A Bacterial Artificial Chromosome Contig Spanning the Major Domestication Locus *Q* in Wheat and Identification of a Candidate Gene

Justin D. Faris,*,[1] John P. Fellers,[†] Steven A. Brooks[‡] and Bikram S. Gill[‡]

**USDA-ARS Cereal Crops Research Unit, Northern Crop Science Laboratory, Fargo, North Dakota 58105, [†]USDA-ARS Plant Science and Entomology Research Unit and [‡]Department of Plant Pathology, Throckmorton Plant Sciences Center, Kansas State University, Manhattan, Kansas 66506*

Manuscript received September 11, 2002
Accepted for publication January 11, 2003

ABSTRACT

The *Q* locus played a major role in the domestication of wheat because it confers the free-threshing character and influences many other agronomically important traits. We constructed a physical contig spanning the *Q* locus using a *Triticum monococcum* BAC library. Three chromosome walking steps were performed by complete sequencing of BACs and identification of low-copy markers through similarity searches of database sequences. The BAC contig spans a physical distance of ~300 kb corresponding to a genetic distance of 0.9 cM. The physical map of *T. monococcum* had perfect colinearity with the genetic map of wheat chromosome arm 5AL. Recombination data in conjunction with analysis of fast neutron deletions confirmed that the contig spanned the *Q* locus. The *Q* gene was narrowed to a 100-kb segment, which contains an *APETALA2* (*AP2*)-like gene that cosegregates with *Q*. *AP2* is known to play a major role in controlling floral homeotic gene expression and thus is an excellent candidate for *Q*.

WHEAT, maize, and rice are all members of the grass family (Poaceae) and provide most of the calories consumed by humans. Domestication of each of these crops occurred quite recently in human history and likely involved the occurrence of mutations at corresponding genetic loci resulting in domesticated traits such as large seed size, reduced shattering, and day-length-insensitive flowering (Paterson *et al.* 1995).

Common wheat (*Triticum aestivum* L. $2n = 6x = 42$, AABBDD genomes) arose ~8000 years ago (Huang *et al.* 2002) as a result of hybridization between the tetraploid wheat *T. turgidum* L. ($2n = 4x = 28$, AABB genomes) and the diploid goat grass *Aegilops tauschii* Coss. ($2n = 2x = 14$, DD genomes; McFadden and Sears 1946). After the formation of the amphiploid, mutations for certain traits such as soft glumes, nonfragile rachis, and the free-threshing character resulted in the domestication of wheat. It is the *Q* gene that confers the square-headed phenotype and free-threshing character of domesticated hexaploid bread wheat (*T. aestivum* ssp. *vulgare*) and tetraploid durum wheat (*T. turgidum* ssp. *durum*). The two major effects of *Q*, free-threshing and square headedness, were once thought to be controlled by two different genes, *q* and *k*, until MacKey (1954) found that the two genes were identical.

Q is thought to be a major regulatory gene for floral development because it exhibits pleiotropic effects on a repertoire of characters important for domestication (MacKey 1954; Muramatsu 1963, 1986). Not only does it confer the free-threshing character, but also it influences glume keeledness, rachis toughness, spike length, spike type, and culm height. Muramatsu (1986) suggested that the range of variation of the characters is very narrow in the absence of *Q* and that they become more obvious in the presence of *Q*. But, the degree to which *Q* influences these characters is dependent on the genetic background and influenced by modifier genes at other loci.

The *q* allele and null mutations result in the speltoid spike phenotype. Wild wheats and hexaploid subspecies such as spp. *spelta*, *vavilovii*, and *macha*, which have the *q* allele, are not free threshing because of tough glumes, and they are susceptible to shattering because they have a fragile rachis (Leighty and Boshnakian 1921; Singh *et al.* 1957; Kabarity 1966). Differences between the ssp. *spelta* and speltoid mutants are slight and probably due to the presence of *q* in ssp. *spelta*, but modifier genes at other loci are likely involved (MacKey 1954).

Q is incompletely dominant to *q*, and plants with the genotype *Qq* have a spike morphology that is intermediate to speltoid and square headed. *Q* is considered to be ineffective in the hemizygous state. Plants with genotypes *Qq* or *Q_* may have spike morphologies that closely resemble speltoid, or they may be more square headed, depending on the genetic background. Muramatsu (1963) showed that *q* has the same effect as *Q* but to a lesser degree. In cytological experiments, he found that four or fewer doses of the *q* allele resulted in plants with speltoid spikes, but five doses resulted in a square-headed spike. In similar experiments, plants that were

[1]*Corresponding author:* USDA-ARS Cereal Crops Research Unit, Northern Crop Science Laboratory, 1307 18th St. N., Fargo, ND 58105. E-mail: justin.faris@ndsu.nodak.edu

monosomic, disomic, trisomic, and tetrasomic for chromosome 5A carrying the *Q* allele resulted in phenotypes that were speltoid, normal (square head), subcompactoid, and compactoid, respectively, indicating that the effects of *Q* on square headedness are dosage dependent (Huskins 1946; Sears 1952, 1954).

Q was found to be present on the long arm of chromosome 5A through cytological experiments of aneuploids (Huskins 1946; Unrau *et al.* 1950; Sears 1952, 1954; MacKey 1954). More recently, its position on recombination-based maps has been determined (Kojima and Ogihara 1998; Kato *et al.* 1999), and more precise physical mapping of the *Q* gene has been carried out using chromosome deletion lines (Miller and Reader 1982; Endo and Mukai 1988; Tsujimoto and Noda 1989, 1990; Ogihara *et al.* 1994; Endo and Gill 1996).

Transposon tagging systems are currently unavailable in wheat, but positional cloning remains an option for gene isolation. The size of the haploid wheat genome is ~16,000 Mb and contains ~80% repetitive DNA sequences. This makes it seemingly laborious to use a map-based approach to clone genes if they are distributed at random throughout the genome. However, recent comparisons of cytogenetic physical maps based on chromosome deletion lines with genetic linkage maps have revealed gene-rich recombination hot spots along wheat chromosomes (Werner *et al.* 1992; Gill *et al.* 1996a,b; Faris *et al.* 2000; Weng *et al.* 2000; Sandhu *et al.* 2001). Each chromosome arm possesses two to three gene-rich recombination hot spots. One such hot spot on chromosome 5B was evaluated in detail by saturating it with molecular markers and assessing recombination within a physical segment that accounted for 4% of the arm (Faris *et al.* 2000). Estimates of the physical-to-genetic distance ratio within the hot spot were <200 kb/cM, a 22-fold increase compared to the genomic average.

Stein *et al.* (2000) since demonstrated that chromosome walking in wheat is feasible by using bacterial artificial chromosome (BAC) clones from the diploid wheat *T. monococcum* ($2n = 2x = 14$, AA genomes). They performed two walking steps resulting in a 450-kb contig spanning the *Lr10* resistance locus and demonstrated perfect colinearity between the physical map of *T. monococcum* and the genetic map of bread wheat on chromosome 1AS. In addition, they overcame the inherent problem of repetitive BAC end sequences by conducting low-pass sequencing of the BACs to identify putative low-copy fragments to use as markers.

We previously reported on the construction of a high-resolution genetic map of the *Q* locus using a genomic targeting approach that combined molecular methodologies with unique wheat cytogenetic stocks (Faris and Gill 2002). Here, we report on the construction of a physical *T. monococcum* BAC contig spanning the *Q* locus and on the identification of a candidate gene.

Figure 1.—Spikes of euploid Chinese Spring (left) and the chromosome deletion line 5AL-7 (right). The Chinese Spring spike is square headed and the 5AL-7 spike is speltoid.

MATERIALS AND METHODS

Plant materials: Populations consisting of 465 F_2 plants derived from Chinese Spring (CS) × CS/*T. turgidum* ssp. *dicoccoides* disomic 5A chromosome substitution line (CS-DIC 5A) and CS/Cheyenne 5A substitution × CS-DIC 5A lines were used for mapping as described in Faris and Gill (2002). CS nullisomic-tetrasomic lines (Sears 1954) for homeologous group 5 chromosomes (N5AT5D, N5BT5D, N5DT5B) and the chromosome deletion lines 5AL-7 and 5AL-23 (Endo and Gill 1996) were used to construct diagnostic Southern blots for verification of the chromosomal location of fragments as described in Faris and Gill (2002). The deletion line 5AL-7 has a speltoid spike because it lacks a distal segment of chromosome 5AL containing the *Q* gene (Figure 1).

For generation of mutants, 10,000 seeds of CS (free-threshing, *QQ* genotype) were exposed to fast neutrons (5 Gy) at the International Atomic Energy Agency, Vienna. From this, 2600 M1 plants were grown to obtain M2 seed. M2 plants were evaluated for spike morphology as well as other characters. Plants with speltoid spikes were considered putative *q* mutants and advanced to the M4 generation for analysis with molecular markers.

BAC library screening and analysis of BAC clones: A BAC library of *T. monococcum* accession DV92 was described in Lijavetsky *et al.* (1999). The library coverage is 5.6 genome equivalents and consists of 276,480 clones arrayed on 15 high-density filters. Hybridization of fragments to the filters was carried out under the same conditions as restriction fragment length polymorphism (RFLP) hybridizations described in Faris *et al.* (2000). Positive clones were identified and subsequently purchased from Dr. Jorge Dubcovsky, University of California-Davis.

BAC plasmid DNA was isolated using standard protocols,

and 1 μg of plasmid DNA was spotted onto a nylon membrane and hybridized with the probe used to detect the clone initially for verification. Approximately 5 μg of plasmid DNA was digested with *Not*I to excise the insert. Sizes of BAC inserts were determined by pulsed-field gel electrophoresis (PFGE) using the CHEF Mapper system (Bio-Rad) according to the manufacturer's instructions.

Sequence analysis and contig assembly: BACs were prepared, sheared, and subcloned according to Brooks *et al.* (2002). Subclone plasmids were sequenced directly with 5 pmol of primer (T3 and T7) in 10-μl reactions using ABI Prism BigDye Terminator (v1.0 and v3.0) Ready Reaction cycle sequencing kits. Completed sequencing reactions were run on an ABI Prism 3700 DNA analyzer, base quality scores called by *phred* version 0.990722.f (Ewing *et al.* 1998; Ewing and Green 1998), and contigs assembled using *phrap* version 0.990319 (http://www.phrap.org). *Consed* version 11.0 (Gordon *et al.* 1998) was used for editing contigs and design of custom primers to complete the sequence of clones spanning gaps. Manual alignment of sequences from clones spanning gaps to assembled contigs was performed with AssemblyLIGN version 1.0.9c (Accelrys). FASTA files of assembled contigs were exported from *Consed* for coding sequence (CDS) identification. GENSCAN 1.0 (http://genes.mit.edu/GENSCAN.html) was used to predict CDSs with maize.smat as the parameter matrix.

Identification of low-copy fragments: Putative gene sequences were defined by results of BLASTn and BLASTx searches against the National Center for Biotechnology Information nonredundant database (Altschul *et al.* 1997; http://www.ncbi.nlm.nih.gov/BLAST/). Sequences were also subjected to searches for similarity to known repetitive sequences in the Triticeae repeat sequence database (TREP; http://www.wheat.pw.usda.gov/ggpages/ITMI/Repeats/index.shtml). If known low-copy genes were present in regions of BACs where markers were desired, then PCR primers were designed on the basis of the gene sequence. The fragments were amplified using standard PCR protocols and used as probes for mapping. If no apparent genes existed in the targeted regions, then sequences that had no similarity to sequences in either database were PCR amplified and hybridized to Southern blots to determine if they were low copy.

Mapping of BAC-derived fragments: Genomic DNA isolation, restriction digestion, gel electrophoresis, probe labeling, hybridization, and membrane washing procedures were performed as described in Faris *et al.* (2000). Images were captured using a Typhoon 9410 phosphorimager (Molecular Dynamics, Sunnyvale, CA). Subfragments of BAC clones ranging from 300 bp to 2 kb were PCR amplified from BACs. PCR products were purified from agarose gels using the QIAquick gel extraction kit (QIAGEN, Chatsworth, CA) and hybridized to diagnostic blots containing DNA of CS, N5AT5D, N5BT5D, N5DT5B, 5AL-7, and 5AL-23 digested with either *Eco*RI or *Hin*dIII. This allowed us to determine if the fragments were low copy *vs.* high copy, and it served as a verification that the BAC-derived fragments were near the *Q* locus on chromosome 5A.

Low-copy fragments were hybridized to the parents of the mapping populations digested with restriction enzymes *Apa*I, *Bam*HI, *Bgl*II, *Dra*I, *Eco*RI, *Eco*RV, *Hin*dIII, *Sac*I, *Sca*I, and *Xba*I to survey for polymorphism. Fragments were subsequently hybridized to digested DNA of the entire CS × CS-DIC 5A population and F_2 plants of the CS-DIC 5A × CS-Cheyenne 5A that possessed recombination events near *Q* using the restriction enzyme giving the clearest polymorphism. The computer program MAPMAKER (Lander *et al.* 1987) V2.0 for Macintosh was used to calculate linkage distances using the Kosambi mapping function (Kosambi 1944) and a LOD of 4.00.

Analysis of mutants: M1:4 putative fast neutron-induced *q* mutants were grown in the greenhouse, and leaf tissue was harvested from two plants of each M4 family for DNA extraction. DNA was digested with either *Hin*dIII or *Sca*I and transferred to nylon membranes by Southern blotting. Probes mapping to the long arm of chromosome 5A, including those generated from the BAC clones, were hybridized to the mutants and visually observed for presence/absence of the 5A fragment.

RESULTS

Faris and Gill (2002) described the targeting of markers to the *Q* locus and construction of a genetic linkage map of the *Q* region with a resolution of 0.1 cM. The closest marker identified was *XksuP16*, which mapped 0.7 cM distal to *Q*. We used this marker to initiate a chromosome walk to the *Q* gene.

Initial BAC library screening: The probe KSUP16 (GenBank accession no. AY170863) was used to screen four high-density *T. monococcum* BAC filters resulting in the identification of one BAC clone (598P15). PFGE of this clone indicated that it was ~130 kb. Upon sequencing of 598P15, we found that the sequence of probe KSUP16 was ~10 kb from one end of the BAC. Therefore, we began searching the sequence of the opposite end of the BAC to try to identify a low-copy sequence that could be useful as a marker. According to database searches, no known genes existed within a region of 70 kb at the end of the BAC opposite *XksuP16*. Much of the 70 kb had similarity to known repetitive sequences (data not shown), but multiple segments had no similarity to any sequences in the databases and were therefore PCR amplified and screened by hybridization to Southern blots. Fourteen primer pairs designed to amplify fragments within the 45 kb of the BAC end opposite *XksuP16* resulted in RFLP probes that all contained highly repetitive sequences. We therefore designed PCR primers to amplify segments that had similarity to a receptor-like kinase (probe P15-4849; GenBank accession no. AY170868) and a hypothetical protein (probe P15-6163; GenBank accession no. AY170869), which were located at distances of 38 and 51 kb from *XksuP16*, respectively. Both sequences were low copy and subsequently used for mapping (Figure 2).

Probe P15-4849 mapped between *XksuP16* and *Q* at distances of 0.2 and 0.4 cM, respectively. From this we were able to orient the BAC and determined that it extended toward the *Q* gene. Marker *P15-6163*, which was only 13 kb from *P15-4849*, mapped between *P15-4849* and *Q* at distances of 0.1 and 0.4 cM, respectively. One F_2 plant was found to contain a double crossover singleton at *P15-6163*, which was responsible for the recombination between *P15-4849* and *P15-6163* (Figure 3). Thus, the initial screening of the BAC library resulted

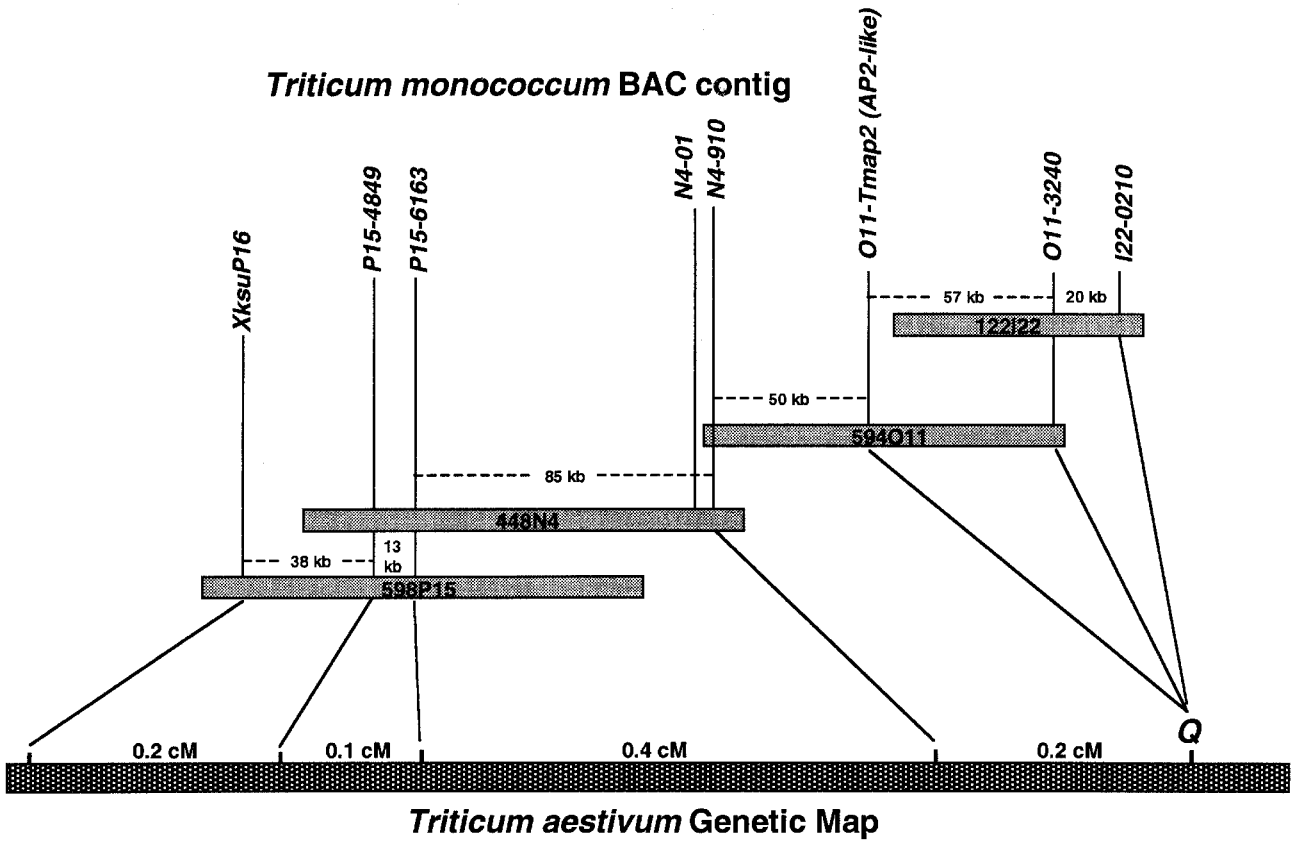

FIGURE 2.—The *T. monococcum* BAC contig anchored to the *T. aestivum* genetic map of the region containing the *Q* locus. Physical and genetic positions of *XksuP16* and BAC-derived markers are shown on the *T. monococcum* contig map and on the *T. aestivum* genetic map.

in 51 kb spanning 0.3 cM, giving a physical-to-genetic distance ratio of 170 kb/cM.

Chromosome walking: *Step 1:* Probe P15-6163 was used to screen the entire *T. monococcum* BAC library, resulting in the identification of three new BACs (170-L14, 448N4, and 619J11). BAC 448N4 was determined to be ~135 kb by PFGE and selected for sequencing. Sequence analysis revealed that it extended ~30 kb beyond the proximal end of BAC 598P15. Within the 30-kb region, BLASTx and BLASTn searches revealed a cluster of hypothetical proteins within 20 kb of the proximal end. Six primer pairs were designed to amplify segments of this region and resulted in the identification of two low-copy probes (N4-01 and N4-910; GenBank accession nos. AY170864 and AY170865, respectively) ~8 kb apart.

Probe N4-910 mapped between *P15-6163* and *Q* at distances of 0.4 and 0.2 cM, respectively (Figure 2). As with *P15-6163*, *N4-910* detected one F_2 plant that contained a double crossover singleton, causing the map distances to inflate (Figure 3). Our first walking step spanned 85 kb corresponding to 0.4 cM, giving a physical-to-genetic ratio of 212.5 kb/cM.

Step 2: Probe N4-910 was used to rescreen the entire *T. monococcum* BAC library, resulting in the identification of three new BAC clones. Plasmid dot-blot analysis revealed that probe N4-01 did not hybridize to BAC 594O11. Because the N4-01 sequence was only ~8 kb from N4-910 (the probe used to screen the library), this indicated that 594O11 had <17 kb of overlap with BAC 448N4. PFGE of 594O11 indicated that it was 110 kb.

Sequence analysis of BAC 594O11 revealed that it contained a large amount of repetitive retroelement-like sequences (data not shown). However, ~50 kb from the N4-910 sequence was a 3.5-kb sequence that had similarity to the *APETALA2* (*AP2*) gene from Arabidopsis and other genes known to possess the highly conserved AP2 DNA-binding domain such as the maize *indeterminate spikelet* (*ids*) and *glossy15* genes. Primers were designed to amplify a fragment of the *AP2*-like gene from BAC 594O11. The resulting PCR product was used as a probe (O11-Tmap2; GenBank accession no. AY170867) for mapping and was found to cosegregate with the *Q* gene. Thus, this walking step spanned ~50 kb of physical distance corresponding to 0.2 cM.

Three fragments at the proximal end of BAC 594O11 that did not have significant similarities to database sequences were PCR amplified and tested as probes, but only one of these, O11-3240, was low copy. The sequence for probe O11-3240 (GenBank accession no. AY170866)

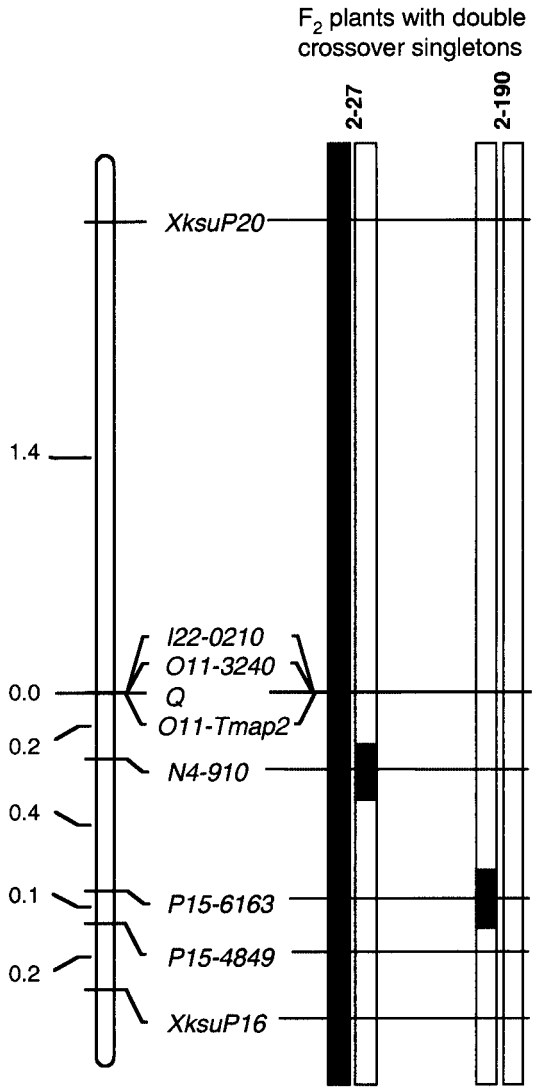

FIGURE 3.—Genotypes of two plants possessing double crossover singletons at markers near the *Q* locus. ■, *T. turgidum* ssp. *dicoccoides*; □, Chinese Spring chromatin. In plant 2-27, marker *N4-910* is homozygous for ssp. *dicoccoides* and adjacent regions are heterozygous. In plant 2-190, marker *P15-6163* is heterozygous and adjacent regions are homozygous for Chinese Spring.

lies ~3 kb from the end of the BAC and ~50 kb from *O11-Tmap2*. But when mapped, it was found to also cosegregate with the *Q* gene. Thus, we found no recombination events between *O11-Tmap2* and *O11-3240*, which were separated by ~50 kb, indicating somewhat suppressed recombination compared to the adjacent distal region.

Step 3: Probe O11-3240 was used to screen the entire *T. monococcum* BAC library but only one BAC clone (122I22) was identified. PFGE indicated that this clone was ~80 kb. Sequencing revealed that it contained a large amount of repetitive sequences, but also much that had no significant similarities to database sequences. Three fragments from 122I22 that lacked similarity to database sequences were PCR amplified and tested as probes, but only one fragment (probe I22-0210; GenBank accession no. AY170870) was low copy. The sequence of probe I22-0210 was ~20 kb from *O11-3240*, and mapping of I22-0210 revealed that it, too, cosegregated with the *Q* gene. Therefore, we had now walked 70 kb without spanning a recombination event in our population, providing further evidence that we were traversing through a region of suppressed recombination.

Analysis of the *AP2*-like gene sequence: As mentioned above, an *AP2*-like gene sequence was identified on the *T. monococcum* BAC 594O11 and found to cosegregate with the *Q* gene in *T. aestivum*. The known functional properties of characterized *AP2*-like genes from various plant species suggests that the *T. monococcum AP2*-like gene is an excellent candidate for *q* (see DISCUSSION). In its efforts to develop and map a unigene set consisting of 10,000 wheat expressed sequence tags (ESTs), the International Triticeae EST Consortium has identified the cDNA (GenBank accession no. BG313955) for the *AP2*-like gene from the free-threshing (*QQ*) wheat CS. The sequence analysis and predicted protein alignments revealed that the CS *AP2*-like gene has at least two amino acids that differ from the *T. monococcum* gene. These differences within the *T. monococcum AP2*-like gene may be responsible for the speltoid phenotype due to altered, or reduced, function.

Analysis of fast neutron *q* mutants: From 2600 fast neutron-induced M1 plants, 26 putative *q* mutants were identified. Two M1:4 plants from each of the 26 families were analyzed using markers on the long arm of 5AL, particularly those derived from the BAC contig at the *Q* locus. Of the 26 putative mutants, all were either homozygous or hemizygous for deletions of the O11-Tmap2 sequence. Phenotypic analysis of these plants indicated that they all had speltoid spikes, confirming the absence of *Q*.

Most *q* deletion mutants contained chromosome deletions that were quite large and some were missing much of the long arm of chromosome 5A. The smallest deletion was observed in plant fndel-143. The plant had an obvious speltoid phenotype and was therefore lacking the expression of the *Q* gene. In addition, DNA analysis revealed that *O11-Tmap2* and several distal markers were absent from this plant, but *O11-3240* and all proximal markers were present (Figure 5). Therefore, the proximal deletion breakpoint in fndel-143 physically separated *O11-3240* from *O11-Tmap2* and eliminated anything proximal to *O11-3240* as a candidate for the *Q* gene. This result confirmed that our *T. monococcum* BAC contig spanned the *Q* locus and narrowed the region for prospective *Q* gene candidates to a segment between *N4-910* and *O11-3240*, which is ~100 kb.

It is difficult to precisely estimate the size of the smallest deletion since we do not have a BAC contig assembled for the entire region spanning the deletion. But,

```
Maize_ids            MVLDLNVASPADSGTSSSSVLNSADGG---FRFGLLGSPVDDDDCSGEMAPGASTGFMTR 57
Barley_AP2-like      MVLDLNVESPADSGTSSSSVLNSADAAG-AFRFGLLGSPDDDD-CSGELAPAAASGFVTR 58
Tmap2                MVLDLNVESPADSGTSSSSVLNSADAAGGGFRFGLLGSPDDDD-CSGEPAPVGSG-FVTR 58
Wheat_AP2-like       ------------------------------------------------------------

Maize_ids            QLFPSPTPPAE--------PEPEPVAAPVPVWQPQRAEDLGMAQKPVAPAKNTRRGPRSR 109
Barley_AP2-like      QLFPAPPPAPG-----VMMGQAPAPPPTAPVWQPRRAEELVVAQR-VAPKKKTRRGPRSR 112
Tmap2                QLFPASPPGHAGAPGMMMGQQAPAPAPMAPVWQPRRAEELVVAQR-VAPAKKTRRGPRSR 117
Wheat_AP2-like       ------------------------------------------------------------

Maize_ids            SSQYRGVTFYRRTGRWESHIWDCGKQVYLGGFDTAHAAARAYDRAAIKFRGLDADINFSL 169
Barley_AP2-like      SSQYRGVTFYRRTGRWESHIWDCGKQVYLGGFDTAHAAARAYDRAAIKFRGLEADINFNL 172
Tmap2                SSQYRGVTFYRRTGRWESHIWDCGKQVYLGGFDTAHAAARAYDRAAIKFRGLEADINFNL 177
Wheat_AP2-like       ----------------------XXXVYLGGFDTAHAAARAYDRAAIKFRGLEADINFNL 37
                                            **************************:*****.*

Maize_ids            SDYEDDLKQMRNWTKEEFVHILRRQSTGFARGSSKYRGVTLHKCGRWEARMGQLLGKKYI 229
Barley_AP2-like      SDYEEDLKQMRNWTKEEFVHILRRQSTGFARGSSKYRGVTLHKCGRWEARMGQLLGKKYI 232
Tmap2                SDYEEDLKQMRNWTKEEFVHILRRQSTGFARGSSKYRGVTLHKCGRWEARMGQLLGKKYI 237
Wheat_AP2-like       SDYEEDLKQMRNWTKEEFVHILRRQSTGFARGSSKYRGVTLHKCGRWEARMGQLLGKKYI 97
                     ****:************************************************************

Maize_ids            YLGLFDSEVEAARAYDRAALRFNGREAVTNFEPSSYNAGDNNLRDTETEAIDDGDAIDLD 289
Barley_AP2-like      YLGLFDSEVEAARAYDRAAIRFNGRDAVTNFDSSSYNG--DATPDVENEAIVDADALDLD 290
Tmap2                YLGLFDSEVEAARAYDRAAIRFNGREAVTNFESSSYNG--DAPPDAENEAIVDADALDLD 295
Wheat_AP2-like       YLGLFDSEVEAARAYDRAAIRFNGREAVTNFESSSYNG--DAPPDAENEAIVDADALDLD 155
                     ******************:*****:*****:.****.  :   *.*.*** *.**:***

Maize_ids            LRISQPNVQDPKRDNTLAGLQPTCDSPESSNTMASQPMSSSSP--WPGYHQNPAVS-FHH 346
Barley_AP2-like      LRMSQPTAHDPKRDNIIAGLQLTFDSPESSTTMVSSQPMSSSS-QWPVHQHGTAVPPQQH 349
Tmap2                LRMSQPTAHDPKRDNIIAGLQLTFDSPESSTTMVSSQPMSSSSSQWPVHQHGTAVAPQQH 355
Wheat_AP2-like       LRMSQPTAHDPKRDNIIAGLQLTFDSPESSTTMLSSQPMSSSSSQWPVHQHGTAVPPQQH 215
                     **:***..:****** :**** * ******.**|*.   ***.  ** ::::..**|.  :*

Maize_ids            QRLYSSACHGFFP--NHQVQERPVERRPELGAQPFPSWAWQAQGSP----HVPLHHSAAS 400
Barley_AP2-like      QRLYPSACHGFYPNVQVQVQERPLEPRPPE-PSSFPGWGWHAQAVPPGSSHSPLLYAAAS 408
Tmap2                QRLYPSACHGFYPNVQVQVQERPMEARPPEQPSSFPGWGWQAQAMPPGSSHSPLLYAAAS 415
Wheat_AP2-like       QRLYPSACHGFYPNVQVQVQERPMEARPPEQPSSFPGWGWQAQAMPPGSSHSPLLYAAAS 275
                     ****.******:*  : ******:* **   ...**.*.*:**. *    * ** ::***

Maize_ids            SGFSTAAGANGGMPLPSHPPAQFPTTTNPFFFP 433
Barley_AP2-like      SGFSTAAGANPAPPAVVPRPSPPLLLPRPPDN- 440
Tmap2                SGFSTAAAGANLAPPPPYPDHHRFYFPRPPDN- 447
Wheat_AP2-like       SGFSTAAAGANLAPPPPYPDHHRFYFPRPPDN- 307
                     *******..    *            ..*
```

FIGURE 4.—Comparison between the maize *ids* gene (GenBank accession no. AF-048900), barley AP2-like gene (GenBank accession no. AY-069953), *T. monococcum O11-Tmap2* (GenBank accession no. AY170867), and *T. aestivum AP2*-like EST (GenBank accession no. BG313955). Amino acids shared by all four predicted proteins are indicated by asterisks in the consensus. Amino acids in boldface type represent the AP2 domain. Amino acids that differ between *T. aestivum* and *T. monococcum* are in boxes.

if the average physical-to-genetic distance ratio within the deleted region is ~330 kb/cM, and the distal breakpoint of fndel-143 occurs half way between *XksuP16* and *Xabg391*, then we could estimate that the deleted segment accounts for ~2 Mb. We are currently performing experiments to conduct detailed analysis of the sizes and distribution of fast neutron-induced deletions in wheat (B. GILL, B. FRIEBE and J. FARIS, unpublished results).

DISCUSSION

Chromosome walking in wheat: Positional cloning of genes in wheat has long been considered infeasible. But recent evidence regarding the distribution of genes and recombination along wheat chromosomes suggests that most of the genes in wheat should be amenable to positional cloning (WERNER *et al.* 1992; GILL *et al.* 1996a,b; FEUILLET and KELLER 1999; FARIS *et al.* 2000; WENG *et al.* 2000; SANDHU *et al.* 2001). STEIN *et al.* (2000) demonstrated that subgenome chromosome walking using a *T. monococcum* BAC library can be used to successfully construct contigs across the A genome chromosomes of hexaploid wheat. Furthermore, they showed that, due to the large percentage of repetitive sequences in the wheat genome, subcloning of BAC ends was not useful, and a much higher efficiency of obtaining low-copy fragments for mapping was obtained through low-pass sequencing and subsequent comparisons of homology to database sequences.

Here, we undertook a similar strategy to assemble a BAC contig spanning the *Q* locus, except that we sequenced selected BACs completely instead of conducting low-pass sequencing. An advantage of low-pass sequencing as opposed to complete sequencing is that it is relatively inexpensive while still allowing the efficient identification of low-copy fragments. Full sequencing of BACs is more expensive, but it allowed us to orient BACs along the chromosome and provided information as to which region(s) of the BACs to initiate the search for

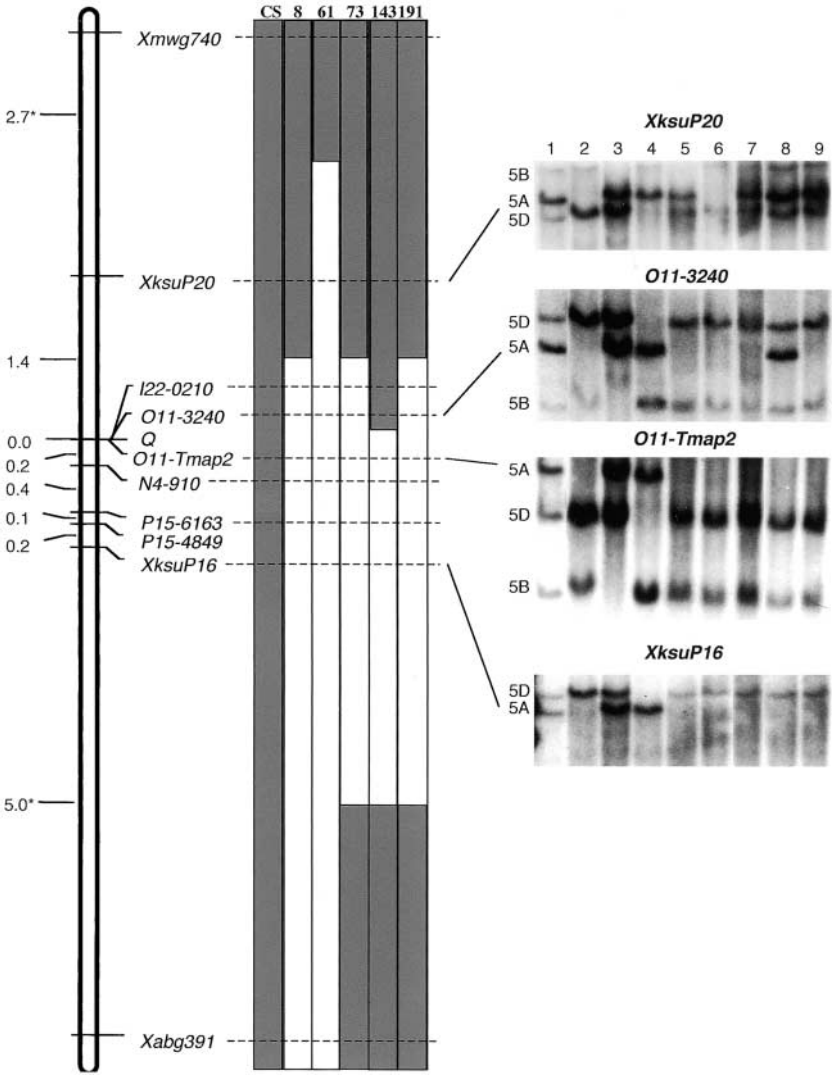

FIGURE 5.—Molecular analysis of deletion sizes of five fast neutron-induced speltoid mutants. The genetic map of the region containing *Q* is to the left. The estimated breakpoints of deletions based on molecular marker analysis are depicted to the right of the genetic map. Shaded and open bars indicate presence and absence of chromatin, respectively. On the right, autoradiograms of several probes used to characterize the deletions: lane 1, Chinese Spring; lane 2, N5AT5D; lane 3, N5BT5D; lane 4, N5DT5B; lane 5, fndel-8; lane 6, fndel-61; lane 7, fndel-73; lane 8, fndel-143; lane 9, fndel-191. Asterisks next to centimorgan distances on the genetic map indicate distances that are not to scale.

low-copy fragments to use as markers. In addition, it allowed us to annotate the BAC sequence and conduct detailed analysis of the genomic sequence, such as the distribution of genes and the types and distribution of repetitive elements (J. FELLERS, J. FARIS, S. BROOKS and B. GILL, unpublished data).

Sequencing of BACs circumvented the need to subclone and sequence BAC ends to use as markers. STEIN *et al.* (2000) subcloned 18 BAC ends but found that only 3 (17%) contained nonrepetitive sequences. We also subcloned several BAC ends before sequencing whole BACs and found all to contain highly repetitive sequences (data not shown).

In total, we sequenced four BACs, resulting in a 300-kb contig spanning the *Q* locus. BAC sequences consisted largely of repetitive elements, but sequence data allowed us to identify a sufficient number of low-copy sequences for mapping and extension of the contig. In total, we developed 29 probes from the 300-kb contig. Of these, 5 had significant similarities to known or putative genes in the database. The remaining 24 had no similarity to anything in the databases, including the TREP database, using BLASTn and BLASTx searches, and only 2 of the 24 were low copy, suggesting that wheat may contain a significant number of yet-unidentified repetitive sequences. This is contrary to STEIN *et al.* (2000), who reported that among the class of probes with no homology to database sequences, very few proved to be repetitive.

Colinearity between chromosomes 5A^m and 5A: STEIN *et al.* (2000) reported observing perfect colinearity between a 350-kb *T. monococcum* BAC contig and a region on the short arm of wheat chromosome 1A. We too observed perfect colinearity between our 300-kb *T. monococcum* BAC contig and the region encompassing the *Q* locus on wheat chromosome 5AL. A total of five probes resolved by physical distance on the *T. monococcum* BAC contig and recombination events on the wheat genetic map showed perfect colinearity. A sixth probe derived from the *T. monococcum* BAC contig was resolved on chromosome 5A of wheat by a fast neutron-induced deletion breakpoint. Therefore, six probes demonstrated perfect colinearity between the investigated region of *T. monococcum* chromosome 5A^m and wheat chro-

mosome 5A. This agrees with the notion that only very minor rearrangements between the A^m genome of *T. monococcum* and the A genome of common wheat have occurred. It also confirms the data presented by Dubcovsky *et al.* (1995) that suggested that the chromosomes of *T. monococcum* are essentially colinear with the A-genome chromosomes of common wheat.

Taken together, the results of this work and that of Stein *et al.* (2000) suggest that *T. monococcum* is an excellent model for the A genome of common wheat. The BAC library of *T. monococcum* should be useful for assembling contigs and deriving markers for map-based cloning of agronomically important genes within gene-rich regions on the A-genome chromosomes of common wheat.

Recombination frequency within the 300-kb contig: Our population size of 465 F_2 individuals representing 930 gametes provides a mapping resolution of 0.1 cM. This is a relatively small population size compared to those used in other map-based cloning experiments in plants. For the map-based cloning of *Mla*, *mlo*, and *Rpg1* genes from barley, the numbers of gametes surveyed were 3600, 4044, and 8518, respectively (Buschges *et al.* 1997; Wei *et al.* 1999; Brueggeman *et al.* 2002). Stein *et al.* (2000) surveyed 6240 gametes to construct a *T. monococcum* BAC contig spanning the *Lr10* leaf rust resistance locus in hexaploid wheat. However, Spielmeyer *et al.* (2000) showed that an *Ae. tauschii* F_2 population of 58 individuals was enough to identify recombinants between most of the mapped markers on 1DS. In the genomic region containing the seed storage proteins, a 110-kb BAC clone that contained three markers and spanned 5 cM was identified, and in the distal adjacent region, two other BAC clones similar in size also contained RFLP markers that were separated by recombination. Therefore, the size of the mapping population required to observe the necessary recombination rate depends on the target region. Durrett *et al.* (2002) presented a simple formula for estimating the number of gametes required for observation for positional cloning, but prior knowledge, or at least an estimate, of the physical-to-genetic distance ratio within the target is required.

Although our population provided a relatively low mapping resolution, recombination was variable across the 300-kb contig. Also, the double crossover singletons observed in one plant each for the markers *P15-6163* and *N4-910* tended to have drastic effects on the map distances. The 300-kb contig spanned 0.9 cM, which results in an overall estimate of a physical-to-genetic distance of ~330 kb/cM. But within the region the estimates range from 130 kb/cM between *P15-4849* and *P15-6163* to 600 kb/cM between *N4-910* and *I22-0210*.

Mechanisms of recombination suppression and activation relative to wheat × ssp. *dicoccoides* populations were reviewed in Faris *et al.* (2000). They found recombination to be somewhat suppressed in a gene-rich region on chromosome 5B in wheat × ssp. *dicoccoides* and durum × ssp. *dicoccoides* populations. Suppressed recombination was not observed in the populations used for this study (Faris and Gill 2002), which involves chromosome 5A of wheat × ssp. *dicoccoides*. On the contrary, the identification of the two singleton double crossovers suggests that in wheat microrecombination hot spots may exist within the recombination hot spots observed at the level of whole chromosome linkage maps. The microrecombination hot spots would be difficult to detect on whole chromosome linkage maps because they tend to lack the necessary resolution and are prone to errors incorporated by statistical linkage analysis programs. But construction and analysis of local BAC contigs anchored to linkage maps provide high resolution at the sequence level. As more local BAC contigs are assembled and anchored in wheat, more sequences containing microrecombination hot spots may be identified.

Another reason that double crossover singletons are not observed more often might be that most map-based gene isolation studies use only the individuals having apparent recombination between the most closely linked markers and the target gene for mapping and extension of the walk. The time, labor, and expense involved with mapping every marker on every individual within a very large mapping population, especially when the goal is to isolate a targeted gene, does not make it practical to do so.

We intend to investigate further the sequences of the two plants containing the double crossover singleton events and to identify the point of recombination. In other research, we found a similar case in which two recombination events occurred within 1350 bp in the same plant (L. Huang, S. Brooks, J. Fellers and B. Gill, unpublished results). In the two cases presented in this research, probes P15-6163 and N4-910 both represent putative genes. There is much evidence to support the notion that recombination hot spots occur within or near genes (reviewed by Schnable *et al.* 1998). Intragenic recombination frequencies may be influenced by various factors including *trans*-acting factors (Timmermans *et al.* 1997), transposon insertions (Dooner 1986; Xu *et al.* 1995), and the presence of small base-pair heterologies between allelic combinations (Borts and Haber 1989; Dooner and Martinez-Ferez 1997). In yeast, specific short DNA sequences required for recombination hot-spot activity have been identified (reviewed by Smith 1994).

***APETALA2* is a candidate gene for *Q*:** A predicted 3.5-kb open reading frame (ORF) near the middle of BAC 594O11 had a high degree of similarity to the maize *ids1* gene, the *AP2* gene from Arabidopsis, and other similar proteins. *AP2* is a floral homeotic gene that is distinguished by a plant-specific DNA-binding motif referred to as the AP2 domain. The *AP2* family of transcription factors has been implicated in a wide range of plant development roles. In Arabidopsis, *AP2* controls

the establishment of floral meristem identity (Irish and Sussex 1990; Bowman *et al.* 1993), the specification of floral organ identity (Komaki *et al.* 1988; Bowman *et al.* 1989; Kunst *et al.* 1989; Jofuku *et al.* 1994), and the temporal and spatial regulation of flower homeotic gene expression (Drews *et al.* 1991). Jofuku *et al.* (1994) suggested that *AP2* is genetically upstream of the major flower-specific homeotic genes that regulate Arabidopsis flower development. Okamura *et al.* (1997) defined a large family of Arabidopsis genes containing the AP2 domain and showed that the expression of many of these genes was regulated by *AP2*.

In maize, the genes *ids1* and *glossy15* have also been found to contain the highly conserved AP2 domain. The *ids1* gene governs the number of floral meristems produced (Chuck *et al.* 1998), while *glossy15* functions to repress adult leaf characters in juvenile plants (Moose and Sisco 1997).

It has long been thought that *Q* is a major regulatory gene influencing inflorescence architecture in addition to having pleiotropic effects on other traits. The various traits that *Q* has been shown to influence include the free-threshing habit, glume keeledness, rachis toughness, spike type, spike length, and culm height (MacKey 1954; Muramatsu 1963, 1986; Singh 1969; Kato *et al.* 1999). The degree to which *Q* influences these characters is dependent on the genetic modifiers at other loci, and more profound effects of modifiers were observed in tetraploid wheats compared to hexaploids presumably due to the lack of the D genome in tetraploids (Muramatsu 1986).

Molecular genetic experiments have also provided some evidence that *Q* is a regulatory gene. Faris and Gill (2002) and Kojima *et al.* (2000) conducted mRNA differential display and cDNA-amplified fragment length polymorphism analysis, respectively, comparing lines harboring chromosome deletions for the *Q* locus. Both experiments resulted in the identification of fragments with increased expression in the lines containing *Q* compared to the lines lacking it, and many of the fragments mapped to loci on chromosomes other than 5A.

It therefore seems likely that *Q*, much like *AP2* in Arabidopsis, is situated genetically upstream of other genes involved in determining floral-related characters in wheat. *T. monococcum* is non-free-threshing and considered to possess the *q* allele. The fact that the *AP2*-like gene in *T. monococcum* contains some sequence differences compared to other *AP2*-like genes cloned thus far, including the EST from the free-threshing (*QQ*) hexaploid wheat, suggests it may be reduced, or altered, in function. It may be that this putative reduction in function is responsible for the non-free-threshing speltoid (*qq*) phenotype in *T. monococcum*.

Although much evidence suggests that the *AP2*-like gene is *Q*, including the observation of complete colinearity between *T. monococcum* chromosome $5A^m$ and *T. aestivum* chromosome 5A, we cannot rule out the possibility that *T. monococcum* has a null allele of *q*. It is possible that a small ORF of several kilobases could exist near the *AP2*-like gene in *T. aestivum* but not in *T. monococcum*. In this case, the ORF would be lacking in the deletion mutants along with the *AP2* sequence. Expression analysis, sequencing, and subsequent transformation of the *AP2*-like gene from free-threshing wheat will shed much light on the notion that the *AP2*-like gene in wheat is actually *Q*.

Fast neutron mutagenesis: We identified 26 fast neutron-induced speltoid mutants out of 2600 plants analyzed, for a mutation rate of 1% at the *Q* locus. Singh (1969) reported a rate of 0.5% mutations at the *Q* locus using fast neutrons. He also conducted mutagenesis experiments using X rays, isotopes (^{32}P and ^{35}S), and gamma rays and found speltoid mutants occurring at rates ranging from 0.1 to 7.6%. Therefore, it seems that *Q* lies within a highly mutable locus. As Faris *et al.* (2000) discussed, it may be that the chromatin structure at recombination hot spots is such that it is highly accessible to recombination machinery. If this tends to be the case, then it seems likely that such chromatin regions would also be prone to breakage.

It is difficult to precisely estimate the sizes of the fast neutron-induced chromosome deletions involving the *Q* locus. The largest deletions span >70 cM of genetic distance, and the smallest span <6 cM, which may correspond to ~2 Mb. Li *et al.* (2001) described a reverse genetics method to identify and isolate fast neutron-induced deletions in Arabidopsis using a PCR-based approach. The sizes of 36 deletions characterized ranged from 0.8 to 12 kb with most of them <6 kb, and a 2.5-kb deletion was observed in a rice target gene. Therefore, the deletion sizes observed in Arabidopsis and rice were much smaller than those analyzed in this study. As a polyploid, wheat has a very high buffering capacity and can tolerate large chromosome deletions and aberrations. Such large deletions would likely be lethal in diploid species such as Arabidopsis and rice. Presumably, screening of much larger mutagenized wheat populations would result in the identification of deletions similar in size to those observed in Arabidopsis.

The large sizes of the *Q* locus deletions observed in this study prohibited the provision of proof that the *AP2*-like gene is actually *Q*. Analysis of deletion mutants alone narrows the candidate region to a segment slightly larger than our BAC contig. However, one of the deletion breakpoints occurred between our candidate gene (*O11-Tmap2*) and *O11-3240*, which were not separated by recombination events. Thus, analysis of the deletion mutants in conjunction with recombination data provided proof that the BAC contig does in fact span the *Q* locus.

Together, these data define the region containing *Q* as a segment <100 kb. The majority of the 100-kb segment consists of repetitive elements, and no other

known genes besides the *AP2*-like gene are present. This provides further evidence that the *AP2*-like gene may actually be *Q*.

The *AP2*-like gene proves to be a very promising candidate gene for *Q*, and it will be the immediate focus of future research. The isolation of *Q* and subsequent analysis will provide many insights regarding floral homeotic gene regulation and expression in wheat and knowledge of polyploid gene regulation in general. Studies can be done to determine if *Q* arose first in hexaploid *T. aestivum* or in the tetraploid *T. turgidum* or if it arose in both species independently. Furthermore, the debate as to whether *Q* arose as a modification or a duplication of *q*, or if *q* arose from *Q*, can be addressed. Finally, *Q* may be used in a practical sense to render synthetic amphiploids free threshing. Ultimately, the *Q* gene may be used essentially to domesticate wild grasses as new crops that could be cultivated for specific purposes.

The authors thank Duane Wilson, David Birdsall, Angie Matthews, and Erik Doehler for technical assistance. This is contribution no. 03-147-J from the Kansas Agricultural Experiment Station.

LITERATURE CITED

Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang *et al.*, 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:** 3389–3402.

Borts, R. H., and J. E. Haber, 1989 Length and distribution of meiotic gene conversion tracts and crossovers in *Saccharomyces cerevisiae*. Genetics **123:** 69–80.

Bowman, J. L., D. R. Smyth and E. M. Meyerowitz, 1989 Genes directing flower development in *Arabidopsis*. Plant Cell **1:** 37–52.

Bowman, J. L., J. Alvarez, D. Weigel, E. M. Meyerowitz and D. R. Smyth, 1993 Control of flower development in *Arabidopsis thaliana* by *APETALA1* and interacting genes. Development **119:** 721–743.

Brooks, S. A., L. Huang, B. S. Gill and J. P. Fellers, 2002 Analysis of 106 kb of contiguous DNA sequence from the D genome of wheat reveals high gene density and a complex arrangement of genes related to disease resistance. Genome **45:** 963–972.

Brueggeman, R., N. Rostoks, D. Kudrna, A. Kilian, F. Han *et al.*, 2002 The barley stem rust-resistance gene *Rpg1* is a novel disease resistance gene with homology to receptor kinases. Proc. Natl. Acad. Sci. USA **99:** 9328–9333.

Buschges, R., K. Hollricher, R. Panstruga, G. Simons, M. Wolter *et al.*, 1997 The barley mlo *Mlo* gene: a novel control element of plant pathogen resistance. Cell **88:** 695–705.

Chuck, G., R. B. Meeley and S. Hake, 1998 The control of maize spikelet meristem fate by the *APETALA2*-like gene *indeterminate spikelet1*. Genes Dev. **12:** 1145–1154.

Dooner, H. K., 1986 Genetic fine structure of the *Bronze* locus in maize. Genetics **113:** 1021–1036.

Dooner, H. K., and I. M. Martinez-Ferez, 1997 Recombination occurs uniformly within the *bronze* gene, a meiotic recombination hotspot in the maize genome. Plant Cell **9:** 1633–1646.

Drews, G. N., J. L. Bowman and E. M. Meyerowitz, 1991 Negative regulation of the Arabidopsis homeotic gene *AGAMOUS* by the *APETALA2* product. Cell **65:** 991–1002.

Dubcovsky, J., M. Luo and J. Dvorak, 1995 Differentiation between homoeologous chromosomes 1A of wheat and 1Am of *Triticum monococcum* and its recognition by the wheat *Ph1* locus. Proc. Natl. Acad. Sci. USA **92:** 6645–6649.

Durrett, R. T., K. Y. Chen and S. D. Tanksley, 2002 A simple formula useful for positional cloning. Genetics **160:** 353–355.

Endo, T. R., and B. S. Gill, 1996 The deletion stocks of common wheat. J. Hered. **87:** 295–307.

Endo, T. R., and Y. Mukai, 1988 Chromosome mapping of speltoid suppression gene of *Triticum aestivum* L. based on partial deletion in the long arm of chromosome 5A. Jpn. J. Genet. **63:** 501–506.

Ewing, B., and P. Green, 1998 Base-calling of automated sequencer traces using *Phred*. II. Error probabilities. Genome Res. **8:** 186–194.

Ewing, B., L. Hillier, M. C. Wendl and P. Green, 1998 Base-calling of automated sequencer traces using *Phred*. I. Accuracy assessment. Genome Res. **8:** 175–185.

Faris, J. D., and B. S. Gill, 2002 Genomic targeting and high-resolution mapping of the domestication gene *Q* in wheat. Genome **45:** 706–718.

Faris, J. D., K. M. Haen and B. S. Gill, 2000 Saturation mapping of a gene-rich recombination hot spot in wheat. Genetics **154:** 823–835.

Feuillet, C., and B. Keller, 1999 High gene density is conserved at syntenic loci of small and large grass genomes. Proc. Natl. Acad. Sci. USA **96:** 8265–8270.

Gill, K. S., B. S. Gill, T. R. Endo and E. V. Boiko, 1996a Identification and high-density mapping of gene-rich regions in chromosome group 5 of wheat. Genetics **143:** 1001–1012.

Gill, K. S., B. S. Gill, T. R. Endo and T. Taylor, 1996b Identification and high-density mapping of gene-rich regions in chromosome group 1 of wheat. Genetics **144:** 1883–1891.

Gordon, D., C. Abajian and P. Green, 1998 *Consed*: a graphical tool for sequence finishing. Genome Res. **8:** 195–202.

Huang, S., A. Sirikhachornkit, X. Su, J. Faris, B. Gill *et al.*, 2002 Genes encoding acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the Triticum/Aegilops complex and the evolutionary history of polyploid wheat. Proc. Natl. Acad. Sci. USA **99:** 8133–8138.

Huskins, C. L., 1946 Fatuoid, speltoid and related mutations of oats and wheat. Bot. Rev. **12:** 457–514.

Irish, V. F., and I. M. Sussex, 1990 Function of the *apetala-1* gene during *Arabidopsis* floral development. Plant Cell **2:** 741–753.

Jofuku, K. D., B. G. W. den Boer, M. Van Montagu and J. K. Okamuro, 1994 Control of Arabidopsis flower and seed development by the homeotic gene *APETALA2*. Plant Cell **6:** 1211–1225.

Kabarity, A., 1966 On the origin of the new cultivated wheat II. Cytogenetical studies on the karyotypes of some *Triticum macha* varieties. Beitr. Biol. Pfanzen **42:** 339–346.

Kato, K., H. Miura and S. Sawada, 1999 QTL mapping of genes controlling ear emergence time and plant height on chromosome 5A of wheat. Theor. Appl. Genet. **98:** 472–477.

Kojima, T., and Y. Ogihara, 1998 High-resolution RFLP map of the long arm of chromosome 5A in wheats and its synteny among cereals. Genes Genet. Syst. **73:** 51–58.

Kojima, T., Y. Habu, S. Iida and Y. Ogihara, 2000 Direct isolation of differentially expressed genes from a specific chromosome region of common wheat: application of the amplified fragment length polymorphism-based mRNA fingerprinting (AMF) method in combination with a deletion line of wheat. Mol. Gen. Genet. **263:** 635–641.

Komaki, M. K., K. Okada, E. Nishino and Y. Shimura, 1988 Isolation and characterization of novel mutants of *Arabidopsis thaliana* defective in flower development. Development **104:** 195–203.

Kosambi, D. D., 1944 The estimation of map distances from recombination values. Ann. Eugen. **12:** 172–175.

Kunst, L., J. E. Klenz, J. Martinez-Zapater and G. W. Haughn, 1989 *AP2* gene determines the identity of perianth organs in flowers of *Arabidopsis thaliana*. Plant Cell **1:** 1195–1208.

Lander, E. S., P. Green, J. Abrahamson, A. Barlow, M. J. Daly *et al.*, 1987 MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics **1:** 174–181.

Leighty, C. E., and S. Boshnakian, 1921 Genetic behaviour of the spelt form in crosses between *Triticum spelta* and *Triticum aestivum*. J. Agric. Res. **7:** 335–364.

Li, X., Y. Song, K. Century, S. Straight, P. Ronald *et al.*, 2001 A fast neutron deletion mutagenesis-based reverse genetics system for plants. Plant J. **27:** 235–242.

Lijavetsky, D., G. Muzzi, T. Wicker, B. Keller, R. Wing *et al.*, 1999 Construction and characterization of a bacterial artificial

chromosome (BAC) library for the A genome of wheat. Genome **42:** 1176–1182.

MacKey, J., 1954 Neutron and X-ray experiments in wheat and a revision of the speltoid problem. Hereditas **40:** 65–180.

McFadden, E. S., and E. R. Sears, 1946 The origin of *Triticum speltoides* and its free-threshing hexaploid relatives. J. Hered. **37:** 81–89, 107–116.

Miller, T. E., and S. M. Reader, 1982 A major deletion of part of chromosome 5A of *Triticum aestivum*. Wheat Inf. Serv. **88:** 10–12.

Moose, S., and P. Sisco, 1997 Glossy15, an APETALA2-like gene from maize that regulates leaf epidermal cell identity. Genes Dev. **10:** 3018–3027.

Muramatsu, M., 1963 Dosage effect of the spelta gene q of hexaploid wheat. Genetics **48:** 469–482.

Muramatsu, M., 1986 The *vulgare* super gene, *Q*: its universality in *durum* wheat and its phenotypic effects in tetraploid and hexaploid wheats. Can. J. Genet. Cytol. **28:** 30–41.

Ogihara, Y., K. Hasegawa and H. Tsujimoto, 1994 High-resolution cytological mapping of the long arm of chromosome 5A in common wheat using a series of deletion lines induced by gametocidal (Gc) genes of *Aegilops speltoides*. Mol. Gen. Genet. **244:** 253–259.

Okamura, J. K., B. Caster, R. Villarroel, M. Van Montagu and K. D. Jofuku, 1997 The AP2 domain of *APETALA2* defines a large new family of DNA binding proteins in *Arabidopsis*. Proc. Natl. Acad. Sci. USA **94:** 7076–7081.

Paterson, A. H., Y. R. Lin, Z. Li, K. F. Schertz, J. W. Doebley *et al.*, 1995 Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. Science **269:** 1714–1715.

Sandhu, D., J. A. Champoux, S. N. Bondareva and K. S. Gill, 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.

Schnable, P. S., A.-P. Hsia and B. J. Nikolau, 1998 Genetic recombination in plants. Curr. Opin. Plant Biol. **1:** 123–129.

Sears, E. R., 1952 Misdivision of univalents in common wheat. Chromosoma **4:** 535–550.

Sears, E. R., 1954 The aneuploids of common wheat. MO Agr. Exp. Sta. Res. Bull. **572:** 1–59.

Singh, M. P., 1969 Some radiation induced changes at the 'Q' locus in bread wheat. Caryologia **22:** 119–126.

Singh, H. B., E. Anderson and B. P. Pal, 1957 Studies in the genetics of *Triticum vavilovii* Jackub. Agron. J. **49:** 4–11.

Smith, G. R., 1994 Hotspots of homologous recombination. Experientia **50:** 234–242.

Spielmeyer, W., O. Moullet, A. Laroche and E. Lagudah, 2000 Highly recombinogenic regions at seed storage protein loci on chromosome 1DS of *Aegilops tauschii*, the D-genome donor of wheat. Genetics **155:** 361–367.

Stein, N., C. Feuillet, T. Wicker, E. Schlagenhauf and B. Keller, 2000 Subgenome chromosome walking in wheat: a 450-kb physical contig in *Triticum monococcum* L. spans the *Lr10* resistance locus in hexaploid wheat (*Triticum aestivum* L.). Proc. Natl. Acad. Sci. USA **97:** 13436–13441.

Timmermans, M. C. P., O. P. Das, J. M. Bradeen and J. Messing, 1997 Region-specific *cis*- and *trans*-acting factors contributing to genetic variability in meiotic recombination in maize. Genetics **146:** 1101–1113.

Tsujimoto, H., and K. Noda, 1989 Structure of chromosome 5A of wheat speltoid mutants induced by the gametocidal genes of *Aegilops speltoides*. Genome **32:** 1085–1090.

Tsujimoto, H., and K. Noda, 1990 Deletion mapping by gametocidal genes in common wheat: position of speltoid suppression (*Q*) and β-amylase (β-*Amy-A2*) genes on chromosome 5A. Genome **33:** 850–853.

Unrau, J., W. E. Smith and R. C. McGinnis, 1950 Spike density, speltoidy and compactoidy in hexaploid wheat. Can. J. Res. Commun. **28:** 273–276.

Wei, F., K. Gobelman-Werner, S. M. Morroll, J. Kurth, L. Mao *et al.*, 1999 The *Mla* (powdery mildew) resistance cluster is associated with three NBS-LRR gene families and suppressed recombination within a 240-kb DNA interval on chromosome 5S (1HS) of barley. Genetics **153:** 1929–1943.

Weng, Y., N. A. Tuleen and G. E. Hart, 2000 Extended physical maps and a consensus physical map of the homoeologous group-6 chromosomes of wheat (*Triticum aestivum* L. em Thell.). Theor. Appl. Genet. **100:** 519–527.

Werner, J. E., T. R. Endo and B. S. Gill, 1992 Towards a cytogenetically based physical map of the wheat genome. Proc. Natl. Acad. Sci. USA **89:** 11307–11311.

Xu, X., A.-P. Hsia, L. Zhang, B. J. Nikolau and P. S. Schnable, 1995 Meiotic recombination break points resolve at high rates at the 5′ end of a maize coding sequence. Plant Cell **7:** 2151–2161.

Communicating editor: A. Paterson