



## AFLP-derived SCARs facilitate construction of a 1.1 Mb sequence-ready map of a region that spans the *Vf* locus in the apple genome

Mingliang Xu and Schuyler S. Korban\*

Department of Natural Resources and Environmental Sciences, University of Illinois, 310 Madigan Building, 1201 West Gregory Drive, Urbana, IL 61801, USA (\*author for correspondence; e-mail: s-korban@uiuc.edu)

Received 28 December 2001; accepted in final form 15 May 2002

**Key words:** apple scab, bacterial artificial chromosome (BAC), contig, amplified fragment length fragment (AFLP), sequence characterized amplified region (SCAR), *Vf* gene

### Abstract

The availability of high-density anchored markers is a prerequisite for reliable construction of a deep coverage BAC contig, which leads to creation of a sequence-ready map in the target chromosomal region. Unfortunately, such markers are not available for most plant species, including woody perennial plants. Here, we report on an efficient approach to build a megabase-size sequence-ready map in the apple genome for the *Vf* region containing apple scab resistance gene(s) by targeting AFLP-derived SCAR markers to this specific genomic region. A total of 11 AFLP-derived SCAR markers, previously tagged to the *Vf* locus, along with three other *Vf*-linked SCAR markers have been used to screen two apple genome BAC libraries. A single BAC contig which spans the *Vf* region at a physical distance of approximately 1,100 kb has been constructed by assembling the recovered BAC clones, followed by closure of inter-contig gaps. The contig is  $\sim 4 \times$  deep, and provides a minimal tiling path of 16 contiguous and overlapping BAC clones, thus generating a sequence-ready map. Within the *Vf* region, duplication events have occurred frequently, and the *Vf* locus is restricted to the ca. 290 kb region covered by a minimum of three overlapping BAC clones.

### Introduction

Construction of a megabase-size sequence-ready bacterial artificial chromosome (BAC) or yeast artificial chromosome (YAC) contig for a defined chromosomal region plays a key role in map-based gene cloning, physical map development, genetic structure characterization, and initiation of sequencing projects. Initially, contig construction of a target region is heavily dependent on chromosome walking (Nakamura *et al.*, 1997; Niederfuhr *et al.*, 1998; Patocchi *et al.*, 1999; Schmidt *et al.*, 1995; Stein *et al.*, 2000). However, this method is a very tedious and labor-intensive effort, and it is frequently hampered by the large amount of DNA traversed as well as by the presence of repetitive DNA sequences (Tanksley *et al.*, 1995). Alternatively, high-density markers (RFLPs or PCR-based markers) mapped within a target chromosomal region have been introduced in building BAC or YAC contigs (Cao

*et al.*, 1999; Marra *et al.*, 1997; Soderlund *et al.*, 2000; Vollrath and Jaramillo-Babb 1999; Zhu *et al.*, 1999). With high-density anchored markers, most positive clones can be recovered from large-insert DNA libraries. The marker contents and restriction fingerprints of selected clones are then used to build initial contigs for the target region. Inter-contig gaps, if present, can be readily closed with a few steps of chromosome walking, thus resulting in megabase-size contigs and creation of sequence-ready maps (Cao *et al.*, 1999; Vollrath and Jaramillo-Babb 1999).

In plant species, high-density markers are only available in a few model plants, such as *Arabidopsis*, tomato, and rice (Haanstra *et al.*, 1999; Lukowitz *et al.*, 2000; Temnykh *et al.*, 2000). Even in these model plants, additional markers are frequently required to saturate a target region in an attempt to construct contigs. For the vast majority of plant species, including woody perennial plants, the crucial step

prior to any map-based gene cloning or genomic analysis effort is the generation of high-density markers. This however, remains a daunting task (Tanksley *et al.*, 1995; Zhang and Wing 1997).

The apple, *Malus × domestica* Borkh., has become a model woody perennial angiosperm for genomic research due to its economic importance and its relatively small genome size (769 Mb/haploid) (Patocchi *et al.*, 1999). Most cultivated apples are diploids ( $2n = 34$ ), self-incompatible, and display a juvenile period of 6 to 10 years or more (Korban and Chen 1992). The *Vf* gene, originating from the crabapple *Malus floribunda* 821, confers resistance to apple scab incited by the fungal pathogen *Venturia inaequalis* (Cke.) Wint. (Bénaouf and Parisi, 2000). The disease severely damages both tree and fruit, and is known as one of the most serious diseases of apples. The *Vf* gene has been introgressed into the cultivated apple using a pseudo-testcross breeding strategy whereby all scab-susceptible cultivars serve as the 'recurrent parent' during each cycle of sexual hybridization (Korban and Chen 1992). It is impossible to generate F<sub>2</sub> or F<sub>3</sub> populations or any backcross populations in an apple genetics program. Following five to seven cycles of hybridization, several apple cultivars carrying the *Vf* gene are then named and released (Crosby *et al.*, 1992). The *Vf* gene exhibits race-specific resistance by conferring resistance to races 1 to 5 of *V. inaequalis*, but not to two newly-discovered races (race 6 identified in Germany and race 7 identified in France) (Parisi *et al.*, 1993; Parisi and Lespinasse 1996; Bénaouf and Parisi, 2000). In addition, minor or modifying genes can influence expression of the *Vf* gene resulting in a range of symptoms in *Vf*-containing plants (Gessler 1989). In the past decade, several efforts have been undertaken to tag the *Vf* gene with biochemical and molecular markers (Bournival and Korban 1987; Manganaris *et al.*, 1994; Yang and Korban 1996; Tartarini 1996; Tartarini *et al.*, 1999; Gardiner *et al.*, 1996; Gianfranceschi *et al.*, 1996; Yang *et al.*, 1997 a,b; Hemmat *et al.*, 1998). The closest markers flanking the *Vf* gene, thus far, have been OPM18 (left of the *Vf*) and OPAM19/OPAL07 (right of the *Vf*) with a physical distance of approximately 550 kb (Patocchi *et al.*, 1999). These three markers have been already used as starting points for chromosomal walking to construct a BAC contig spanning the *Vf* locus (Patocchi *et al.*, 1999).

In our effort toward map-based cloning of the *Vf* gene and structural analysis of the *Vf* region, 15 AFLP markers have been generated to tag the *Vf* gene

with an average distance of one marker every 40 kb (Xu and Korban 2000). Of these 15 AFLP markers, 11 have been successfully converted into SCAR markers (Xu *et al.*, 2001a). In addition, two BAC libraries have been constructed from *M. floribunda* 821 and 'GoldRush' (a *Vf*-containing apple cultivar) (Xu *et al.*, 2001b,c), respectively. Each library represents approximately five haploid-genome equivalents.

The objective of the present study is to assess the feasibility of constructing a BAC contig for the *Vf* region using AFLP-derived SCAR markers. The strategy described in this study, therefore, can expand the utility of AFLP-derived SCARs in building BAC or YAC contigs for those chromosomal regions containing a gene(s) of interest, but are poorly covered with molecular markers.

## Materials and methods

### *BAC libraries, SCAR markers and library screening*

Two BAC libraries have been constructed for identifying BAC clones derived from the *Vf* region. One library is constructed from *M. floribunda* 821, the original source of the *Vf* gene, and the other from a *Vf*-containing apple cultivar, 'GoldRush' (Xu *et al.*, 2001b,c). Each library represents approximately 5 × haploid-genome equivalents.

A total of 14 *Vf*-linked SCARs, including 11 AFLP-derived SCAR markers (ACS-1 to -11) and three RAPD-derived SCAR markers (OPM18, OPAM19, and OPAL07), have been used to screen the two BAC libraries. All SCAR markers, except for OPM18, reside *in cis* along the *Vf* region, and thus can only identify BAC clones derived from the *Vf* region, but not from its homologous region. OPM18 can detect BAC clones from both the *Vf* and its homologous regions. A genetic linkage map of the *Vf* region has been previously developed (Xu *et al.*, 2001a). This map shows that three SCAR markers (ACS-3, -7, and -9) are inseparable from *Vf*, two (ACS-6 and OPM18) are located left, and nine (ACS-1, -2, -4, -5, -8, -10, -11, OPAM19, and OPAL07) are located right of the *Vf* locus at genetic distances of 0.4 cM and 0.2 cM, respectively.

A three-step PCR-based screening procedure was adopted for identifying positive BAC clones from the two BAC libraries as previously described (Xu *et al.*, 2001b,c). The PCR reaction for SCAR markers was as follows: 35 cycles of denaturing at 94 °C for 30 s,

annealing at 65 °C for 30 s, and extension at 72 °C for 1 min.

#### *Estimation of insert sizes of BAC clones*

BAC clones were grown overnight at 37 °C in a 2.0 ml LB medium containing chloramphenicol (12.5 µg/ml). The BAC DNA was extracted using a mini-alkaline lysis procedure (Sambrook *et al.*, 1989). One-tenth of the volume of a DNA sample was digested with *NotI* to recover BAC inserts. The insert size was estimated against a λ DNA ladder after size-fractionation on a clamped homogeneous electrical field (CHEF) electrophoresis apparatus (BIO-RAD) at 5.7 V/cm for 15 h with a 5- to 15-s pulse time at a 120° angle.

#### *End-labeled double-digest fingerprinting method*

DNA fingerprinting was carried out according to the method described by Klein *et al.*, (2000) with some modifications. A DNA fingerprinting cocktail solution was prepared by mixing 86 µl sterile water, 24 µl 10× multi-core buffer (Promega), 5 µl 1 mM dGTP, 5 µl 1 mM dTTP, 6 µl *Bam*HI (10 U/µl), 6 µl *Hae*III (10 U/µl), 4 µl AMV reverse transcriptase (10 U/µl), and 1 µl α-<sup>32</sup>P-dATP. Each DNA sample (30–60 ng), in a 3 µl TE buffer, was mixed with an equal volume of cocktail solution, and incubated for 2 h at 37 °C. Following this reaction, 6 µl loading buffer (98% formamide, 0.1% bromophenol blue, 0.1 xylene cyanol, 10 mM EDTA, pH8.0) was added and thoroughly mixed. The mixture was heated for 3 min at 90 °C, then promptly cooled down by placing it on ice. A 2 µl mixture of each DNA sample was subjected to electrophoresis on a 7% Long Ranger sequencing gel (FMC Bio Products) at 65 W for 2 h. The gel was dried at 80 °C for 2 h, and then exposed to an X-ray film for one week.

#### *Complete BamHI-digest fingerprinting method*

Each DNA sample (approximately 250 ng) was digested with 2 U *Bam*HI for 2 h at 37 °C. Digested DNA fragments were loaded onto 1% Ultra Pure agarose gel, and were subjected to electrophoresis at 1.5 V/cm for 20 h in a cold room (4 °C). Recovered *Bam*HI-digested DNA fragments of positive BAC clones were used to calculate the sizes of overlapping regions, and to assemble a BAC contig in the *Vf* region.

#### *End rescue of BAC inserts*

A method of direct cloning and sequencing of BAC insert ends was adopted with some modifications (Chen and Gmitter, 1999). BAC DNA was completely double-digested with *NotI/DraI* or *NotI/EcoRV*. Digested fragments were ligated to *NotI/EcoRV*-digested pBlueScriptII KS (Stratagene), and then transformed into *E. coli* competent cells by electroporation using a BioRad Gene Pulser. White colonies were then picked, and individually dissolved in 30 µl TE buffer. After heating at 100 °C for 10 min and centrifugation at 5,000 rpm for another 10 min, the DNA-containing supernatant was directly used as template DNA to check whether or not the recombinant plasmid contained the right/left insert end. Primers used to find the left insert end are BACBL (5'-GAATTTCGAGCTCGGTACC-3'), designed based on the sequence upstream of the *Bam*HI cloning site of the BAC vector pBeloBAC11 and the KS primer (5'-TCGAGGTCGACGGTATC-3') from the pBlueScriptII KS. While primers used to identify the right insert end are BACBR (5'-TGCAGGTCGACTCTAGAG-3'), designed based on the sequence downstream of the *Bam*HI cloning site of the BAC vector pBeloBAC11 and the KS primer. The PCR reaction was performed at 94 °C for 2 min followed by 35 cycles of 1 min denaturing at 94 °C, 1 min annealing at 65 °C, and 3 min elongation at 72 °C. PCR products were subjected to electrophoresis on a 1% agarose gel. Clones giving rise to amplicons using left-end search primers (BACBL and KS primers) contained putative left insert ends; while, clones giving rise to amplicons using right-end search primers (BACBR and KS primers) contained putative right insert ends. Plasmid DNAs of the selected clones were extracted using Plasmid Mini Kit (QIAGEN), and double-digested with *XhoI/Bam*HI. Digested DNA fragments served to further validate clones containing right/left insert ends. If a clone did not contain any BAC insert end, only one DNA fragment corresponding to the linear vector pBlueScriptII KS would be detected. Whereas, if a clone contained the right/left end, two or more DNA fragments would be detected, one corresponding to the linear vector pBlueScriptII KS, and the other(s) corresponding to the left/right end. Recombinant plasmids containing either left ends or right ends were sequenced using a T3 primer located on the vector pBlueScriptII KS. The end sequences were used to design new PCR-based probes for screening BAC libraries.

## Results

### *Characterization of positive BAC clones identified from the two libraries*

A collection of 32 positive BAC clones was identified from two BAC libraries by using the three-step screening procedure (Xu *et al.*, 2001b,c). The insert sizes of these BACs were estimated by CHEF electrophoresis (Table 1). Interestingly, all eight clones detected by OPM18 (left to *Vf*) had a single *NotI* site within their inserts indicating that OPM18 was very close to the *NotI* site. The recovery of positive BAC clones from both BAC libraries resulted in a high level of redundancy, which facilitated contig construction. For example, SCAR markers ACS-6, OPM18, and OPAL07 detected seven, eight, and eight positive clones, respectively (Table 1). Immediately neighboring SCARs frequently identified common BAC clones. These clones in turn served as bridging clones to connect clones detected by different SCARs, resulting in an array of contiguously overlapping BAC clones along the *Vf* region. For example, uninterrupted bridging clones were detected from ACS-6 (left to *Vf*) until ACS-8 (right to *Vf*), including three clones (*M1-F-12*, *M2-J-3*, and *M40-I-12*) for ACS-6 and OPM18, one clone (*M4-P-11*) for OPM18 and ACS-3, two clones (*M61-M-16* and *G11-J-23*) for ACS-3 and ACS-7, one clone (*G7-C-18*) for ACS-7 and ACS-9, and one clone (*G53-N-7*) for ACS-9 and ACS-8 (Table 1, Figure 1). These bridging clones led to the creation of the first array of contiguously overlapping clones starting from ACS-6 (left of *Vf*) and ending at ACS-8 (right of *Vf*), which obviously traversed the *Vf* locus. Another two sets of contiguously overlapping clones were located to the right of the *Vf* gene, covering the intervals of ACS-10/ACS-4 and ACS-11/ACS-1, respectively. No common clones were detected in the two sets of immediately neighboring markers, ACS-8/ACS-10 and ACS-4/ACS-11. Therefore, it was concluded that physical distances between ACS-8 and ACS-10 or between ACS-4 and ACS-11 were longer than those of the BAC inserts, or there were no single BAC clones available in the BAC libraries to cover both ACS-8 and ACS-10 or ACS-4 and ACS-11.

Based on our previous linkage map and the bridging clones identified in this study, the order of the 14 SCAR markers along the *Vf* region can be roughly deduced (from left to right). ACS-6 is the left-most SCAR marker, followed by OPM18; while, the three

co-segregating SCAR markers are now placed in the following order, ACS-3\_ACS-7\_ACS-9; and finally the nine SCAR markers right to *Vf* are ordered as such: ACS-8\_ACS-10\_ACS-4\_ACS-11\_ACS-2/ACS-5\_OPAM19\_OPAL07\_ACS-1 (Table 1).

### *Initial contig assembly by end-labeled double-digest fingerprinting method*

All 32 positive BAC clones were subjected to fingerprinting analysis based on end-labeled double-digestion followed by electrophoresis on a Long Ranger sequencing gel. This strategy allowed accurate comparison of fingerprints generated from different BAC clones. Two bands derived from the BAC vector pBeloBAC11 were present in all 32 BAC clones including a small-size band (57 bp, left to the *BamHI* cloning site) and a large-size band (259 bp, right to the *BamHI* cloning site) (Figure 2). Many fingerprints were shared among adjacent BAC clones indicating overlapping relationships among these BAC clones, for example, *M2-J-3* and *M40-I-12*. Some BAC clones had exactly the same fingerprinting patterns (for example, *M12-I-5* and *M60-K-11*), and they might be derived from the same chromosomal region (Figure 2). Completely different fingerprinting patterns were observed for the following three BAC clones, *G70-F-18*, *M66-E-9*, and *M81-E-10* (Figure 2). To investigate the reason(s) for this observation, the three BAC clones were re-amplified using their corresponding markers. No amplicons were observed for *G70-F-18* using its four corresponding markers: ACS-11, ACS-2, ACS-5, and OPAM19. It was obvious that contamination occurred during picking of clone *G70-F-18*. Whereas, PCR products were present for *M66-E-9* and *M81-E-10*, following amplification with their respective SCAR markers OPAM19 and OPAL07. Therefore, the possible origins of these two clones were further investigated.

In order to verify the identity of *G70-F-18* BAC clone, the original clone stored in a 384-well plate was streaked onto LB media, and colonies were individually checked using the above four corresponding markers until the true *G70-F-18* BAC clone was detected. For *M66-E-9* and *M81-E-10*, both left- and right-insert ends were isolated and sequenced. PCR-based probes were developed based on end sequences, and used to amplify all 32 positive clones. Apart from their cognate BAC clones, no other clones gave rise to amplicons. This strongly substantiated the finding that these two clones were not derived from the *Vf* region,

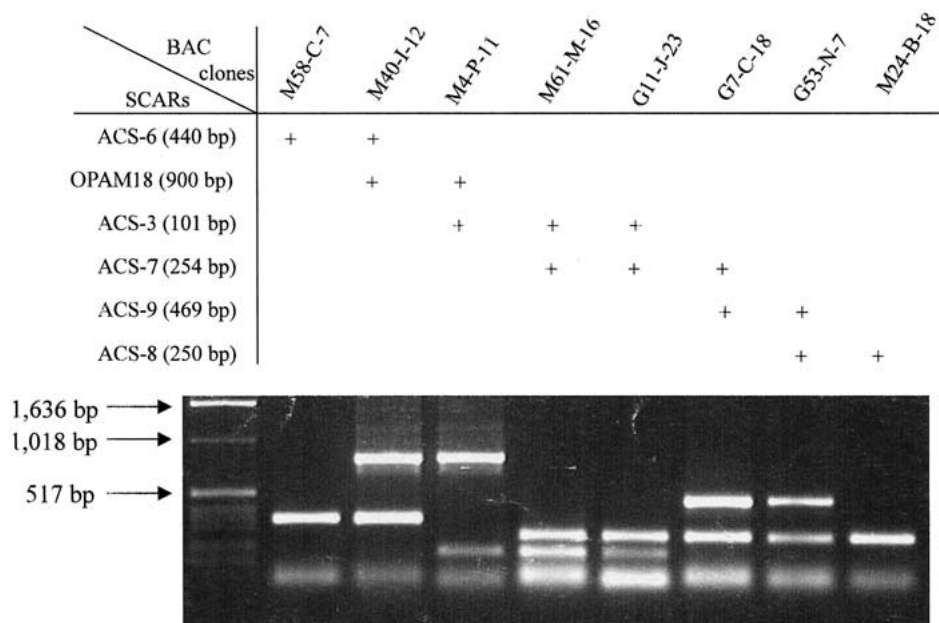


Figure 1. An array of contiguous and overlapping BAC clones was created from SCAR ACS-6 to ACS-8 by a series of bridging clones. From left to right: *M40-I-12* bridges ACS-6 and OPAM18, *M4-P-11* bridges M18 and ACS-3; *M61-M-16* and *G11-J-23* bridge ACS-3 and ACS-7; *G7-C-18* bridges ACS-7 and ACS-9; and *G53-N-7* bridges ACS-9 and ACS-8.

regardless of the fact that these clones were identified using *Vf*-linked SCARs.

Inconsistent fingerprinting patterns have been observed for the following three adjacent clones: *M12-I-5*, *M60-K-11*, and *M70-G-2*. These three clones share many common fingerprints among each other, unfortunately, almost half of the fingerprints can not be validated by their neighboring clones. It is very likely that these three BAC clones are derived from the *Vf* homologous region, rather than the *Vf* region itself, since the SCAR marker OPAM18 can amplify both the *Vf* and its homologous region.

Except for the above five non-*Vf* region derived BAC clones (*M66-E-9*, *M81-E-10*, *M12-I-5*, *M60-K-11*, and *M70-G-2*) and one contaminated clone (*G70-F-18*), the remaining 26 clones have been assembled into three BAC contigs, corresponding to the above three unambiguous arrays of overlapping BAC clones (Figure 2). The first contig consists of 14 clones, starting from ACS-6 and ending at ACS-8. The *Vf* gene is located within this contig. The second contig has only three clones spanning the ACS-10/ACS-4 interval. The third contig is composed of nine clones, bordered by ACS-11 and ACS-1. No shared fingerprints had been observed between the right-most clone in the first contig (*M24-B-18*) and the left-most clone in the second contig (*M72-E-9*). Likewise, no shared

fingerprints are present between the rightmost clone in the second contig (*M89-I-20*) and the leftmost clone in the third contig (*G20-G-18*). It is obvious that two inter-contig gaps are present among these three contigs, and must be sealed. Thus, we have pursued an end-rescue strategy for these right- or left- most BAC clones to initiate chromosome walking to close these two gaps.

#### *End-rescue to close inter-contig gaps and elongation of the contig*

In the first gap, insert ends of the two border clones (*M24-B-18* of the first contig and *M72-E-9* of the second contig) were isolated by end-rescue and then sequenced. Accordingly, four new PCR-based probes were developed based on both left- and right-end sequences of the two clones (Table 2). These new probes were tested against clones from the first and second BAC contigs (Table 3). The presence/absence of PCR products on tested BAC clones involved in these two contigs allowed us to determine the origin of BAC clones and the orientation of each probe. The probe located opposite to the gap must amplify the cognate BAC clone and its overlapping BAC clones; whereas, the probe located toward the gap can amplify the cognate BAC clone, but not the overlapping clones. For

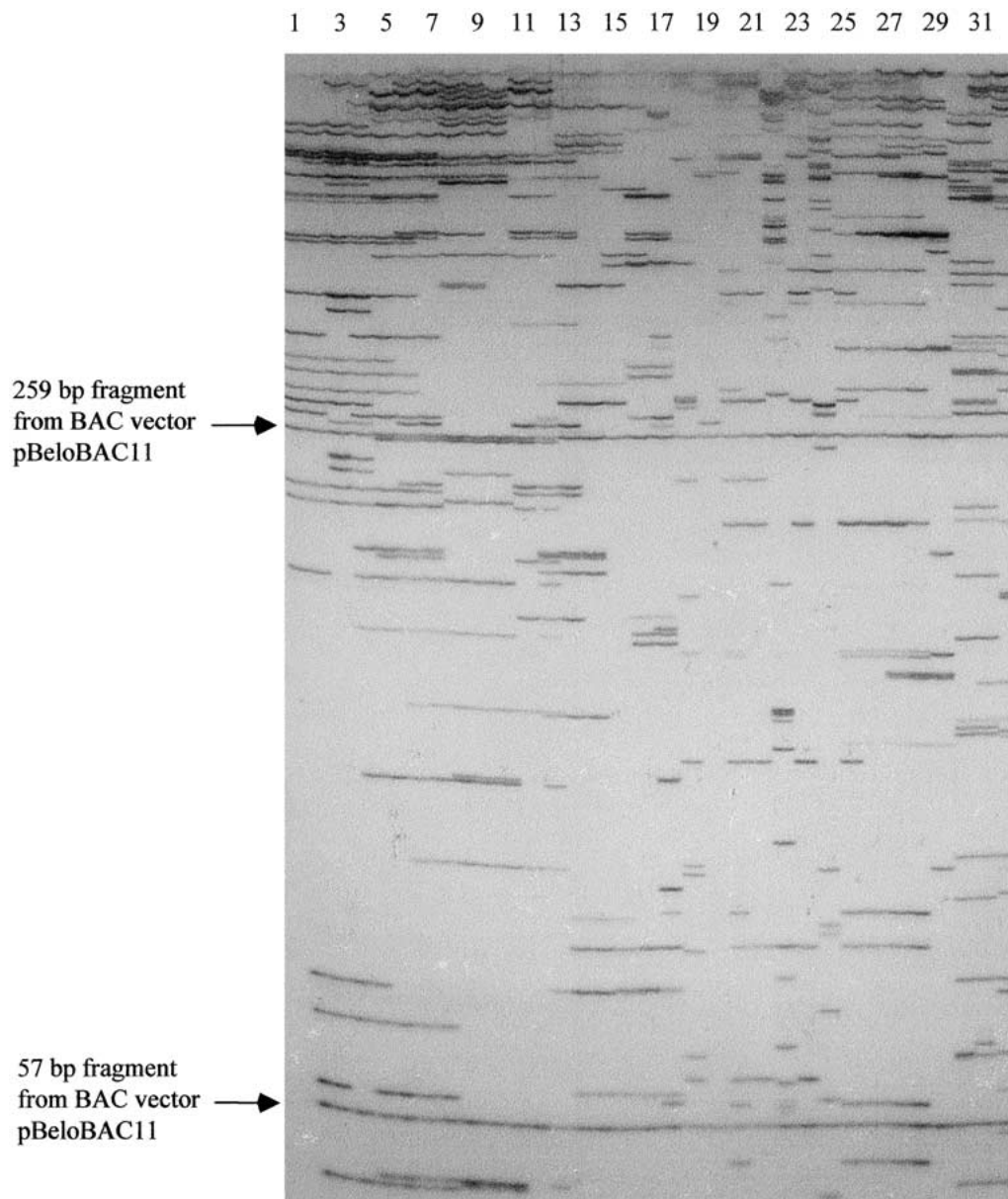


Figure 2. Autoradiograph of a  $^{32}\text{P}$ -labeled DNA fingerprinting gel. Each lane represents fingerprints of a single BAC clone digested with *Bam*HI and *Hae*III followed by labeling at the *Bam*HI terminus. Lane 1 to 32 correspond to 32 positive BAC clones selected from both *M. floribunda* 821 and 'GoldRush' BAC libraries using 14 *Vf*-linked SCAR markers: *M85-H-20*, *G4-I-23*, *M61-C-12*, *M58-C-7*, *M1-F-12*, *M2-J-3*, *M40-I-12*, *M12-I-5*, *M60-K-11*, *M70-G-2*, *M34-L-22*, *M4-P-11*, *M61-M-16*, *G11-J-23*, *G7-C-18*, *G53-N-7*, *M24-B-18*, *G20-G-18*, *G70-F-18*, *G9-N-2*, *G92-J-1*, *M66-E-9*, *M72-M-3*, *M81-E-10*, *M89-H-11*, *G44-C-14*, *M15-N-1*, *M72-E-14*, *M60-I-19*, *M72-E-9*, *G32-M-9*, and *M89-I-20*. Three small contigs can be assembled according to the fingerprinting patterns. The first contig consists of 14 BAC clones, corresponding to lane 1 to 17, except for lane 8 (*M12-I-5*), 9 (*M60-K-11*), and 10 (*M70-G-2*); the second contig has three clones, corresponding to lane 30 to 32; and the third contig is composed of 9 clones from lane 18 to 29, except for lane 19 (contaminated *G70-F-18*), lane 22 (*M66-E-9*), and lane 24 (*M81-E-10*).

Table 1. Positive BAC clones identified from both *Malus floribunda* 821 and 'GoldRush' BAC libraries using 14 *Vf*-linked SCAR markers\*.

Order of SCARs	<i>Malus floribunda</i> 821 BAC library	'GoldRush' BAC library
Left to the <i>Vf</i> gene:		
ACS-6	M58-C-7 (140 kb), M61-C-12 (90 kb), M85-H-20 (105 kb), <i>M1-F-12 (135 kb)</i> , <i>M2-J-3 (160 kb)</i> , <i>M40-I-12 (140 kb)</i>	G4-I-23 (110 kb)
M18	<i>M1-F-12 (135 kb)</i> , <i>M2-J-3 (160 kb)</i> , <i>M40-I-12 (140 kb)</i> , M12-I-5 (100 kb), M34-L-22 (100 kb), M60-K-11 (100kb), M70-G-2 (90 kb), <b>M4-P-11 (170 kb)</b>	
Co-segregating with the <i>Vf</i> gene:		
ACS-3	<b>M4-P-11 (170 kb)</b> , <b>M61-M-16 (115 kb)</b>	<b>G11-J-23 (85 kb)</b>
ACS-7	<b>M61-M-16 (115 kb)</b>	<b>G11-J-23 (85 kb)</b> , <b>G7-C-18 (90 kb)</b>
ACS-9		<b>G7-C-18 (90 kb)</b> , <b>G53-N-7 (95 kb)</b>
Right to the <i>Vf</i> gene:		
ACS-8	M24-B-18 (120 kb)	<b>G53-N-7 (95 kb)</b>
ACS-10	<u>M72-E-9 (100 kb)</u>	<u>G32-M-9 (125 kb)</u>
ACS-4	<u>M72-E-9 (100 kb)</u> , M89-I-20 (100 kb)	<u>G32-M-9 (125 kb)</u>
ACS-11		<u>G20-G-18 (115 kb)</u> , <u>G70-F-18 (140 kb)</u>
ACS-2		<u>G20-G-18 (115 kb)</u> , <u>G70-F-18 (140 kb)</u> <u>G9-N-2 (145 kb)</u> , <u>G92-J-1 (115 kb)</u>
ACS-5		<u>G20-G-18 (115 kb)</u> , <u>G70-F-18 (140 kb)</u> <u>G9-N-2 (145 kb)</u> , <u>G92-J-1 (115 kb)</u>
OPAM19	M66-E-9 (105 kb)	<u>G70-F-18 (140 kb)</u> , <u>G9-N-2 (145 kb)</u> <u>G92-J-1 (115 kb)</u>
OPAL07	M72-M-3 (100 kb), M81-E-10 (95 kb), M89-H-11 (100 kb), <u>M15-N-1 (100 kb)</u> , <u>M72-E-14 (85 kb)</u>	<u>G9-N-2 (145 kb)</u> , <u>G92-J-1 (115 kb)</u> G44-C-14 (120 kb)
ACS-1	<u>M15-N-1 (100 kb)</u> , <u>M72-E-14 (85 kb)</u> , M60-I-19 (50 kb)	

\*BAC names beginning with M represent BAC clones selected from the *M. floribunda* 821 BAC library; while, BAC names beginning with G represent BAC clones selected from the 'GoldRush' BAC library. The insert sizes of identified BAC clones were estimated using CHEF electrophoresis, and written in parenthesis after names of BAC clones. Bridging clones left to the *Vf* gene are marked in *italics*; bridging clones co-segregating with the *Vf* gene are marked in **bold**; and bridging clones right to the *Vf* gene are underlined.

the border clone *M24-B-18*, the left-end probe amplified five overlapping clones, and was oriented opposite to the gap; while, the right-end probe amplified the cognate clone *M24-B-18*, seven non-overlapping BAC clones, and both *M. floribunda* 821 and the scab-susceptible bulk, and was oriented toward the gap (Tables 2 and 3). Unfortunately, PCR-based screening of the two BAC libraries using this right-end probe gave rise to amplicons for almost all super-pools, and failed to identify any positive BAC clones. Thus, the right-end probe of *M24-B-18* was deemed to be derived from a repetitive sequence, and chromosome

walking could not proceed using this probe. For the clone *M72-E-9*, the right-end probe amplified the cognate clone *M72-E-9* and one of its overlapping clones *G32-M-9*; while, the left-end probe amplified *M72-E-9* and another overlapping clone *M89-I-20*. Thus, *M72-E-9* must be located between *G32-M-9* and *M89-I-20*, the orientation of the two end probes could not yet be determined until the two gaps were closed. Screening of the two BAC libraries using both right-end and left-end probes revealed four (*M61-E-12*, *M81-G-7*, *M81-P-10*, and *M86-F-16*) and three (*M30-*

Table 2. List of BAC insert-end primers used for chromosome walking and contig extension.

BAC insert-end primers	Sequences		Position
The first gap			
M24-B-18			
Left-end	GGTCACTACATCGCAAGTC	GCTACTAGGGTTTGCTTTGG	Opposite to the gap
Right-end	AACTGGTATTGCCGTCTTCC	GGCATGGTTGTAAGCTTGTG	Toward the gap
G53-N-7			
Left-end	GCAGAAGGTATAGAATTTTGAAAC	TACTTCTCACACACCCCTTG	Toward the gap
Right-end	GGATATTGTGGATAAGGCTC	CTGCAACACATCCAGGCAAG	Opposite to the gap
G28-C-23			
Left-end	CCCAAAAACGCCTTAATATATG	GATATCGTTTTTGCATTTGCTC	Toward the gap
Right-end	CTGAGATTAACATCCCCCTG	CTTCCCTTACCTGCTATGAG	Opposite to the gap
M72-E-9			
Left-end	GATGATCATCACACATGTC	GATGTGATGTGCACACTG	Opposite to the gap
Right-end	TCACTTTAGATGGTTCTGAC	TTAGAAATCCATGCTCAGGG	Toward the gap
The second gap			
M72-E-9			
Left-end	GATGATCATCACACATGTC	GATGTGATGTGCACACTG	Toward the gap
Right-end	TCACTTTAGATGGTTCTGAC	TTAGAAATCCATGCTCAGGG	Opposite to the gap
G20-G-18			
Left-end	ACAGCCGCTTATTCTCACAG	GGCACTGTGCAGAAGTATGG	Opposite to the gap
Right-end	CTCCTACCTGATTCTGCAGC	AGGTGCAGGATCAGAGATTG	Toward the gap
G70-F-18			
Left-end	TGGCAACCGCTATGATCGTG	AGGAAGGCAGAGTCGAGTCG	Toward the gap
Right-end	GTCACCTCAATTATTTACCCTACTC	GAATTCCTGCCTATCAGACG	Opposite to the gap
G54-G-7			
Left-end	AACCTTTTTATAACTACCTGGC	CGGAAGTCCGAAGAAACACT	Toward the gap
Right-end	CTGTGAAAGAACAACGCTCC	TACAGTGGCTCATCGGTTAG	Opposite to the gap
Contig extension			
M60-I-19			
Left-end	TCGAGCCACGACAGCCTATC	TCTCCACAAAGTAGATCCGG	On the border
Right-end	ATGATCTATCGGCCCGATC	TCTAGAGGATCCACAACCGC	Inside the contig
M63-D-6			
Left-end	ACGTTTCTTACCCAAAGGTG	AGCAGTCCACCATTGTCAAG	Inside the contig
Right-end	AACCTATCACCCGTCACATG	GCTCAGGCATGCCATATATG	On the border
M12-M-15			
Left-end	GGTGAAGCTAGGCACATAAG	TATTCACCACCTCCTCCCTC	On the border
Right-end	GATATCGACGGTCTTGCTGACAC	GGATCCAATGAGAGAAAGGAAG	Inside the contig
M11-A-16			
Left-end	GGCATGCTAGATTAGGACATG	CCTCCTCATCAATGAATGTGG	On the border
Right-end	GGAAATTCGGTTTGTTTG	CCCATGCATTAATTATAAATCAC	Inside the contig

*L-16*, *M47-N-1*, and *G4-G-23*) positive BAC clones, respectively.

Since the border clone *M24-B-18* failed to identify any new overlapping BAC clones, the next border clone *G53-N-7* was selected for chromosomal walking. The left-end probe of *G53-N-7* was positioned

toward the gap, and showed very specific to the *Vf* region (Tables 2 and 3). Screening of the two BAC libraries using this probe allowed identification of a single positive clone (*G28-C-23*) from the 'GoldRush' BAC library (Table 3). Fingerprinting analysis further confirmed that *G28-C-23* overlapped with its

Table 3. Presence/absence of PCR products of BAC clones in the first and second initial contigs using insert-end primers derived from BAC clones located in the first inter-contig gap\*.

The ordered BAC clones in the first and second initial contigs	Left border				Right border			
	M24-B-18		G53-N-7		G28-C-23		M72-E-9	
	L	R	L	R	L	R	L	R
M85-H-20	-	+	-	-	-	-	-	-
G4-I-23	-	+	-	-	-	-	-	-
M61-C-12	-	-	-	-	-	-	-	-
M58-C-7	-	+	-	-	-	-	-	-
M1-F-12	-	+	-	-	-	-	-	-
M2-J-3	-	+	-	-	-	-	-	-
M40-I-12	-	+	-	-	-	-	-	-
M34-L-22	-	+	-	-	-	-	-	-
M4-P-11	+	-	-	-	-	-	-	-
M61-M-16	+	-	-	+	-	-	-	-
G11-J-23	+	-	-	+	-	-	-	-
G7-C-18	+	-	-	+	-	-	-	-
G53-N-7	+	-	+	+	-	+	-	-
M24-B-18	+	+	+	-	-	+	-	-
G28-C-23			+		+	+		
M60-I-10					+	-		
M86-F-16					+	-		+
M81-P-10					+	-		+
M81-G-7					+	-		+
M61-E-12					+	-		+
G32-M-9	-	-	-	-	-	-		+
M72-E-9	-	-	-	-	-	-	+	+
M89-I-20	-	-	-	-	-	-	+	
M47-N-1					-	-	+	
G4-G-23					-	-	+	
M30-L-16					-	-	+	
<i>M. floribunda</i> 821	+	+	+	+	+	+	+	+
Scab-susceptible bulk	-	+	-	+	-	+	-	-

\*+: Presence of PCR products; -: Absence of PCR products; Blank: no data available.

neighboring clones. At least four fingerprints appeared to be common between *G28-C-23* and four positive clones (*M61-E-12*, *M81-G-7*, *M81-P-10*, and *M86-F-16*) detected by the right-end probe of *M72-E-9*. This indicated that the gap was likely closed, although more evidence was needed to confirm closure of this gap. A new round of chromosome walking was launched using the newly identified clone *G28-C-23*. The left-end probe of *G28-C-23* was found to be oriented toward the gap, and was very specific to the *Vf* region (Tables 2 and 3). This probe detected the following five positive BAC clones, *M60-I-10*, *M61-E-12*, *M81-G-7*,

*M81-P-10*, and *M86-F-16*. Among those, four BAC clones (*M61-E-12*, *M81-G-7*, *M81-P-10*, and *M86-F-16*) were exactly the same as those identified by the right-end probe of *M72-E-9*. Fingerprinting analysis further confirmed the overlapping relationships among these BAC clones by the presence of many shared fingerprints (Figure 3). Thus far, the first gap had been sealed.

In the second gap, chromosome walking was launched from the right border clone *G20-G-18*. The right-end probe of *G20-G-18* was positioned toward the gap (Tables 2 and 4); however, this probe ampli-

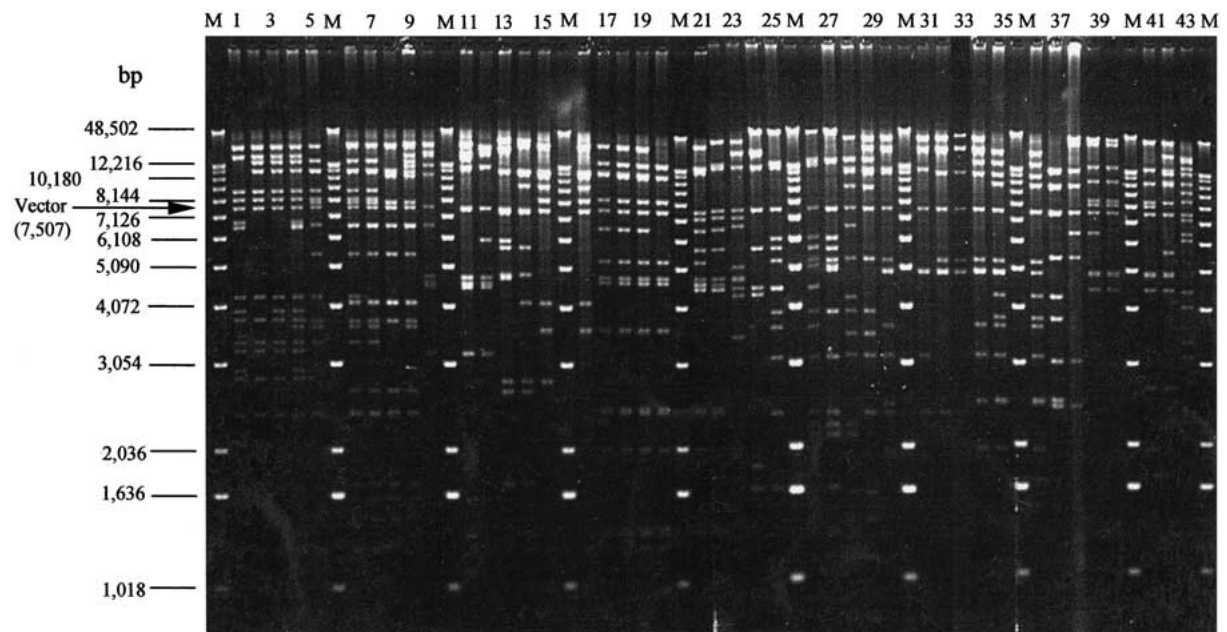


Figure 3. An agarose gel of *Bam*HI-digested DNA fragments for 43 BAC clones of a megabase-size BAC contig on the *Vf* chromosomal region. M: molecular marker (lambda DNA plus 1 kb DNA ladder marker). Lanes 1 to 43 correspond to 43 positive BAC clones: *M61-C-12*, *M85-H-20*, *G4-I-23*, *M58-C-7*, *M1-F-12*, *M2-J-3*, *M40-I-12*, *M34-L-22*, *M4-P-11*, *M61-M-16*, *G11-J-23*, *G7-C-18*, *G53-N-7*, *M24-B-18*, *G28-C-23*, *M60-I-10*, *M86-F-16*, *M81-P-10*, *M81-G-7*, *M61-E-12*, *G32-M-9*, *M72-E-9*, *M89-I-20*, *M47-N-1*, *G4-G-23*, *M30-L-16*, *G54-G-7*, *G70-F-18*, *G20-G-18*, *G9-N-2*, *G92-J-1*, *M72-M-3*, *M89-H-11*, *G44-C-14*, *M15-N-1*, *M72-E-14*, *M60-I-19*, *M63-D-6*, *M12-M-15*, *M70-P-19*, *M41-L-8*, *M11-A-16*, and *M2-E-13*. All these 43 BAC clones created a single megabase-size BAC contig on the *Vf* region.

fied almost one-third of the super-pools of the two BAC libraries, and failed to identify any overlapping clones. Instead, another border clone, *G70-F-18*, was selected for chromosome walking. The left-end probe of *G70-F-18* was located toward the gap (Tables 2 and 4), and amplified a number of super-pools. PCR products from all positive super-pools were then subjected to complete digestion with several restriction enzymes. Fortunately, patterns of restriction fragments displayed differences between PCR products derived from the *Vf* region from those derived from non-*Vf* regions (data not shown). Thus, one positive clone, *G54-G-7*, was obtained from the 'GoldRush' BAC library that overlapped with *G70-F-18*, and was used for the next step of chromosome walking. The left-end probe of *G54-G-7* was placed toward the gap (Tables 2 and 4), and detected three positive BAC clones (*M47-N-1*, *M30-L-16*, and *G4-G-23*) from both *M. floribunda* 821 and 'GoldRush' BAC libraries. Interestingly, all three BAC clones had already been identified using the left-end probe of *M72-E-9*. The overlapping relationship among these positive BAC clones was confirmed by fingerprinting analysis, resulting in closure of the second gap (Figure 3).

Apart from the closure of the two inter-contig gaps, the contig was also extended on the right side by chromosome walking. The right-most BAC clone *M60-I-19* in the third contig was used to initiate chromosome walking. The left-end probe of *M60-I-19* was placed on the right border, and detected an overlapping clone, *M63-D-6* (Tables 2 and 5). A *Not*I recognition site was found within the insert of *M63-D-6*. The other *Not*I site left to the *Vf* locus has already been identified, and is very close to the SCAR OPM18. The physical distance between these two *Not*I sites was estimated to be 870 kb (Patocchi *et al.*, 1999). In the second walk, two new overlapping clones, *M12-M-15* and *M70-P-19*, were detected by using the left-end probe of *M63-D-6*. In the third walk, another two new overlapping clones, *M11-A-16* and *M41-L-8*, were detected by using the left-end probe of *M12-M-15*. Furthermore, the left-end probe of *M11-A-16* identified its overlapping clone, *M2-E-13* (Table 5). The overlapping relationships among these newly identified BAC clones were confirmed by fingerprinting analysis (Figure 3). Thus, the contig was successfully extended to over one megabase in size (Figures 3 and 4).

Table 4. Presence/absence of PCR products of BAC clones in the second and third initial contigs using insert-end primers derived from BAC clones located in the second inter-contig gap\*.

The ordered BAC clones in the second and third initial contigs	Left border		Right border					
	M72-E-9		G54-G-7		G70-F-18		G20-G-18	
	L	R	L	R	L	R	L	R
M86-F-16		+	-	-				
M81-P-10		+	-	-				
M81-G-7		+	-	-				
M61-E-12		+	-	-				
G32-M-9		+	-	-	-	-	-	-
M72-E-9	+	+	-	-	-	-	-	-
M89-I-20	+		-	-	-	-	-	-
M47-N-1	+		+	-				
G4-G-23	+		+	-				
M30-L-16	+		+	-				
G54-G-7			+	+	+			
G70-F-18			-	+	+	+		
G20-G-18	-	-	-	-	-	+	+	+
G9-N-2	-	-	-	-	-	+	+	-
G92-J-1	-	-	-	-	-	+	+	-
M72-M-3	-	-	-	-	-	-	+	-
M89-H-11	-	-	-	-	-	-	-	-
G44-C-14	-	-	-	-	-	-	+	-
M15-N-1	-	-	-	-	-	-	-	-
M72-E-14	-	-	-	-	-	-	-	-
M60-I-19	-	-	-	-	-	-	-	-
<i>M. floribunda</i> 821	+	+	+	+	+	+	+	+
Scab-susceptible bulk	-	-	+	-	-	+	-	-

\*+: Presence of PCR products; -: Absence of PCR products; Blank: no data available.

#### Construction of a megabase-size BAC contig and a sequence-ready map for the Vf region

To assess clone integrity and produce a megabase-size BAC contig for the Vf region, fingerprints of all BAC clones involved in the contig were generated by a complete *Bam*HI-digestion (Figure 3). A restriction fragment (7,507 bp), corresponding to the linear BAC vector pBeloBAC11, was present in all BAC clones. All other restriction fragments were generated from BAC inserts, and could be validated by several overlapping clones. This indicated the contig had sufficient depth to produce a sequence-ready map.

The size of this BAC contig has been estimated to be ca. 1,100 kb, based on insert sizes and fingerprinting patterns of all BAC clones and the estimated physical distance (870 kb) between two *Not*I sites (one very close to OPM18, and the other near ACS-1). Our

contig extended beyond these two *Not*I sites. The contig is, therefore,  $\sim 4 \times$  deep, and meets all criteria set by McPherson (1997) to produce a sequence-ready map. So far, a minimum of 16 tiling BAC clones covering the Vf region have been produced for sequencing (Figure 4).

#### Duplication events in the Vf region and identification of Vf-containing candidate clones

From the present study, it could be concluded that duplication events frequently occurred in the Vf region. First, shared fingerprints frequently appeared among non-overlapping BAC clones. For instance, at least four *Bam*HI-digested DNA fragments were noted to be common between *M61-M-16* and *M72-E-9* (Figure 3). Second, some newly developed PCR-based probes from BAC insert ends could amplify non-

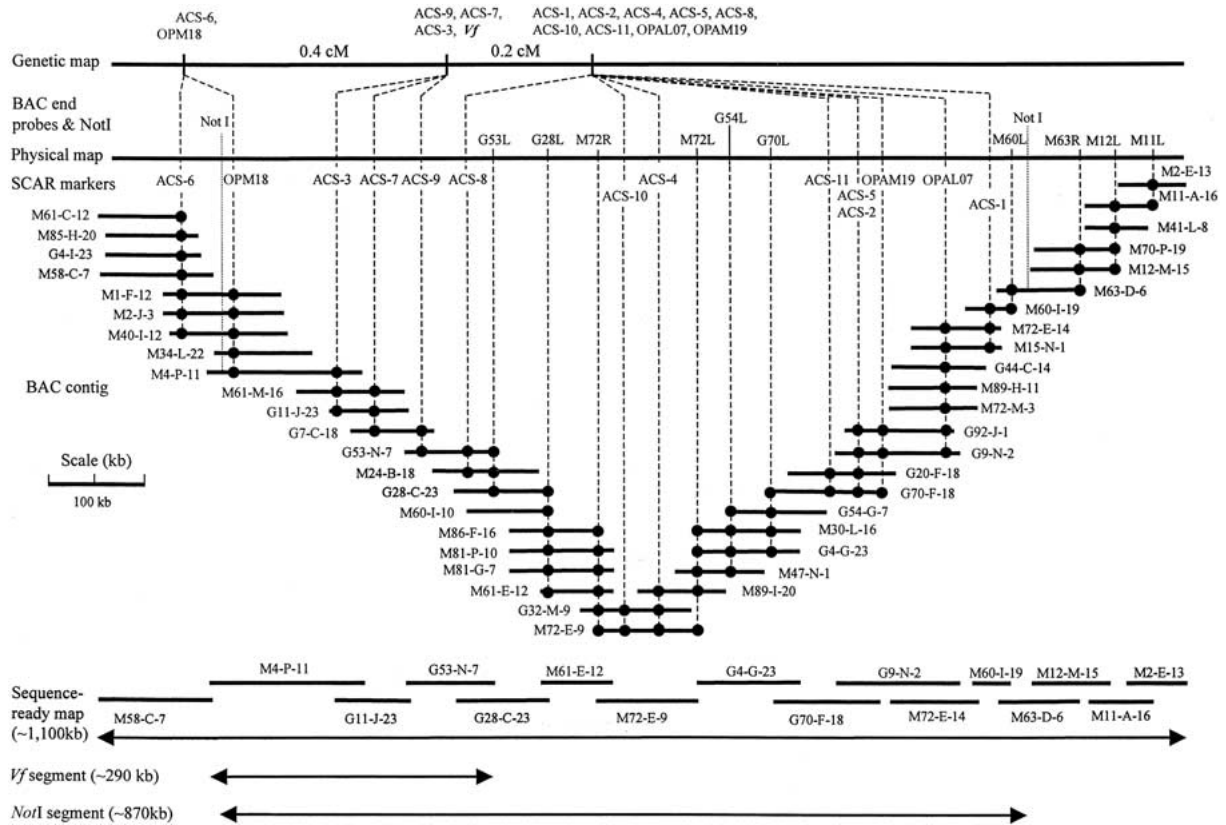


Figure 4. Genetic and physical maps, a megabase-size BAC contig, and a sequence-ready map of the *Vf* region.

overlapping clones, such as the right-end probe of *M24-B-18* (Table 3). The results demonstrated the presence of a multi-copy feature for these insert ends. Third, two BAC clones, *M66-E-9* and *M81-E-10*, identified by *Vf*-linked SCAR markers were *de facto* not derived from the *Vf* region, suggesting that highly homologous sequences involving these SCARs were present in other chromosomal regions. The occurrence of duplication events in the *Vf* region was also supported by our previous results, whereby identical sequences between two AFLP markers, *ET3MG10-1* (co-segregating with *Vf*) and *ET3MG10-2* (right to *Vf*), were obtained, except for a 6-bp deletion in *ET3MG10-2* (Xu *et al.*, 2001a). Moreover, in our previous study, fiber-FISH signals revealed that the insert size of *M61-M-11* was around 200 kb, much longer than the actual size of 115 kb estimated by CHEF electrophoresis, and indicating presence of repetitive DNA sequences (Xu *et al.*, 2001b).

Based on the genetic map (Figure 4), three close crossover events flanking the *Vf* locus must have occurred. The first two (left to *Vf*) are located between

OPM18 and ACS-3, and the third (right to *Vf*) is located between ACS-9 and ACS-8. Therefore, the *Vf* locus is restricted to a narrow region between OPM18 (left) and ACS-8 (right). A minimum of three contiguous and overlapping BAC clones (from left to right: *M4-P-11*, *G11-J-23*, and *G53-N-7*) covers the *Vf* locus with a physical distance of ca. 290 kb. These three clones can then be used to screen a cDNA library to search for the candidate *Vf* gene. This also indicates that the three crossover events have occurred in the very vicinity of the *Vf* locus in our mapping population containing only 500 scab-resistant individuals.

### Discussion

In the present study, a very efficient and robust approach is presented for constructing a megabase-size BAC contig for a poorly mapped chromosomal region in the apple genome. In the construction process, we have taken full advantage of the AFLP technique for saturating the target region, screening BAC libraries, and assembling a megabase-size BAC contig. This

Table 5. Presence/absence of PCR products of BAC clones in the third initial contig using insert-end primers derived from BAC clones located on the right border of the third initial contig\*.

The ordered BAC clones in the third initial contig	Right border of the third initial contig							
	M60-I-19		M63-D-6		M12-M-15		M11-A-16	
	L	R	L	R	L	R	L	R
G20-G-18	-	-	-	-	-	-	-	-
G9-N-2	-	-	-	-	-	-	-	-
G92-J-1	-	-	-	-	-	-	-	-
M72-M-3	-	-	-	-	-	-	-	-
M89-H-11	-	+	-	-	-	-	-	-
G44-C-14	-	+	-	-	-	-	-	-
M15-N-1	-	+	-	-	-	-	-	-
M72-E-14	-	+	-	-	-	-	-	-
M60-I-19	+	+	+	-	-	-	-	-
M63-D-6	+		+	+	-	+	-	-
M12-M-15				+	+	+	-	+
M70-P-19				+	+	-	-	+
M41-L-8					+		-	+
M11-A-16					+		+	+
M2-E-13							+	
<i>M. floribunda</i> 821	+	+	+	+	+	+	+	+
Scab-susceptible bulk	-	+	-	+	+	+	+	+

\*+: Presence of PCR products; -: Absence of PCR products; Blank: no data available.

viable approach can then be extended to other plant genomes where high-density linkage maps around a target region are not readily available.

In saturating the target region, the AFLP technique allowed us to tag the *Vf* region with 15 closely-linked AFLP markers (~40 kb per marker) within a period of six months (Xu and Korban, 2000). In contrast, only three *Vf*-linked markers (OPM18, OPAM19, and OPAL07) have been obtained using the RAPD method following intensive efforts in several laboratories for a period of more than 10 years (Yang and Korban 1996; Tartarini 1996; Tartarini *et al.*, 1999; Gardiner *et al.*, 1996; Gianfranceschi *et al.*, 1996; Yang *et al.*, 1997 a,b; Hemmat *et al.*, 1998). In fact, it has been nearly impossible to further saturate the *Vf* region with additional RAPD markers. With a physical distance of 550 kb between OPM18 and OPAM19/OPAL07, chromosome walking was the only option available for contig construction (Patocchi *et al.*, 1999).

For BAC library screening, AFLP-derived SCARs have proven to be highly reliable, specific to the *Vf* region, and allow rapid identification of positive clones from BAC libraries. All positive BAC clones detected

by our 11 AFLP-derived SCARs have been confirmed to be derivatives of the *Vf* region. This is in contrast to previous efforts using the three RAPD-derived SCAR markers (OPM18, OPAM19, and OPAL07) where 'false positive' clones were often obtained. Three BAC clones (*M12-I-5*, *M60-K-11*, and *M70-G-2*) detected by OPM18 are derived from the *Vf* homologous region rather than from the *Vf* region itself. Another two clones, *M66-E-9* and *M81-E-10*, whose origins remain unknown, have been detected by the SCAR markers OPAM19 and OPAL07. The least effective probes for screening BAC libraries are those developed from BAC insert ends. At times, BAC insert ends reside on repetitive sequences, and the derived probes have failed to identify any positive clone from the two BAC libraries, such as the right-end probe of *M24-B-18*. The presence of repetitive sequences has been a major obstacle in chromosome walking, as reported by others (Tanksley *et al.*, 1995; Stein *et al.*, 2000).

For BAC contig assembly, AFLP-derived SCAR markers are dense with an average physical distance of only ~70 kb per marker (estimated by dividing the size of contig by the number of SCARs), and are

shorter than the average insert size of  $\sim 100$  kb of a BAC clone. As a consequence, the majority of BAC clones derived from the *Vf* region have been readily recovered (Table 1; Figures 1 and 2). This, in turn, has greatly facilitated the construction of a BAC contig covering the *Vf* region. As shown in this study, three small BAC contigs have been rapidly constructed after screening of the two BAC libraries. Following closure of the two inter-contig gaps by chromosomal walks, a single megabase-size BAC contig spanning the *Vf* region has been constructed. The chromosomal structure of the *Vf* region must be quite complex and rich in repetitive sequence as severe duplication events have occurred in this region. This may be a universal phenomenon for chromosomal regions harboring disease resistance genes (Meyers *et al.*, 1998; Richter and Ronald 2000; Wise 2000). Construction of BAC contigs for such chromosomal regions by chromosome walking is a major challenge, and at times impossible. This has been also observed in the present study whereby the most difficult and time-consuming effort involved closure of the two small inter-contig gaps by several steps of chromosome walking. At times, no chromosome walking has been achieved due to the presence of repetitive DNA sequences along insert ends of BAC clones.

It has been reported that AFLP markers generated by using a specific set of restriction enzymes, for example *EcoRI/MseI*, may not be evenly distributed throughout the entire genome (Young *et al.*, 1999). In this study, all SCAR markers within the short chromosomal *Vf* region are clustered in three hot spots corresponding to the three initial BAC contigs. Some are so close that distances between neighboring SCARs are significantly shorter than the average expected distance ( $\sim 70$  kb). For instance, the BAC clone *G70-F-18* (120 kb) covers four SCARs including ACS-11, ACS-5, ACS-2, and OPAM19, and the average distance among these SCARs is  $\sim 30$  kb. On the other hand, the distance for each of the two inter-contig gaps (intervals of ACS-8/ACS-10 and ACS-4/ACS-11) is  $\sim 160$  kb, and therefore it is much longer than the average expected distance ( $\sim 70$  kb). If the gaps are too large to be closed, or closure of the gaps is difficult due to the presence of repetitive sequences, an alternative set of restriction enzymes, for instance *PstI/MseI*, are recommended to generate more randomly-scattered AFLP markers.

In plant genome sequencing projects, building a sequence-ready map for an entire chromosome is frequently hindered by the presence of gaps scattered

along the chromosome. Closure of these gaps is crucial for successful genome sequencing. The contig building strategy presented in this study is highly useful for closing such gaps by generating high-density AFLP markers within gaps followed by contig construction. In comparison with a 'narrow-down' screening strategy (Xu and Korban, 2000), some modifications are required for preparing 'positive' and 'negative' bulks while searching for AFLP markers. For a given gap, a positive bulk is prepared by pooling individuals carrying the closest flanking markers (RFLP or PCR-based markers). While, a negative bulk is made up of individuals without the closest flanking markers. In general, if the genetic distance of a gap is less than 1 cM, a single negative bulk is sufficient for screening AFLP markers as double-crossover events rarely occur within such a narrow interval. If a gap extends to a genetic distance of more than 1 cM, double-crossover events within the gap are likely to occur, and must be taken into consideration. Under this circumstance, no AFLP markers can be identified within an interval limited by two crossovers if a double-crossover individual(s) is unknowingly included in the negative bulk. Alternately, several negative bulks consisting of different individuals are required, since it is unlikely that each negative bulk may contain a double-crossover individual(s), and all double-crossovers may occur in the same chromosomal region. If AFLP bands are present both in the positive parent line and the positive bulk, but absent in the negative parental line and in at least one of the negative bulk(s), these bands are likely to be derived from the region within the gaps. Following linkage mapping using a segregant population, AFLP markers located within the gap can then be identified. A double-crossover individual(s) can be identified in the negative bulk if the individual(s) shows promising AFLP markers. By increasing the number of AFLP primer pairs and using different sets of restriction enzymes, the numbers of AFLP markers that can be generated are unlimited. Therefore, it is possible to saturate gaps with AFLP markers. Following conversion of AFLP markers into SCARs, a global sequence-ready map can be built for sequencing the entire genome. Klein *et al.*, (2000) have used an alternative method for exploiting the AFLP technique in constructing integrated genetic and physical maps for sorghum by pooling BAC clones in six different matrices and using generated AFLP markers to assemble BAC contigs.

Some AFLP markers, for example *EA2MG11-1* in our previous report (Xu *et al.*, 2001a), can't be con-

verted into SCAR markers due to lack of nucleotide polymorphism between the two parental lines. However, the sequence information of such AFLP markers is still very useful for designing PCR-based probes used for screening BAC libraries. It is important to point out that such AFLP-derived probes will detect BAC clones from both the target region and its homologous region. Therefore, an additional step is needed to differentiate between 'real' BAC clones (those from the target region) from 'false' clones (those from the homologous region). In the present study, a probe derived from *EA2MG11-1* has detected a single BAC clone (*M61-M-16*) from the *M. floribunda* 821 BAC library and three (*G7-C-18*, *G11-J-23*, and *G25-A-3*) BAC clones from the 'GoldRush' BAC library. Of these four BAC clones, three (*M61-M-16*, *G7-C-18*, and *G11-J-23*) are the same as those detected by the ACS-3 SCAR markers, and are thus derived from the *Vf* region; while, the remaining clone, *G25-A-3*, is derived from the homologous region as confirmed by the fingerprinting analysis (data not shown). This suggests that all AFLP markers can be used for contig construction. This is especially useful when the genetic diversity between two parental lines is low, and when highly homologous sequences involving AFLP markers are present between two parental lines.

A highly redundant BAC library is essential for successful construction of a BAC contig for a target region. In the present study, positive BAC clones from either *M. floribunda* 821 or 'GoldRush' libraries are not enough to cover the whole *Vf* region, although each library has provided approximately five-fold haploid-genome equivalents. It is only after combining the positive BAC clones from both libraries that we have been able to construct a BAC contig covering the entire *Vf* region. Thus, a 10-fold redundant BAC library is needed to obtain a large-size BAC contig of the *Vf* region.

### Acknowledgements

This research was supported by a grant from the Illinois Council for Food and Agricultural Research-External Program, and from the Victor Kelley Memorial Fund.

### References

- Bénaouf, G. and Parisi, G. 2000. Genetics of host-pathogen relationships between *Venturia inaequalis* races 6 and 7 and *Malus* species. *Phytopathology* 90: 236-242.
- Bournival, B.L. and Korban, S.S. 1987. Electrophoretic analysis of genetic variability in the apple. *Scient. Hort.* 31: 233-243.
- Cao, Y., Kang, H.L., Xu, X., Wang, M., Dho, S.H., Huh, J.R., Lee, B.J., Kalush, F., Bocskai, D., Ding, Y., Tesmer, J.G., Lee, J., Moon, E., Jurecic, V., Baldini, A., Weier, H.U., Doggett, N.A., Simon, M.I., Adams, M.D. and Kim, U.J. 1999. A 12-Mb complete coverage BAC contig map in human chromosome 16p13.1-p11.2. *Genome Res.* 9: 763-774.
- Chen, C.X. and Gmitter, F.G. 1999. Direct cloning and sequencing of bacterial artificial chromosome (BAC) insert ends based on double digestion. *Plant Mol. Biol. Repr.* 17: 231-238.
- Crosby, J.A., Janick, J., Pecknold, P.C., Korban, S.S., O'Connor, P.A., Ries, S., Goffreda, J. and Voordeckers, A. 1992. Breeding apples for scab resistance: 1945-1990. *Fruit Var. J.* 46: 145-166.
- Gardiner, S.E., Bassett, H.C.M., Noiton, D.A.M., Bus, V.G., Hofstee, M.E., White, A.G., Ball, R.D., Forster, R.L.S. and Rikkerink, E.H.A. 1996. A detailed linkage map around an apple scab resistance gene demonstrates that two disease resistance classes both carry the *Vf* gene. *Theor. Appl. Genet.* 93: 485-493.
- Gessler, C. 1989. Genetics of interaction *Venturia inaequalis-Malus*: The conflict between theory and reality. Pages 168-190 in: *Integrated Control of Pome Fruit Diseases*, Vol. II. C. Gesseler, D.J. Butt, and B. Koller, eds IOBC/WPRS, Zurich, Switzerland.
- Gianfranceschi, L., Koller, B., Seglias, N., Kellerhals, M. and Gessler, C. 1996. Molecular selection in apple for resistance to scab caused by *Venturia inaequalis*. *Theor. Appl. Genet.* 93: 199-204.
- Haanstra, J.P.W., Wye, C., Verbakel, H., Meijer-Dekens, F., van den Berg, P., Odinet, P., van Heusden, A.W., Tanksley, S., Lindhout, P. and Peleman, J. 1999. An integrated high density RFLP-AFLP map of tomato based on two *Lycopersicon esculentum* x *L-pennellii* F-2 populations. *Theor. Appl. Genet.* 99: 254-271.
- Hemmat, M., Weeden, N.F., Aldwinckle, H.S. and Brown, S.K. 1998. Molecular markers for the scab resistance (*Vf*) region in apple. *J. Amer. Soc. Hort. Sci.* 123: 992-996.
- Klein, P.E., Klein, R.R., Cartinhour, S.W., Ulanich, P.E., Dong, J., Obert, J.A., Morishige, D.T., Schleuter, S.D., Childs, K.L., Ale, M. and Mullet, J.E. 2000. A high-throughput AFLP-based method for constructing intergrated genetic and physical maps: Progress towards a sorghum genome map. *Genome Res.* 10: 789-807.
- Korban, S.S. and Chen, H. 1992. Apple. In: Litz R. and F.A. Hammerschlag (Eds) *Biotechnology of perennial fruit crops*. CAB Intl., Wallingford, UK, pp. 203-227.
- Lukowitz, W., Gillmor, C.S. and Scheible, W.R. 2000. Positional cloning in *Arabidopsis*. Why it feels good to have a genome initiative working for you. *Plant Physiol.* 123: 795-805.
- Manganaris, A.G., Alston, F.H., Weeden, N.F., Aldwinckle, H.S., Gustafson, H.L. and Brown, S.K. 1994. Isozyme locus *Pgm-1* is tightly linked to a gene (*Vf*) for scab resistance in apple. *J. Amer. Soc. Hort. Sci.* 119: 1286-1288.
- Marra, M.A., Kucaba, T.A., Dietrich, N.L., Green, E.D., Brownstein, B., Wilson, R.K., McDonald, K.M., Hillier, L.W., McPherson, J.D. and Waterston, R.H. 1997. High throughput fingerprint analysis of large-insert clones. *Genome Res.* 7: 1072-1084.
- McPherson, J.D. 1997. Sequence ready-or not. *Genome Res.* 7: 1111-1113.

- Meyers, B.C., Chin, D.B., Shen, K.A., Sivaramakrishnan, S., Lavelle, D.O., Zhang, Z. and Michelmore, R.W. 1998. The major resistance gene cluster in lettuce is highly duplicated and spans several megabases. *Plant Cell*, 10: 1817-1832.
- Nakamura, S., Asakawa, S., Ohmido, N., Fukui, K., Shimizu, N. and Kawasaki, S. 1997. Construction of an 800-kb contig in the near-centromeric region of the rice blast resistance gene *Pi-ta* (2) using a highly representative rice BAC library. *Mol. Gen. Genet.* 254: 611-620.
- Niederfuhr, A., Hummerich, H., Gawin, B., Boyle, S., Little, P.F.R. and Gessler, M. 1998. A sequence-ready 3-Mb PAC contig covering 16 breakpoints of the Wilms tumor aniridia region of human chromosome 11p13. *Genomics* 53: 155-163.
- Patocchi, A., Vinatzer, B.A., Gianfranceschi, L., Tartarini, S., Zhang, H.B., Sansavini, S. and Gessler, C. 1999. Construction of a 550 kb BAC contig spanning the genomic region containing the apple scab resistance gene *Vf*. *Mol. Gen. Genet.* 262: 884-891.
- Parisi, L. and Lespinasse, Y. 1996. Pathogenicity of *Venturia inaequalis* strains of race 6 on apple clones (*Malus sp.*). *Plant Dis.* 80: 1179-1183.
- Parisi, L., Lespinasse, Y., Guillaumes, J. and Krueger, J. 1993. A new race of *Venturia inaequalis* virulent to apples with resistance due to the *Vf* gene. *Phytopathology* 83: 533-537.
- Richter, T.E. and Ronald, P.C. 2000. The evolution of disease resistance genes. *Plant Mol. Biol.* 42: 195-204.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. 'Molecular Cloning: A Laboratory Manual' 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Schmidt, R., West, J., Love, K., Lenehan, Z., Lister, C., Thompson, H.Y., Bouchez, D. and Dean, C. 1995. Physical map and organization of *Arabidopsis thaliana* chromosome 4. *Science* 270: 480-483.
- Soderlund, C., Humphray, S., Dunham, A. and French, L. 2000. Contigs built with fingerprints, markers, and FPC V4.7. *Genome Res.* 10: 1772-1787.
- Stein, N., Feuillet, C., Wicker, T., Schlagenhauf, E. and Keller, B. 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.
- Tanksley, S.D., Ganai, M.W. and Martin, G.B. 1995. Chromosome landing: a paradigm for map-based gene cloning in plants with large genomes. *Trends Genet.* 11: 63-68.
- Tartarini, S. 1996. RAPD markers linked to the *Vf* gene for scab resistance in apple. *Theor. Appl. Genet.* 92: 803-810.
- Tartarini, S., Gianfranceschi, L., Sansavini, S. and Gessler, C. 1999. Development of reliable PCR markers for the selection of the *Vf* gene conferring scab resistance in apple. *Plant Breed.* 118: 183-186.
- Temnykh, S., Park, W.D., Ayres, N., Cartinhour, S., Hauck, N., Lipovich, L., Cho, Y.G., Ishii, T. and McCouch, S.R. 2000. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa L.*). *Theor. Appl. Genet.* 100: 697-712.
- Vollrath, D. and Jaramillo-Babb, 1999. A sequence-ready BAC clone contig of a 2.2-Mb segment of human chromosome 1q24. *Genome Res.* 9: 150-157.
- Wise, R.P. 2000. Disease resistance: what's brewing in barley genomics. *Plant Dis.* 84: 1160-1170.
- Xu, M.L. and Korban, S.S. 2000. Saturation mapping of the apple scab resistance gene *Vf* using AFLP markers. *Theor. Appl. Genet.* 101: 844-851.
- Xu, M.L., Huaracha, E. and Korban, S.S. 2001a. Development of sequence characterized amplified regions (SCARs) from AFLP markers tightly linked to the *Vf* gene in apple. *Genome* 44: 63-70.
- Xu, M.L., Song, J.Q., Cheng, Z.K., Jiang, J.M. and Korban, S.S. 2001b. A bacterial artificial chromosome (BAC) library of *Malus floribunda* 821 and contig construction for positional cloning of the apple scab resistance gene *Vf*. *Genome* 44: 1104-1113.
- Xu, M.L., Song, J.Q., Jiang, J.M. and Korban, S.S. 2001c. Constructing a bacterial artificial chromosome library of the apple cultivar GoldRush. *Acta Hort.* (in press).
- Yang, H. and Korban, S.S. 1996. Screening apples for OPD20/600 using sequence-specific primers. *Theor. Appl. Genet.* 92: 263-266.
- Yang, H., Korban, S.S., Kruger, J. and Schmidt, H. 1997a. A randomly amplified polymorphic DNA (RAPD) marker tightly linked to the scab-resistance gene *Vf* in apple. *J. Amer. Soc. Hort. Sci.* 122: 47-52.
- Yang, H., Korban, S.S., Kruger, J. and Schmidt, H. 1997b. The use of a modified bulk segregant analysis to identify a molecular marker linked to a scab resistance gene in apple. *Euphytica* 94: 175-182.
- Young, W.P., Schupp, J.M. and Keim, P. 1999. DNA methylation and AFLP marker distribution in the soybean genome. *Theor. Appl. Genet.* 99: 785-792.
- Zhang, H.B. and Wing, R.D. 1997. Physical mapping of the rice genome with BACs. *Plant Mol. Biol.* 35: 115-127.
- Zhu, H., Blackmon, B.P., Sasinowski, M. and Dean, R.A. 1999. Physical map and organization of chromosome 7 in the rice blast fungus, *Magnaporthe grisea*. *Genome Res.* 9: 739-750.