Ronald,J., Akey,J.M., Whittle,J., Smith,E.N., Yvert,G. and Kruglyak,L. (2005) Simultaneous genotyping, gene-expression measurement, and detection of allele-specific expression with oligonucleotide arrays. *Genome Res.*, **15**, 284–291.

29. West,M.A.L., van Leeuwen,H., Kozik,A., Kliebenstein,D.J., Doerge,R.W., St Clair,D.A. and Michelmore,R.W. (2006) High-density haplotyping with microarray-based expression and single feature polymorphism markers in Arabidopsis. *Genome Res.*, **16**, 787–795.

30. Kim,S., Zhao,K.Y., Jiang,R., Molitor,J., Borevitz,J.O., Nordborg,M. and Maijoram,P. (2006) Association mapping with single-feature polymorphisms. *Genetics*, **173**, 1125–1133.